EFFICIENCY OF PRIMARY HEALTH CARE IN LOW-AND MIDDLE-INCOME COUNTRIES: CASE STUDIES FROM BANGLADESH

Thesis submitted to the University of London for examination for the degree of

Doctor of Philosophy

Damian G Walker

Health Policy Unit

Department of Public Health and Policy

London School of Hygiene and Tropical Medicine

2006

"What saves us is efficiency".

(Joseph Conrad, *Heart of Darkness, 1902*)

ABSTRACT

Most of the research concerned with the economics of health systems has focussed on allocative efficiency. Specifically, much effort has been devoted to the development and application of techniques of economic evaluation. The consideration of technical efficiency has figured less prominently in the search for 'solutions' to the problems of health systems. Those working on the economic evaluation of health care interventions have adopted the assumption that interventions are being, or will be, produced in a technically efficient manner.

The aim of this thesis is to challenge this assumption and illustrate the potential implications of assuming technical efficiency when allocating scarce resources. Two case studies from Bangladesh are presented: vaccination services in Dhaka City and primary health care in rural Bangladesh. The specific objectives of this thesis are to: estimate the cost of these services using standard costing methods; and analyse the same data sets using parametric (stochastic frontier analysis) and non-parametric (data envelopment analysis) techniques in order to identify whether, and to what degree, the services were being delivered efficiently.

Applying efficiency measurement techniques illustrated that standard costing methods disguise a high degree of inefficiency. By investigating production practices, costs related to inefficiencies can be identified and addressed. The thesis illustrates that if something is deemed worth doing then it should be carried out in a way which ensures the optimum use of scarce resources. An exclusive focus on switching resources from less cost-effective to more cost-effective activities will not realise the full benefits in terms of improved allocative efficiency if providers on the ground are not producing services at lowest cost. Recommendations are made for policy-makers on how technical efficiency can be improved. Recommendations for future research are also made.

STATEMENT

I formally declare that the work presented in this thesis is the result of my own work.

[redacted]

Damian G Walker

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF BOXES

LIST OF APPENDICES

LIST OF PUBLISHED PAPERS

See Chapter 1, section 1.4: Responsibility for completion of work included in the thesis.

An early version of Chapter 6 was published by Khan et al. (2004) and can be found in Appendix 1

- Khan MM, Khan SH, **Walker D**, Fox-Rushby J, Cutts F, Akramuzzaman SM (2004) Cost of delivering child immunization services in urban Bangladesh: a study based on facility-level surveys. *Journal of Health, Population and Nutrition* 22(4): 404-412

An early version of Chapter 7 was published by Valdmanis et al. (2003) and can be found in Appendix 2

- Valdmanis V, **Walker D**, Fox-Rushby JA (2003) Are vaccination sites in Bangladesh scale efficient? *International Journal of Technology Assessment in Health Care* 19(4): 692-697.

Another paper was published based on the data set used in Chapters 6 and 7 by Dervaux et al. (2003) and can be found in Appendix 3

- Dervaux B, Leleu H, Valdmanis V, **Walker D**. Parameters of control when facing stochastic demand: a DEA approach applied to Bangladeshi vaccination sites *International Journal of Health Care Finance and Economics* 2003; 3(4): 287-299

LIST OF CONFERENCE PRESENTATIONS GIVEN BASED ON THE THESIS

- **Walker D**, Valdmanis V, Fox-Rushby JA. Are vaccination sites in Bangladesh scale efficient? Paper presented at the meeting of the Health Economics Study Group, London, July 2002
- **Walker D**, Dervaux B, Leleu H, Valdmanis V. Parameters of control when facing stochastic demand: A DEA approach applied to Bangladeshi vaccination sites. Oral presentation at the 4th International Health Economics Association Conference, San Francisco, June 2003
- **Walker D**. Tackling the issue of technical efficiency: a critical review of the cost and cost-effectiveness of routine immunisation services in low- and middle-income

countries. Paper presented at the meeting of the Health Economics Study Group, Glasgow, July 2004

- **Walker D**. Accounting for technical inefficiencies in cost-effectiveness analysis. Poster presentation at the 5th European Conference on Health Economics, London, September 2004
- **Walker D**. The efficiency of primary health centres in rural Bangladesh. Poster presentation at the 5th European Conference on Health Economics, London, September 2004
- **Walker D**. The efficiency of health centres in rural Bangladesh. Oral presentation at the 5th International Health Economics Association Conference, Barcelona, July 2005

LIST OF ABBREVIATIONS

ANOVA	analysis of variance
BCG	Bacillus Calmette-Guérin
CBA	cost-benefit analysis
CEA	cost-effectiveness analysis
CMA	cost-minimisation analysis
COLS	correct least-squares
CRS	constant returns to scale
CUA	cost-utility analysis
DALY	disability-adjusted life year
DCC	Dhaka City Corporation
DEA	data envelopment analysis
DFID	Department for International Development
DPT	diphtheria, pertussis, tetanus
EPI	Expanded Programme on Immunization
ESP	essential services package
FP	family planning
FTE	full-time equivalent
FVC	fully vaccinated child
GCEA	generalized cost-effectiveness analysis
GNI	gross national income
GoB	Government of Bangladesh
HNPSP	Health Nutrition and Population Sector Programme
HPSP	Health and Population Sector Programme
ICDDR,B	ICDDR,B: Centre for Health and Population Research
IRS	increasing returns to scale
LR	likelihood ratio
LSHTM	London School of Hygiene and Tropical Medicine
MCH	maternal and child health
ML	maximum likelihood
MOHFW	Ministry of Health and Family Welfare
NGO	non-government organisation
NICE	National Institute for Clinical Excellence
NIRS	non-increasing returns to scale

OLS	ordinary least squares
OPV	oral polio vaccine
SFA	stochastic frontier analysis
TT	tetanus toxoid
UHC	upazila health complex
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
VRS	variable returns to scale
WHO	World Health Organization

ACKNOWLEDGEMENTS

Many people have contributed in important ways to the work in this thesis and to supporting me during its development. First and foremost, I would like to thank my supervisor, Professor John Cairns. John became my supervisor in May 2004, at a crucial time in my thesis and provided constant encouragement and support. In addition, I would like to thank Dr Julia Fox-Rushby and Dr Vivian Valdmanis, who were my co-supervisors for this PhD before they went to Brunel University and the University of the Sciences in Philadelphia respectively. In between supervisors, Professor Anne Mills provided guidance and support, for which I am extremely grateful.

Members of my advisory committee were Professor Charles Normand and Dr Colin Sanderson, who provided support at the upgrading stage of my thesis. Many other colleagues at the London School of Hygiene and Tropical Medicine read and commented on sections of the thesis or provided valuable intellectual and moral support. These include Stephen Jan, Warren Stevens, Virginia Wiseman, Justin Parkhurst, Lorna Guinness, Lesong Conteh, Paul Coleman, Richard Grieve, Catherine Goodman, John Browne and Kara Hanson.

In Bangladesh, I was also supported by extremely kind colleagues. In particular I thank, Shahadat Hossain, Mahmud Khan, Suhaila Khan, Nazme Sabina and Shakil Ahmed, at ICDDR,B: Centre for Health and Population Research. A large number of individuals made important contributions through granting access to their organisations and data. I am grateful for the patience, effort and hospitality I received from the staff of the health facilities.

Financial support for the projects from which the data were collected was received from the World Health Organization and the UK Department for International Development.

I would also like to thank my friends and family, in particular my Mum and Dad, for always being there for me. It would have much harder to complete this thesis without their moral and practical support.

Chapter 1

INTRODUCTION

This thesis explores whether, and to what degree, health care is delivered efficiently. It compares and contrasts different efficiency measurement techniques, and applies them to the delivery of primary health care in urban and rural Bangladesh. It considers the impact an assumption of technical efficiency may have on the interpretation of cost-effectiveness ratios. And finally, it provides recommendations to policy-makers before presenting an agenda for future research.

Resources are scarcest in low- and middle-income countries such as Bangladesh, thus their inefficient use exacts a much higher penalty in terms of foregone health benefits in these settings than it does in high-income countries. It is essential that resources are used as efficiently as possible.

1.1 Introduction

In recent years, most of the research concerned with the economics of health systems in low- and middle-income countries has focussed on aspects of allocative efficiency. Specifically, a desire to enhance the allocative efficiency of health systems has led to much effort being devoted to the development and application of the techniques of economic evaluation, and in particular cost-effectiveness analysis[1] (CEA), which aim to allow comparison among alternative health interventions (e.g. Jamison et al. 1993; Jamison et al. 2006). The consideration of technical efficiency, however, has figured

[1] It should be noted that in practice, there has been a blurring of the distinctions between CEA and cost-utility analysis (CUA), with the latter seen as an extension of the former (Musgrove 2000). Hence, the cost-effectiveness literature often encompasses both these approaches. Indeed, use of the term CEA in this thesis will encompass both approaches.

less prominently in the search for 'solutions' to the problems of health systems in low- and middle-income countries. In the best traditions of economics, those working on the economic evaluation of health care interventions have tended to adopt the assumption that the interventions they are examining are being, or will be, produced in a technically efficient manner (Hensher 2001).

1.2 Case studies, and thesis aim and objectives

The overall aim of the thesis is to contribute to the methodological development of cost, and more broadly cost-effectiveness, analysis of health care programmes by exploring whether, and to what degree, health care is delivered in a technically efficient manner.

Two case studies were chosen from projects described below. The first uses data from the delivery of vaccination services in Dhaka City. Data were collected from a sample of 132 vaccination delivery units. The second case study uses data collected from 36 health centres in rural Bangladesh. More details on these case studies are provided in Chapters 5, 6 and 8. It should be noted that while immunisation is an integral component of primary health care in Bangladesh, it is not routinely delivered through rural health centres in Bangladesh (although they are used on a regular basis as outreach sites). Therefore, although the case studies both consider aspects of primary health care, they unfortunately do not cover the same activities[2]. However, the decision to include both case studies was largely influenced by the candidate's upgrading committee meeting in March 2003. The committee urged the candidate to supplement his early collaborative work on the efficiency of vaccination services in Dhaka (see below) to

[2] Data were available from nine sub-district hospitals that are responsible for organising and delivering vaccination services in rural Bangladesh (see Chapter 5), collected as part of project funded by the UK Department for International Development (DFID) described below. However, for reasons described in Chapter 3, nine facilities is an insufficient sample size for the purpose of this thesis.

ensure the work included in this thesis was substantially his own. However, as is described in more detail in the section 1.4 below, "Responsibility for completion of work included in the thesis", that earlier collaborative work has also been significantly revised by the candidate.

Using these two case studies from Bangladesh, the specific objectives are to:

1. Describe the empirical evidence on the efficiency of health care programmes in low- and middle-income countries and regions;
2. Estimate the cost of delivering vaccination services among a sample of vaccination delivery units in Dhaka City;
3. Estimate the cost of delivering primary health care among a sample of health centres in rural Bangladesh;
4. Estimate the efficiency of delivering these services using data envelopment analysis and stochastic frontier analysis;
5. Describe the variation in efficiency among the units and to explore some of the causes of this variation;
6. On the basis of these findings, describe the potential implications of inefficiency in the delivery of health care programmes;
7. On the basis of these findings, make recommendations on how policy-makers in Bangladesh and elsewhere could improve efficiency, and make recommendations on further research relevant to health care efficiency issues.

The thesis draws from two independent, yet related, projects of which the candidate was a co-investigator. The first project's aim was to estimate the cost-effectiveness of the measles vaccine of the national immunisation programme in Dhaka, Bangladesh. The data were collected in 1999, and analysed and written up in 2000 (Walker et al. 2000). The study was funded by the World Health Organization (WHO) (grant number HQ/98/454419 011638) with contributions to salaries and fieldwork costs from ICDDR,B: Centre for Health and Population Research (ICDDR,B) and the London School of Hygiene and Tropical Medicine (LSHTM). The second project's aim was to evaluate the cost-effectiveness of introducing vaccines against hepatitis b, *Haemophilus influenzae* type b and rotavirus versus the *status quo* of the current programme at existing, and higher, coverage rates in Bangladesh and Peru. The data were collected during 2002-2003, and analysed and written up during 2004-2005. The study was funded by DFID (grant number R7842) with contributions to salaries and fieldwork costs from WHO, ICDDR,B and LSHTM.

Neither of these studies had as an objective the application of parametric and non-parametric efficiency measurement techniques, such as stochastic frontier analysis (SFA) and data envelopment analysis (DEA). Rather, they sought to identify the cost-effectiveness of different vaccination programmes. The implications of this are discussed in Chapter 10. However, it is important to note that the aim of this thesis is not to examine in detail the methodological underpinnings of parametric and non-parametric efficiency measurement techniques. Rather, it is to use these techniques to critique the underlying assumptions of technical and scale efficiency in economic evaluation. Of course, this does not mean that such techniques are not without their own problems and Chapter 3 provides a summary of the main criticisms. However, the focus of this thesis is on using these techniques to critique economic evaluation rather than *vice versa*.

1.3 Contribution of thesis

It is anticipated that this thesis will add to the cost-effectiveness literature by providing an insight into the importance of failing to consider technical and scale efficiency. This study will assess the extent to which incorporating technical efficiency considerations can potentially alter the decision on whether or not to adopt a new technology and / or expand an existing technology. This should provide decision-makers with a clearer indication of the extent to which results generated in one setting are transferable between settings, transferable with adjustment or not transferable at all.

1.4 Responsibility for completion of work included in this thesis

All the work included in this thesis derives from the two studies described above and the candidate recognises the contributions of all who were employed to work on the projects.

Early versions of Chapters 6 and 7 have been published elsewhere by Khan et al. (2004)[3] and Valdmanis et al. (2003)[4]. In terms of the candidate's contributions to each of the papers, they were as follows. For the Khan et al. (2004) paper, the candidate:

- wrote the original proposal from which the paper derives[5];
- designed the data collection tools;
- supervised data collection (Suhaila Khan supervised data collection, entry and cleaning locally);
- performed the analysis in collaboration with Suhaila Khan;

---

[3] Khan MM, Khan SH, Walker D, Fox-Rushby J, Cutts F, Akramuzzaman SM (2004) Cost of delivering child immunization services in urban Bangladesh: a study based on facility-level surveys. *Journal of Health, Population and Nutrition* 22(4): 404-412. See Appendix 1 for a pdf copy of this paper.
[4] Valdmanis V, Walker D, Fox-Rushby JA (2003) Are vaccination sites in Bangladesh scale efficient? *International Journal of Technology Assessment in Health Care* 19(4): 692-697. See Appendix 2 for a pdf copy of this paper.
[5] WHO grant number HQ/98/454419 011638

- drafted the paper in collaboration with Suhaila Khan.

Felicity Cutts and Julia Fox-Rushby were principal investigators at LSHTM, and Mahmud Khan and Syed Md. Akramuzzaman were principal investigators at ICDDR,B, and they reviewed the article critically for final approval. All authors responded to the referees' comments.

Chapter 6 represents a substantially revised version of the paper. In particular, the candidate re-cleaned the data set, resulting in the loss of 22 vaccination delivery units due to missing and / or questionable data. Examiners of this thesis may wish to compare Chapter 6 with the Khan et al. (2004) article which can be found in Appendix 1.

For the Valdmanis et al. (2003) article, the candidate[6]:

- contributed to the conceptualisation of the problem statement along with Vivian Valdmanis;
- contributed substantive knowledge regarding the delivery of vaccination services in Dhaka, which enabled Vivian Valdmanis to design the DEA model;
- contributed to the interpretation of the results;
- contributed to the drafting of the article.

All authors reviewed the article critically for final approval, and all authors responded to the referees.

[6] Note, this paper derives from the same original proposal (Walker et al. 2000). Thus many of the points raised above with respect to the Khan et al. (2004) paper apply here. The list here focuses on points of relevance to the secondary analysis performed of the same data using DEA.

Chapter 7 represents a significantly revised version of the paper. In particular, the candidate used the re-cleaned data set, resulting in the loss of seven vaccination delivery units due to missing and / or questionable data. Furthermore, the candidate included a further two DEA model specifications, and introduced the SFAs. Therefore, the methods, results and discussion sections are substantially revised. In essence, Chapter 6 bears little resemblance to the paper Valdmanis et al. (2003). The examiners of this thesis are welcome to compare Chapter 7 with the Valdmanis et al. (2003) article which can be found in Appendix 2.

With respect to Chapters 8 and 9, the candidate:

- wrote the original proposal from which the paper derives[7];
- designed the data collection tools;
- supervised data collection, entry and cleaning;
- performed all analyses.

Colin Sanderson and Julia Fox-Rushby were principal investigators of the project at LSHTM, and Shahadat Hossain, Nazme Sabina were principal investigators at ICDDR,B. The candidate gratefully acknowledges the contributions these individuals made during the data collection period, particularly Shahadat Hossain and Nazme Sabina. While the analyses presented in Chapters 8 and 9 are thus based on data collected through a joint project, they fell outside of the project's aims and objectives, and the candidate is thus fully responsible for what is presented herewith.

[7] DFID grant number R7842.

1.5 Outline of thesis

Chapter 2 provides an overview of the concepts of efficiency and economic evaluation. It describes the role of CEA in the health sector. The chapter concludes by discussing the assumption of technical efficiency underlying CEA, and begins to consider the impact these assumptions may have on the interpretation of cost-effectiveness ratios, and thus how decisions of how to allocate resources.

Chapter 3 describes SFA and DEA, the main parametric and non-parametric efficiency measurement techniques. However, it begins with a brief introduction to the efficiency concepts developed by Farrell (1957). This chapter also compares and contrasts the strengths and weaknesses of the two approaches. The final section discusses some methodological challenges of measuring efficiency in the health sector, e.g. adjusting for case mix, allowing for variation in technical quality and knowledge of input prices.

The literature review presented in Chapter 4, examines the evidence-base on the efficiency of health care services in low- and middle-income countries. It identifies the range of methods used, models specified, results and recommendations. The literature review identifies a number of key gaps and unanswered questions concerning the measurement of efficiency in low- and middle-income settings.

Chapter 5 provides background and context in which the two case studies examined in this thesis are operating in Bangladesh. It provides an overview of general health status indicators, a description of the Ministry of Health and Family Welfare, which has overall responsibility for health sector policy and planning in Bangladesh, a summary of health care services, both government and non-government, in urban and rural

Bangladesh, alongside an examination of health care expenditure. It concludes by summarising the health policy and planning framework in place in Bangladesh and gives a description of recent health sector reform programmes.

Chapters 6 and 8 estimate the costs of delivering vaccination services and primary health care in urban and rural Bangladesh respectively, using standard costing methods. The chapters describe the variation in unit costs observed among 110 vaccination delivery units and 34 health centres respectively. The same data are examined by using DEA and SFA in chapters 7 and 9. The different techniques are compared and contrasted in order to assess the stability of the findings. In addition, analyses are performed to identify whether selected environmental variables explain some of the variation in efficiency observed in the sample data. These chapters also discuss some of the policy implications of the findings, focussing in particular on the potential savings were technical efficiency improved. And in order to guide mangers of these services, it also provides targets for efficiency improvements. Some suggestions are provided on how these targets might be met.

Chapter 10 is divided into two main sections. The first section discusses methodological issues, in particular limitations of the data, analysis and interpretation. The second section discusses the main findings of the thesis. This section focuses in particular on the implications of the findings on the practice of economic evaluation.

The concluding chapter reflects on what has been presented in the preceding ten chapters and draws lessons from the theoretical and empirical information. It discusses the generalisability of the findings within and beyond Bangladesh. It makes recommendations on how policy-makers in Bangladesh and elsewhere could best approach the issue of inefficiency within the health sector. Finally, areas for future research are outlined.

Chapter 2

EFFICIENCY AND ECONOMIC EVALUATION

By way of background, the first section provides an overview of the concepts of efficiency[8]. The second section describes the role of CEA in the health sector. The third section describes the assumption of technical efficiency underlying CEA, drawing upon a selective review of key methodological guidelines. The fourth section concludes the chapter.

2.1 Concepts of efficiency

It is widely agreed that, given the scarcity of health care resources, it is important that services be produced efficiently. However, it is not always clear what is meant by efficient. Economists use a number of concepts of efficiency[9]. Thus, as Culyer (1992) states, "The term 'efficient' ... needs unpacking, since much confusion about what it is abounds". The basic premise underlying the concept of efficiency is that no output can be produced without resources (inputs) and that these resources are limited in supply. From this, it follows that there is a limit to the volume of output that can be produced. At the most basic level, there is a desire to ensure that the existing inputs are not capable of producing more services. Therefore all definitions of efficiency basically follow from avoidance of waste. The presence of waste obviously implies some persons could be made better off without using more resources. The two main concepts to consider

[8] Methods for measuring efficiency are reviewed in Chapter 3.
[9] A summary of arguments concerning economists' general confusion about efficiency can be found in the Reinhardt (2003).

are technical and allocative efficiency[10]. It will be illustrated that the definitions are applied in different ways depending on whether the unit of analysis is a firm[11] or health system.

2.1.1 Concepts of efficiency applied to firms

2.1.1.1 Technical efficiency

In order to measure efficiency, a norm must be specified. The norm set for measuring technical efficiency is that the minimum amount of resources should be used to produce a given level of output or, alternatively, the maximum amount of output should be produced for a given level of resource use. If more resources than necessary are used to produce a given amount of output, this implies a waste of resources and therefore inefficiency. Thus, the difference in the amount of output that could have been produced from a given amount of resources and the amount of output that was actually produced can be used as a measure of technical inefficiency. Technical inefficiency is therefore a matter of degree depending upon how much unnecessary resources have been used. Central to the measurement of technical efficiency is the notion of the isoquant[12].

This is illustrated in Figure 1 for a simple production process that uses only two inputs, X_a and X_b (for example, these inputs could be doctor- and nurse-hours worked, or doctor-hours and drugs). Any point along the isoquant QQ represents a technically efficient way of combining various quantities of inputs X_a and X_b to produce the same

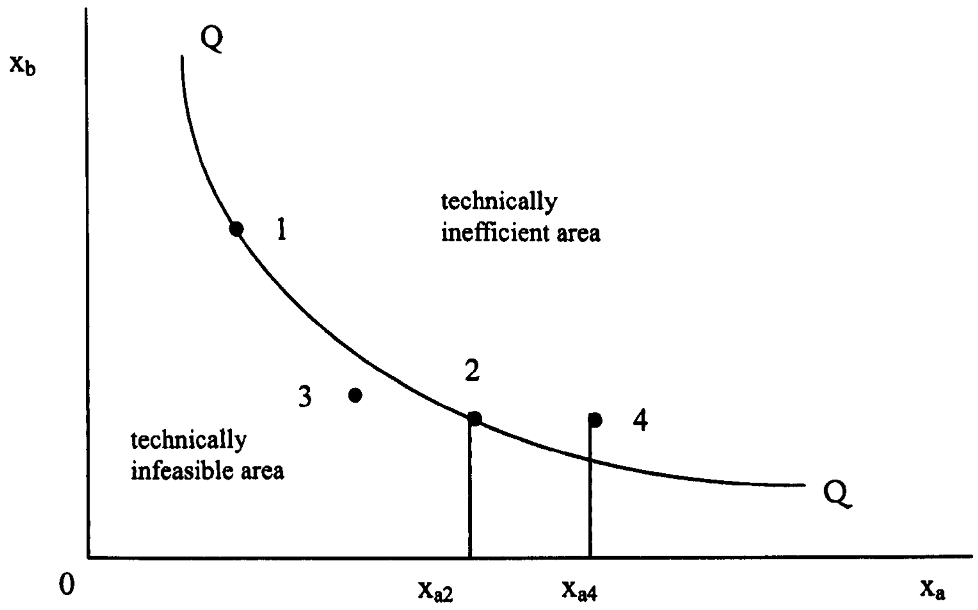
[10] Some textbooks use the terms productive or operational efficiency instead of technical efficiency and the term price or economic efficiency instead of allocative efficiency. The terminology used in this thesis conforms to that used most often in recent production economics literature (e.g. Färe et al. 1994).

[11] The term decision-making unit is sometimes used to describe a productive entity in instances when the term 'firm' may not be entirely appropriate, e.g. when comparing the performance of public vaccination sites, the units are really parts of a firm rather than firms themselves.

[12] An isoquant represents all the possible combinations of inputs, which permit production of the same quantity of health care output – *iso* meaning 'same' and *quant* meaning 'quantity'. The output counterpart to the isoquant is the production possibilities frontier, which depicts the various combinations of inputs that could be used to produce a given level of output.

amount of output Q. For example, while points 1 and 2 differ in the combination of X_a and X_b (production at 1 is more intensive in X_b than at 2), both permit production of the same quantity Q. Points 1 and 2, like all other points on the isoquant QQ are technically efficient because it is not possible to produce Q with smaller quantities of either X_a or X_b, as depicted by the line (there is no room for further gain in technical efficiency). Point 3, like all points to the left of the isoquant, is infeasible, i.e. any reduction in the amounts of X_a and X_b from the amounts represented by the isoquant necessarily translates into a reduction in Q. In contrast, point 4, like all points to the right of the isoquant, constitutes a technically inefficient way of producing Q, i.e. technical efficiency can be improved by moving production from 4 to 2, thereby reducing the amount of X_a from X_{a4} to X_{a2}. In effect, one *modus operandi* is considered more technically efficient than another, if it either produces the same quantity of output using fewer inputs, or produces a greater quantity of outputs using the same resources.

Figure 1: An example of technical efficiency



It is important to note, that it is assumed here that technical quality of care also remains constant along the isoquant. Thus, not only does any combination of inputs X_a and X_b

along the curve permit production of quantity Q of medical care output, but also, any such combination delivers medical care of constant technical quality, i.e., with the same effect on patients' health status (see Chapter 3 for a discussion of the implications of this assumption).

2.1.1.2 Allocative efficiency

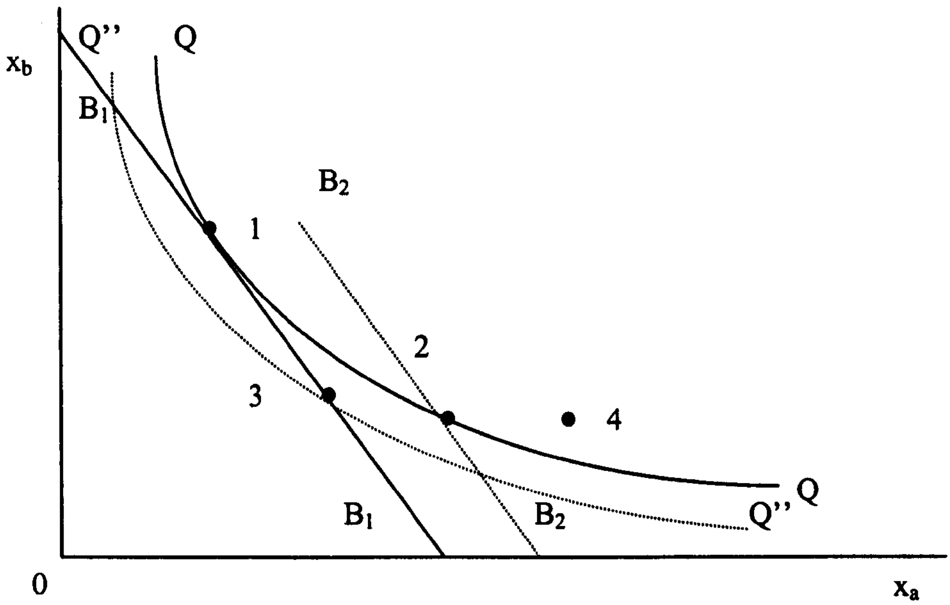
An allocatively efficient firm will combine these inputs in a cost-minimising manner to produce a given level of output, with price ratios being the norms for judging allocative efficiency[13]. With factor input prices given, resources used in production should be combined so as to reflect the corresponding ratio of different factor input prices. A mix of resource use that deviates from the corresponding ratio of given factor input prices is taken as a measure of allocative inefficiency. Any deviation in the mix of resource use from observed price ratios is measurable, and hence, allocative inefficiency becomes a matter of degree, just like technical inefficiency. Although there may be many technically efficient alternatives to produce a given quantity Q, there is only one allocatively efficient way of doing so.

Figure 2 helps to illustrate the fundamental difference and relationship between technical and allocative efficiency. Suppose that the unit prices of inputs X_a and X_b are W_a and W_b respectively. If a health facility is allocated a budget B_1, then B_1 represents the facility's budget constraint. The constraint is given by the equation: $B_1 = (X_a * W_a) + (X_b * W_b)$. Any point along the budget constraint line, such as points 1 and 3, consumes the whole budget B_1. However, point 1 is preferable to 3 because at 1

[13] However, using prices as the criteria for measuring economic efficiency is based on the assumption that firms have no influence on the price. Rather, prices are determined in the market as the outcome of competitive bidding between a large number of consumers and firms. Clearly this may not apply in the health sector.

quantity Q is produced, whereas at point 3 the smaller quantity Q" is produced. Furthermore, of all the technically efficient points along the frontier QQ, point 1 is the most allocatively efficient way of producing quantity Q. Point 2 is as technically efficient as 1, but is less allocatively efficient, since production at 2 requires a budget of B_2, higher than B_1. Graphically, the allocatively efficient point (point 1) corresponds to the tangency between the budget constraint and the isoquant. Thus, technical efficiency is a necessary but not sufficient condition for allocative efficiency. In general, two types of circumstances, discussed above, can lead to allocative inefficiency: technical inefficiency and technically efficient production that uses a mix of inputs that is not cost minimising.[14]

Figure 2: An example of allocative efficiency



Finally, when taken together, technical efficiency and allocative efficiency determine the degree of economic, or overall, efficiency. Thus, if a firm uses its resources in a technically and allocatively efficient way, then it can be said to have achieved economic

[14] There is a third cause of economic inefficiency, referred to as social economic inefficiency that can arise when the input prices faced by facility managers (e.g. personnel wages or pharmaceutical products) depart from social (or shadow) prices.

efficiency. Alternatively, to the extent that either technical or allocative inefficiency is present, the firm will be operating at less than total economic efficiency.

2.1.2 Concepts of efficiency applied to health systems

As described above, allocative efficiency is traditionally used to describe the optimal, i.e. cost-minimising, mix of inputs to a production process given their respective prices. However, as interventions are inputs to the production of health, allocative efficiency can be viewed as choosing the optimal mix of interventions for any given level of expenditure (Tan-Torres Edejer et al. 2003). Technical efficiency is thus viewed as minimising the cost of delivering an intervention, referred to above as allocative efficiency, which perhaps illustrates a cause of some of the confusion among economists noted by Reinhardt (2003). As stated previously, technical efficiency traditionally describes a situation where the minimum quantity of inputs is used to produce a given level of output, or conversely, the maximum quantity of output is produced given available inputs, i.e. the cost of producing these levels of output might not be minimised (see Figure 2). However, it is implicitly assumed when viewing technical efficiency as minimising the cost of delivering an intervention that this traditional definition has been met – indeed, the achievement of allocative efficiency in the traditional sense requires technical efficiency to be met. Given that CEA generally deals with interventions as the unit of analysis rather than a health facility or health system, it is perhaps not surprising that the terms cost-effectiveness and technical efficiency have been viewed as synonymous by some economists. Thus, as Tan-Torres Edejer et al. (2003) state

> "… inefficiencies in the production of health may derive from two sources: problems with technical efficiency – how an intervention is delivered – and problems with allocative inefficiency – which set of interventions is provided".

2.2 Role of cost-effectiveness analysis in the health sector

CEA is a form of economic evaluation. The different forms of economic evaluation are cost-minimisation analysis (CMA), CUA and cost-benefit analysis (CBA). The basic task of any economic evaluation is to identify, measure, value and compare, the costs and consequences of the alternatives being considered (Drummond et al. 1997). The various evaluation techniques estimate costs in a similar fashion, but differ in the measurement of health outcomes (see Box 1). Costs refer to the value of opportunities, or benefits foregone, from not employing resources elsewhere. Benefits are gauged by the consequences of a health programme on people's well-being or health status. The different ways of measuring benefits lead to a trade-off between the scope for potential use and the practicality of various evaluation techniques.

As illustrated above, when the unit of analysis is an intervention, as it generally is when performing an economic evaluation, allocative efficiency can be viewed as choosing the optimal mix of interventions for a given level of expenditure – optimal in the sense that they maximise health gain. In this broader definition of efficiency, different health care interventions with different objectives and outcomes must be compared, e.g. malaria versus tuberculosis control, or more generally, how should the Ministry of Health's budget be distributed between programmes? It thus follows that, while interventions may have different objectives and outcomes of interest, these must be converted into commensurable units if the optimal mix is to be defined. For this reason, CUA, which uses more complex measures of outcomes, can be used to assess allocative efficiency within the health sector. However, this form of economic evaluation is still restricted to comparisons of programmes within the health sector, so strictly speaking only deals with *quasi*-allocative assessments.

Box 1: Different types of economic evaluation

Cost-minimisation analysis: compares two or more interventions that have identical outcomes (e.g. number of cases treated) are assessed to see which provides the cheapest way of delivering the same outcome.

Cost-effectiveness analysis: measures the outcome of interventions in terms of 'natural units' e.g. for national immunisation programmes, this could be the number of disease cases averted.

Cost-utility analysis: these evaluations use a measure of utility reflecting people's preferences. The outcomes are then expressed in terms of measures such as quality-adjusted life-years or disability-adjusted life-years (DALYs).

Cost-benefit analysis: expresses outcomes (e.g. the number of lives saved) in terms of monetary units.

In theory, CBA has the widest scope of the four types of analysis because outcomes are monetised enabling inter-sectoral comparisons, i.e. in principle it can address how a government budget should be distributed between different ministries. In practice however, the valuation of health benefits is difficult and thus preference for CEA over other types of analysis for evaluating health care programmes has emerged since the late 1970s in both developed and developing countries (Elixhauser et al. 93; Elixhauser et al. 98; Stone et al. 2005; Walker and Fox-Rushby, 2000; Warner and Hutton 1980).

On the other hand, all of the different types of economic evaluation can be used to assess technical efficiency, which, can be viewed as maximising the achievement of a given objective within a given budget when the unit of analysis is an intervention[15], e.g. vaccination of children through fixed, outreach or mobile clinics.

As noted above, a perceived strength of CEA is that it can help identify technically efficient alternatives. Thus the next section describes in more detail the micro-economic assumption of technical efficiency underlying economic evaluation. In doing so, it reviews the extent to which a selection of influential economic evaluation guidelines provide appropriate direction to analysts for identifying technically efficient interventions, and by definition, allocatively efficient health systems.

2.3 The assumption of technical efficiency in economic evaluation

It has been noted that CEA guidelines fail to explicitly consider the concept of technical efficiency (Donaldson et al. 2002), with perhaps the notable exception of the recent publication of the World Health Organization's generalised cost-effectiveness analysis (GCEA) guidelines, which conversely, explicitly states that the GCEA guidelines do not consider technical efficiency:

> The main objective of this type of economic evaluation [GCEA] is to provide policy makers with information on the relative cost-effectiveness of a given set of interventions. Thus it addresses issues of allocative efficiency of scarce health care resources. *Technical efficiency is assumed in this type of analysis.* (Italics mine) (Baltussen et al. 2002)

Worryingly, the authors wrote this after stating that, "It is not useful for policy makers to know the cost-effectiveness of inefficient interventions" (Baltussen et al. 2002). However, this could have something to do with the fact that the terms cost-effectiveness

[15] Or alternatively, the ability to produce a given output at the lowest possible cost.

and technical efficiency have been used interchangeably by some economists, again suggesting that technical efficiency is assumed rather than assessed (Donaldson 1990).

As noted above, the 'cost' in cost-effectiveness refers to the value of opportunities, or benefits foregone, from not employing resources elsewhere. This requires that costs reflect overall, or economic, efficiency. Thus, if some of the resources used do not contribute to the improvement of health outcomes, these resources should be identified and excluded in the costing of the health care programme; including these costs would mean that the costs reported no longer reflect opportunity costs. Unfortunately, National Institute for Clinical Excellence (NICE) (2001) guidelines offer little advice on how analysts should measure opportunity costs. As Birch and Gafni (2002) point out when offering a critique of previous NICE guidelines:

> "...a general problem that underlies many aspects of the guidelines relates to the limited attention given to the concept of opportunity cost ... the solution to the problem of using market prices that do not reflect opportunity costs is to use other data which also do not reflect opportunity costs ...".

Rather than provide a more appropriate definition of opportunity costs, the most recent guidelines from NICE state that they prefer unit costs to reflect the financial costs to the National Health Service and Personal Social Services, rather than the opportunity costs. If the aim of economic evaluations is to move resource use towards technical and allocative efficiency, opportunity costs are required. Thus the NICE position appears to be inconsistent with economic theory. Using financial costs in economic evaluations and decision-making may lead to inefficient resource allocation.

If costs vary because some centres are inefficient whereas others are relatively efficient, then it is inappropriate to use costs that are representative of all the centres concerned. Some centres may be producing the health care programmes concerned in an efficient manner, whereas in other centres, resources may be wasted. If the costs included are to represent opportunity costs then the costs from the efficient centres are those that are relevant.

An economic evaluation should ideally be able to recognise departures from allocative efficiency. However, there is little evidence to suggest that the guiding principles for economic evaluation consider technical efficiency. If the costs used do not represent opportunity costs, because they incorporate inefficiency in the provision of health care programmes, then the study may produce misleading estimates of the relative cost-effectiveness of each health care programme. The issue of how to ensure that the costs used in economic evaluations approximate opportunity costs is not addressed by many of the guidelines including the recent methodological guidance issued by NICE and WHO (NICE 2001; Tan-Torres Edejer 2003).

2.4 Summary

- Technical efficiency traditionally describes when the minimum quantity of inputs is used to produce a given level of output, or conversely, the maximum quantity of output is produced given available inputs;
- Allocative efficiency is traditionally used to describe the optimal mix of inputs to a production process given their respective prices;
- However, as interventions are inputs to the production of health, allocative efficiency can be viewed as choosing the optimal mix of interventions for any given

level of expenditure, and technical efficiency can thus be viewed as minimising the cost of delivering an intervention;

- However, it is implicitly assumed when viewing technical efficiency as minimising the cost of delivering an intervention that this traditional definition has been met;
- CEA is a form of economic evaluation. The other forms of economic evaluation are CMA, CUA and CBA. The basic task of any economic evaluation is to identify, measure, value and compare, the costs and consequences of the alternatives being considered;
- CEA fails to explicitly consider technical efficiency, assuming instead that the cost of providing a particular level of service is minimised;
- Failing to account for differing levels of technical, and therefore by definition allocative, efficiency among providers or health systems in different countries could have significant implications for the validity of the results of CEAs;
- This thesis will use parametric and non-parametric efficiency measurement technqiues, described in the next chapter, to challenge this assumption and to explore the implications of it.

The next chapter presents DEA and SFA, two alternative approaches for measuring efficiency.

Chapter 3

ALTERNATIVE METHODS FOR MEASURING EFFICIENCY IN THE HEALTH SECTOR

This chapter presents the methods of two alternative approaches for measuring efficiency; DEA and SFA. However, it begins with a brief introduction to the efficiency concepts developed by Farrell (1957).[16] The primary purpose of this first section is to outline a number of commonly used efficiency measures and to discuss how they can be calculated relative to a given technology, which is generally represented by some form of frontier function. The second section introduces the basic DEA models, namely the constant returns to scale (CRS) and variable returns to scale (VRS) models. This section also considers how allowance for environmental variables can be made. The third section provides an overview of stochastic frontier modelling, and again considers how allowance for environmental variables can be made. In addition, this section describes the maximum likelihood estimation procedure used in SFA, reviews alternative functional forms and presents methods for hypothesis testing. The fourth section considers appropriate sample sizes and the dimensionality issue, which are aspects relevant to both approaches. The fifth section compares and contrasts the strengths and weaknesses of the two approaches. The final section discusses some methodological challenges of measuring efficiency in the health sector, e.g. adjusting for case mix, allowing for variation in technical quality and knowledge of input prices.

[16] A more detailed treatment is provided by Lovell (1993) and Färe et al. (1994).

3.1 Introduction

The consideration of efficiency measurement to a large extent began with Farrell (1957), who drew upon the work of Debreu (1951) and Koopmans (1951) to define a simple measure of firm efficiency which could account for multiple inputs. Farrell (1957) proposed that the efficiency of a firm consists of two components: technical efficiency, which reflects the ability of a firm to obtain maximal output from a given set of inputs, and allocative efficiency, which reflects the ability of a firm to use the inputs in optimal proportions given their respective prices and the production technology. These two measures are then combined to provide a measure of total economic efficiency[17].

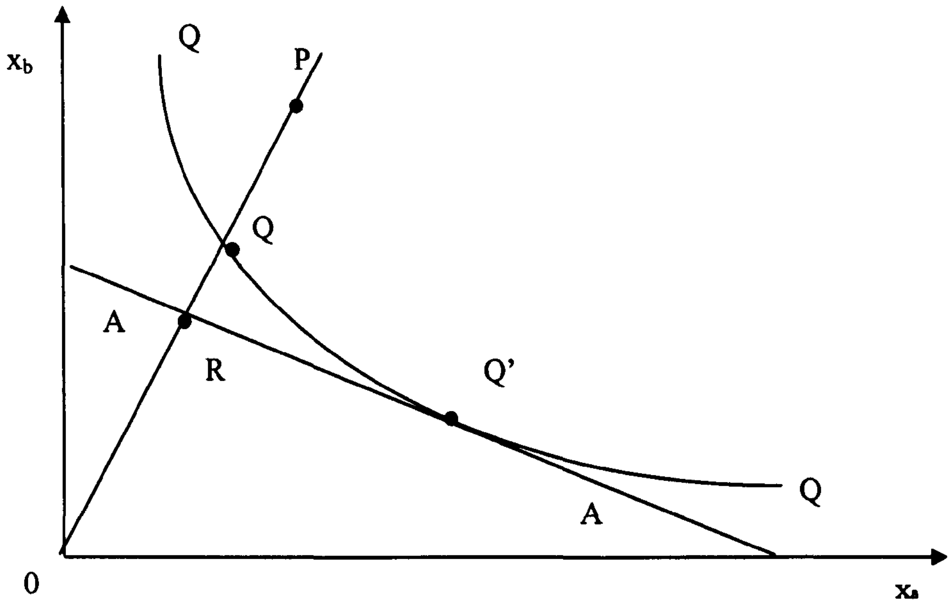
3.1.1 Input-orientated measures

Farrell illustrated his ideas using a simple example involving firms which use two inputs (x_a and x_b) to produce a single output (y), under the assumption of CRS (which allows the technology to be represented using unit isoquants[18]). Knowledge of the unit isoquants of fully efficient firms[19], represented by QQ in Figure 3, permits the measurement of technical efficiency. If a given firm uses quantities of inputs, defined by the point P, to produce a unit of output, the technical inefficiency of that firm could be represented by the distance QP, which is the amount by which all inputs could be proportionally reduced without a reduction in output. This is usually expressed in percentage terms by the ratio QP / 0P, which represents the percentage by which all inputs need to be reduced to achieve technically efficient production. The technical

[17] Farrell used the term price efficiency instead of allocative efficiency, and the term overall efficiency instead of economic efficiency.
[18] A firm is said to exhibit CRS if an increase in the proportion of inputs by one unit will result in a one unit increase in the proportion of outputs.
[19] This is not known in practice and thus must be estimated from observations on a sample of firms in the industry under consideration.

efficiency (TE) of a firm is most commonly measured by the ratio 0Q / 0P, which is equal to one minus QP / 0P. It will take a value between zero and one, and hence provide an indicator of the degree of technical efficiency of the firm. A value of one indicates the firm is fully technically efficient. For example, the point Q is technically efficient because it lies on the isoquant.

Figure 3: Technical and allocative efficiency



If the input price ratio, represented by the slope of the isocost line, AA, in Figure 3, is also known, allocative efficiency can also be calculated. The allocative efficiency (AE) of the firm operating at P is defined as the ratio AE = 0R / 0Q, because the distance RQ represents the reduction in production costs that would occur if production were to occur at the allocatively (and technically) efficient point Q', instead of at the technically efficient, but allocatively inefficient, point Q.

The total economic efficiency (EE) is defined as the ratio EE = 0R / 0P, where the distance RP can also be interpreted in terms of a cost reduction. The product of technical and allocative efficiency provides the measure of overall or economic efficiency:

$$EE = TE \times AE = (0Q / 0P) \times (0R / 0Q) = 0R / 0P$$

Note that all three measures are bounded by zero and one.

These efficiency measures assume that the production function is known. In practice this is not the case, and the efficient isoquants must be estimated from sample data. Farrell (1957) suggested the use of either a non-parametric piece-wise-linear convex isoquant, constructed such that no observed point lies to the left or below it (e.g. DEA), or a parametric function, such as the Cobb-Douglas[20] form, fitted to the data (e.g. SFA), again such that no observed point lies to the left or below it.

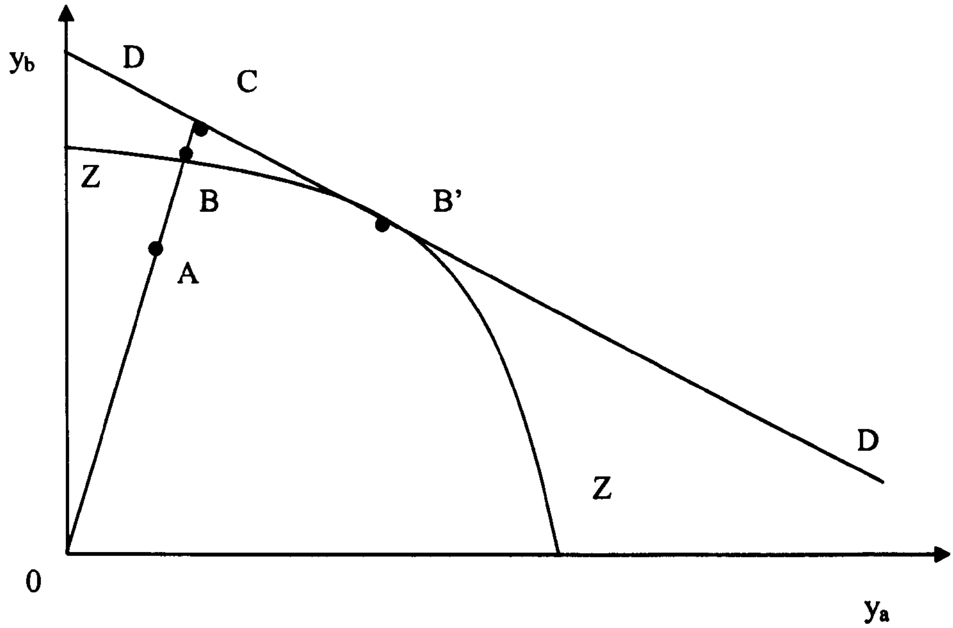
3.1.2 Output-orientated measures

The above input-orientated technical efficiency measure addresses the question: "By how much can input quantities be proportionally reduced without changing the output quantities produced?". One could alternatively ask the question: "By how much can output quantities be proportionally expanded without altering the input quantities used?". This gives output-orientated measures as opposed to the input-orientated measures discussed above.

[20] The Cobb-Douglas function is a particular type of production function, of the form output (measured in appropriate units) = $L^a * K^{(1-a)}$, where L is units of labour used, K is units of capital used, and a is a value between 0 and 1 (Varian 1992). Such a function has two particularly interesting properties: it exhibits CRS (that is, a given percentage change in both inputs will produce the same percentage change in the output) and diminishing returns to a single factor (that is, if one factor is increased by equal successive amounts while the other is held constant, the amount of additional output produced by each additional unit of the variable factor will gradually decline). Although these characteristics are not universally true for all situations, they are fairly representative of a large number of cases.

One can illustrate output-orientated measures by considering the case where production involves two outputs (y1 and y2) and a single input (x1). If the input quantity is fixed at a particular level, the technology can be represented by a production possibilities frontier in two dimensions. This example is depicted in Figure 4 where the line ZZ is the production possibilities frontier and the point A corresponds to an inefficient firm. Note that an inefficient firm operating at point A lies below the frontier, because ZZ represents the upper bound of production possibilities.

Figure 4: Technical and allocative efficiency from an output orientation



The Farrell output-orientated efficiency measures (see Färe et al. 1994) are defined as follows. In Figure 4, the distance AB represents technical inefficiency, i.e. the amount by which outputs could be increased without requiring extra input. Hence a measure of output-orientated technical efficiency is the ratio TE = 0A / 0B. If price information is available then the isorevenue line DD' can be shown, and allocative efficiency is defined to be AE = 0B / 0C, which has a revenue increasing interpretation. Finally, overall or economic efficiency is defined as the product of these two measures:

EE = (0A / 0C) = (0A / 0B) x (0B / 0C) = TE x AE.

Again, all these three measures are bounded by zero and one. The output- and input-orientated measures are equivalent measures of technical efficiency only when CRS exists (Coelli et al. 1998). It is important to note that technical efficiency has been measured along a ray from the origin to the observed production point. Hence these measures hold the relative proportions of inputs (or outputs) constant. One advantage of these radial efficiency measures is that they are unit invariant, i.e. changing the units of measurement does not change the value of the efficiency measure.

3.2 Data envelopment analysis

3.2.1 The constant returns to scale model

DEA involves the use of linear programming methods to construct a non-parametric piece-wise surface, or frontier, over the data. Efficiency measures are then calculated relative to this surface. A comprehensive review of the methodology is presented by Cooper et al. (2003).

The piece-wise-linear convex hull approach to frontier estimation, proposed by Farrell (1957), did not receive wide attention until the paper by Charnes, Cooper and Rhodes (1978), in which the term DEA was first used. Since then there has been a large number of papers which have extended and applied the DEA methodology (Hollingsworth 2003; Worthington 2004).

Charnes et al. (1978) proposed a model which had an input-orientation and assumed CRS. Subsequent papers have considered alternative sets of assumptions, such as Banker et al. (1984), in which a VRS model was proposed. The VRS model is discussed in the next sub-section, but here a description of the input-orientated CRS model is provided.

Some notation is required to begin with. Assume there are data on K inputs and M outputs for each of N firms. For the i-th firm these are represented by the column vectors x_i and y_i respectively. The KxN input matrix, X, and the MxN output matrix, Y, represent the data for all the firms.

$\min_{\theta\lambda} \theta$

subject to $\quad Y\lambda \geq y_i$

$\qquad X\lambda \geq \theta x_i$

$\qquad \lambda \geq 0$

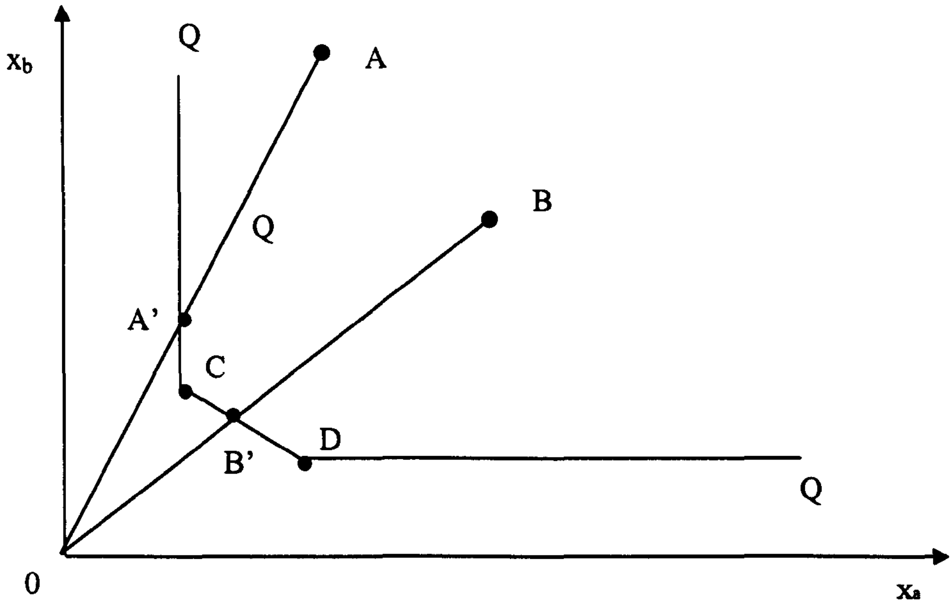
where θ is a scalar and λ is a Nx1 vector of constants. The value of θ obtained will be the efficiency score for the i-th firm. It will satisfy $\theta \leq 1$, with a value of 1 indicating a point on the frontier and hence a technically efficient firm, according to the Farrell (1957) definition. The linear programming problem must be solved N times, once for each firm in the sample. A value of θ is then obtained for each firm. Essentially, the above linear programming problem takes the i-th firm and then seeks to radially contract the input vector, x_i, as far as possible, while still remaining within the feasible input set.

3.2.1.1 Slacks, peers and targets

The piece-wise linear form of the non-parametric frontier in DEA can cause a few difficulties in efficiency measurement. Problem arise because of the sections of the piece-wise linear frontier which run parallel to the axes which do not occur in most parametric functions. To illustrate the problem see Figure 5 below, where the units using input combinations C and D are the two efficient units which define the frontier, and firms A and B are inefficient units. The Farrell (1957) measure of technical efficiency gives the efficiency of firms A and B as 0A' / 0A and 0B' / 0B respectively. However, it is questionable as to whether the point A' is an efficient point because the amount of input x_b used could be reduced (by the amount CA') and still produce the same quantity Q. This is known as input slack in the literature, although some authors use the term input excess (Coelli et al. 1998). The related concept of output slack also occurs.

In DEA, health centres which obtain an efficiency score of one (1.00) are regarded as (relatively) efficient, while those with scores less than one are classified as (relatively) inefficient. An input-orientated specification illustrates the level of radial contraction of inputs necessary for an 'inefficient' health centre to become 'efficient'. This level is established on the basis of efficient health centres that have been found to operate under similar technology. These efficient health centres are known as peers, with which a relatively inefficient health centre is compared (Coelli et al. 1998). Every efficient health centre is peer at least once, for itself, but may be act as a peer for other inefficient health centres. A score of zero denotes that the health centre was not efficient in that particular specification. On the other hand, a score of one indicates that the health centre has a unique combination of input(s) and output(s) compared to other health centres in the sample. DEA studies also refer to targets as well peers. These concepts are illustrated below in Figure 5. It should be noted, for example, that unit B could possibly reduce the consumption of both inputs to the point B' without reducing output. This projected point lies on a line joining points C and D. Firms C and D are therefore referred to as the peers of firm B. Furthermore, the targets of firm B are the coordinates of the efficient projection point B'.

Figure 5: Efficiency measurement and input slacks



3.2.2 The variable returns to scale model and scale efficiencies

The CRS assumption is only appropriate when all firms are operating at an optimal scale. In reality there are usually reasons to suspect that a firm may not be operating at optimal scale (Jacobs and Baladi 1996; Elbasha and Messonnier 2004). The use of the CRS specification when not all firms are operating at the optimal scale, results in measures of technical efficiency which are confounded by scale efficiency. Therefore, Banker, et al. (1984) suggested an extension of the CRS DEA model to account for VRS situations. The use of the VRS specification permits the calculation of technical efficiency devoid of these scale efficiency effects.

The CRS linear programming model can be modified to account for VRS by adding the convexity constraint $N1'\lambda = 1$ to provide:

$\min_{\theta\lambda} \theta$

subject to $Y\lambda \geq y_i$

$X\lambda \geq \theta x_i$

$N1'\lambda = 1$

$\lambda \geq 0$

where N1 is an Nx1 vector of ones. This approach forms a convex hull of intersecting planes which envelope the data points more tightly than the CRS conical hull, and thus provides technical efficiency scores which are greater than or equal to those obtained using the CRS model (Coelli et al. 1998).

The convexity constraint ($N1'\lambda = 1$) essentially ensures that an inefficient firm is only benchmarked against firms of a similar size. That is, the projected point (for that firm) on the DEA frontier will be a convex combination of observed firms. This convexity restriction is not imposed in the CRS case. Hence in a CRS DEA, a firm may be benchmarked against firms which are substantially larger (or smaller) than it. In this instance the λ-weights will sum to a value greater than (or less than) one.

If it is considered that the technology exhibits VRS, then it is possible to obtain a scale efficiency measure for each firm. This is achieved by conducting both a CRS and a VRS DEA. The technical efficiency scores obtained from the CRS DEA are then decomposed into two components, one due to scale inefficiency and the other due to 'pure' technical inefficiency. If there is a difference in the CRS and VRS technical efficiency scores for a particular firm, then this indicates that the firm has scale inefficiency. Scale inefficiency (SE) can be calculated by dividing the CRS technical efficiency (TE_{CRS}) score by the VRS technical efficiency TE_{VRS} score, where all these measures are bounded by zero and one:

$$SE = \frac{TE_{CRS}}{TE_{VRS}}$$

One shortcoming of this measure of scale efficiency is that the value does not indicate whether the firm is operating in an area of increasing or decreasing returns to scale[21]. This can be determined by running an additional DEA problem with non-increasing returns to scale (NIRS) imposed. This is carried out by altering the DEA model presented above by substituting the $N1'\lambda = 1$ restriction with $N1'\lambda \leq 1$, to provide:

$$\min_{\theta\lambda} \theta$$

$$\begin{aligned} \text{subject to} \quad & Y\lambda \geq y_i \\ & X\lambda \geq \theta x_i \\ & N1'\lambda \leq 1 \\ & \lambda \geq 0 \end{aligned}$$

The constraint $N1'\lambda \leq 1$ ensures that the i-th firm will not be benchmarked against firms which are substantially larger than it, but may be compared with firms smaller than it.

3.2.3 Adjusting for the environment

The results of DEA may be misleading because of the favourable or unfavourable environments in which some firms operate, such that there will always be some inherent inefficiency. Fried et al. (2002) recently summarised the existing approaches to incorporating environmental effects in DEA. These approaches can be grouped somewhat loosely into one-stage models and two-stage models. Single-stage approaches were developed by Banker and Morey (1986a; 1986b) for non-discretionary

[21] A firm is said to exhibit CRS when a unit increase in inputs yields a proportionate unit increase in output. Increasing returns occur if a unit increase in input yields a proportionately larger increase in output, and decreasing returns when a unit increase in input yields a proportionately smaller increase in output.

environmental variables (such as quasi-fixed inputs and / or outputs whose magnitudes are temporarily constrained by contractual arrangements), and also for categorical environmental variables (such as form of ownership, e.g. government, private, etc.). The approach to non-discretionary variables is to include them together with the inputs and outputs, but to restrict the optimisation to either inputs or outputs. An obvious requirement is that the direction of the impact on producer performance of each non-discretionary variable must be known in advance. The approach to categorical variables is to restrict the comparison set to other producers in the same or higher (or the same or lower) categories. This of course requires that the categories be nested, and reduces the size of the comparison set for most producers, thereby reducing the discriminatory power of the model. Both approaches are purely deterministic, and so are incapable of incorporating the effect of statistical noise on producer performance. A more detailed commentary on the two single-stage approaches is provided by Cooper et al. (2003).

The typical two-stage approach follows a first stage DEA exercise based on inputs and outputs with a second stage regression analysis seeking to explain variation in first stage efficiency scores in terms of a vector of observable environmental variables. Timmer (1971) pioneered this approach, and several subsequent studies have improved upon Timmer's second stage by using limited dependent variable regression techniques (because efficiency scores are bounded, and frequently achieve their upper bound). For example McCarty and Yaisawarng (1993) went a step further, by using the second stage regression residuals to adjust the first stage efficiency scores. Pastor (1995) suggested a novel variation on the two-stage approach by proposing a double DEA format. In the first stage he applied either input-oriented DEA to the inputs and environmental variables or output-oriented DEA to the outputs and environmental variables. He then replaced either the inputs or the outputs by their radial projections, in order to eliminate the effect of the environmental variables. In the second stage he again applied DEA to an expanded data set consisting of the originally efficient observations, the originally inefficient observations, and the radial projections of the originally inefficient observations. A comparison of the second stage efficiency scores of the originally inefficient observations with those of the radial projections of the originally inefficient observations reveals the impact of the environmental variables on producer performance.

It is also possible to extend the basic two-stage approach, as Fried et al. (1999) have done. In their approach an initial DEA evaluation is followed by a second stage tobit regression analysis to obtain predictions of the impacts of the environmental variables on the first stage performance evaluations. In the third stage, the original data are adjusted to account for these environmental impacts, and the DEA evaluation is repeated. The virtue of this approach is that the second stage is stochastic. The shortcoming of this approach is that the data adjustment accounts for environmental impacts, but not for the impact of statistical noise.

Coelli et al. (1998) recommend the two-stage approach because it has the advantages that it can accommodate more than one variable, and continuous and categorical variables. It does not make prior assumptions regarding the direction of the influence of the categorical variable and it is possible to conduct hypothesis tests to see if the variables have a significant influence upon efficiencies. And finally, the method is relatively easy to calculate and therefore transparent.

3.3 Stochastic frontier analysis

The following review of stochastic frontier modelling and efficiency measurement is brief. The purpose of this section is to provide an introduction to the method. More detailed examinations can be found in Kumbhakar and Lovell (2000). It is important to recall that Farrell (1957) proposed a measure of the efficiency of a firm that consists of two components: technical efficiency, which reflects the ability of a firm to obtain maximal output from a given set of inputs, and allocative efficiency, which reflects the ability of a firm to use the inputs in optimal proportions, given their respective prices. These two measures are then combined to provide a measure of total economic efficiency. These efficiency measures assume that the production function of fully efficient firms is known. However, because the production function is never known in practice, Farrell (1957) suggested that the function be estimated from sample data using either a non-parametric piece-wise linear technique or a parametric function, such as the Cobb-Douglas form. The first suggestion was taken up by Charnes et al. (1978), resulting in the development of the DEA approach reviewed above. The latter parametric approach was followed up by Aigner et al. (1977), subsequently resulting in the development of the stochastic frontier model.

Aigner and Chu (1968) considered the estimation of a parametric frontier production function of Cobb-Douglas form, using data on a sample of N firms. The model is defined by:

$$\ln(y_i) = x_i\beta - u_i \quad i = 1, 2, \ldots, n \tag{1.1}$$

where:

$\ln(y_i)$ is the logarithm of the (scalar) output for the i-th firm

x_i is the (K+1)-row vector, whose first element is "1" and the remaining elements are the logarithms of the K-input quantities used by the i-th firm

$\beta=(\beta_0,\beta_1, \ldots, \beta_K)'$ is a (K+1)-column vector of unknown parameters to be estimated

u_i is a non-negative random variable, associated with technical inefficiency in production of firms in the industry involved.

The ratio of the observed output for the i-th firm, relative to the potential output, defined by the frontier function, given the input vector, x_i, is used to define the technical efficiency of the i-th firm:

$$TE_i = \frac{y_i}{\exp(x_i\beta)} = \frac{\exp(x_i\beta - u_i)}{\exp(x_i\beta)} = \exp(-u_i) \tag{1.2}$$

This measure is an output-orientated Farrell measure of technical efficiency, which takes a value between zero and one. It indicates the magnitude of the output of the i-th firm relative to the output that could be produced by a fully-efficient firm using the same input vector. The technical efficiency, defined by equation (1.2), can be estimated by the ratio of the observed output, y_i, to the estimated value of the frontier output, $\exp(x_i\beta)$, obtained by estimating β using linear programming, where $\sum_{i=1}^{N} u_i$ is minimised, subject to the constraints $u_i \geq 0$, $i=1, 2, \ldots, N$.

Afriat (1972) specified a model similar to that of equation (1.1), except that the u_is were assumed to have a gamma[22] distribution and the parameters of the model were estimated

[22] A distribution used for continuous random variables which are constrained to be greater or equal to 0. It is characterised by two parameters: shape and scale. The gamma distribution is often used to model data which are positively skewed (Everitt 1995).

using the maximum-likelihood (ML) method[23]. Richmond (1974) noted that the parameters of Afriat's model could also be estimated using a method that has become known as correct least-squares (COLS). This method uses the ordinary least-squares (OLS) estimators, which are unbiased for the slope parameters, but the (negatively biased) OLS estimator of the intercept parameter, β_0, is adjusted up, using the sample moments of the error distribution, obtained from the OLS residuals. Schmidt (1976) highlighted that the linear and quadratic programming estimators, proposed by Aigner and Chu (1968), are ML estimators if the u_is are distributed as exponential or half-normal random variables, respectively.

One of the primary criticisms of the above deterministic[24] frontier model is that no account is taken of the possible influence of measurement errors and other 'noise' upon the frontier. All deviations from the frontier are assumed to be the result of technical inefficiency. Aigner et al. (1977) and Meeusen and van den Broeck (1977) independently proposed the stochastic frontier production function, as the solution of the 'noise' problem, in which an additional random error, v_i, is added to the non-negative random variable, u_i, in equation (1.1) to provide:

$$\ln(y_i) = x_i\beta + v_i - u_i, \; i = 1, \ldots, n \tag{1.3}$$

[23] This method is a general method of finding estimated values of parameters. It yields values for the unknown parameters, which maximise the probability of obtaining the observed values. The estimation process involves considering the observed data values as constants and the parameter to be estimated as a variable, and then using differentiation to find the value of the parameter that maximises the likelihood function. First a likelihood function is set up which expresses the probability of the observed data as a function of the unknown parameters. The ML estimators of these parameters are chosen to be those values, which maximise this function. The resulting estimators are those, which agree most closely with the observed data. This method works best for large samples, where it tends to produce estimators with the smallest possible variance. The ML estimators are often biased in small samples (Everitt 1995).

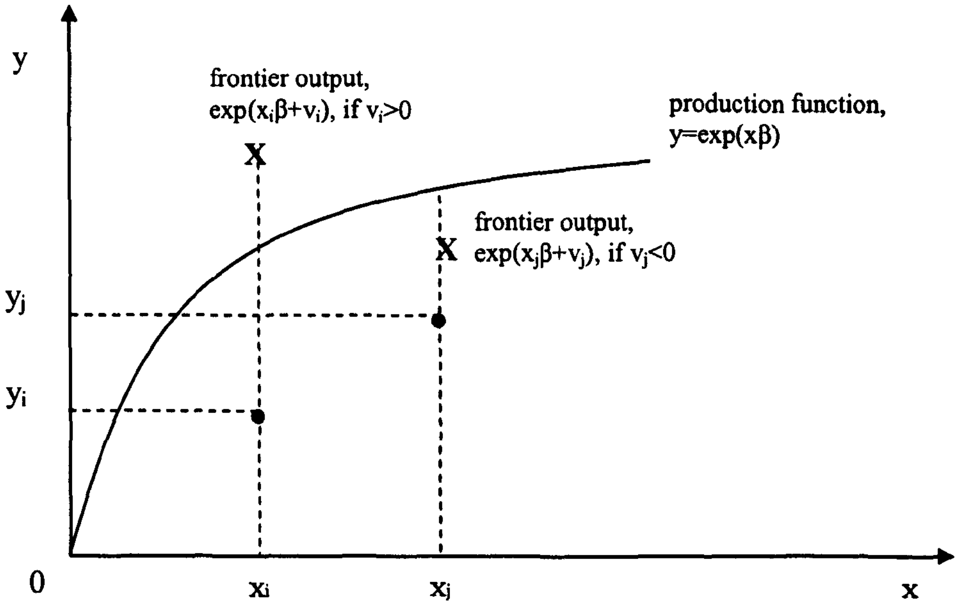
[24] The term deterministic is used because, in the frontier model of equation (1.1), the observed output y_i, is bounded above by the non-stochastic, i.e. deterministic quantity $\exp(x_i\beta)$. Thus, the models of Aigner and Chu (1968), Afriat (1972) and Schmidt (1976) are examples of deterministic frontiers.

The random error, v_i, accounts for measurement error and other random factors, such as the epidemics, the effects of weather and strikes on the value of the output variables, together with the combined effects of unspecified input variables in the production function. Aigner, Lovell and Schmidt (1977) assumed that the v_is were independent and identically distributed normal random variables with mean zero and constant variance, σ_v^2, independent of the u_is, which were assumed to be independent and identically distributed exponential or half-normal random variables.

The model, defined by equation (1.3), is termed the stochastic frontier production function because the output values are bounded above by the stochastic (random) variable, $\exp(x_i\beta+v_i)$. The random error, v_i, can be positive or negative and so the stochastic frontier outputs vary about the deterministic part of the frontier model, $\exp(x_i\beta)$.

The basic features of the stochastic frontier model are illustrated in Figure 6. The input(s) are represented on the horizontal axis and the outputs on the vertical axis. The deterministic component of the frontier model, $y = \exp(x\beta)$, is drawn assuming that diminishing returns to scale apply. The observed outputs and inputs for two firms, i and j, are presented on the graph. The i-th firm uses the level of inputs, x_i, to produce the output(s) y_i. The observed input-output value is indicated by the point marked with • above the value of x_i. The value of the stochastic frontier output, $y_i^* \equiv \exp(x_i\beta+v_i)$, is marked by the point X above the production function because the random error, v_i, is positive. Similarly, the j-th firm uses the level of inputs, x_j, and produces the output(s), y_j. However, the frontier output, $y_j^* \equiv \exp(x_j\beta+v_j)$, is below the production because the random error, v_j, is negative. Of course, the stochastic frontier outputs, y_i^* and y_j^*, are not observed because the random errors, v_i and v_j, are not observable. However, the deterministic part of the stochastic frontier model is seen to lie between the stochastic frontier outputs.

Figure 6: The stochastic frontier production frontier



Source: adapted from Coelli, Rao and Battese (1998)

This stochastic frontier model permits the estimation of standard errors and tests of hypotheses using traditional maximum-likelihood methods.

3.3.1 Maximum likelihood estimation

The parameters of the stochastic frontier production function, defined by equation (1.3), can be estimated using either the ML method or using a variant of the COLS method suggested by Richmond (1974). However, Coelli et al. (1998) state that the ML estimator should be given preference to the COLS estimator because it was found to perform better in a Monte Carlo experiment (Coelli 1995).

Aigner et al. (1977) derived the log-likelihood function for the model defined by equation (1.3), in which the u_is are assumed to be independent and identically distributed truncations (at zero) of a $N(0, \sigma^2)$ random variable, independent of the v_is which are assumed to be independent and identically distributed $N(0, \sigma_v^2)$. Aigner et al. (1977) expressed the likelihood function in terms of the two variance parameters, $\sigma_s^2 \equiv \sigma^2 + \sigma_v^2$ and $\lambda \equiv \sigma/\sigma_v$. Battese and Corra (1997) suggested that the parameter, $\gamma \equiv \sigma^2/\sigma_s^2$, be used because it has a value between zero and one, whereas the λ-parameter could be any non-negative value.

Battese and Corra (1977) showed that the log-likelihood function, in terms of this parameterisation is equal to:

$$\ln(L) = -\frac{N}{2}\ln(\pi/2) - \frac{N}{2}\log(\sigma_s^2) + \sum_{i=1}^{N}\ln[1-\phi(z_i)] - \frac{1}{2\sigma_s^2}\sum_{i=1}^{N}(\ln y_i - x_i\beta)^2 \qquad (1.4)$$

where $z_i = \frac{(\ln y_i - x_i\beta)}{\sigma_s}\sqrt{\frac{\gamma}{1-\gamma}}$; and $\Phi(.)$ is the distribution function of the standard normal random variable.

The ML estimates of β, $\sigma_s^2$and γ are obtained by finding the maximum of the log-likelihood function, defined in equation (1.4). The ML estimators are consistent and asymptomatically efficient (Aigner et al. 1977).

The computer programme, Frontier Version 4.1 (Coelli 1996b), can be used to obtain the ML estimates for the parameters of this model[25]. The programme follows a three-step estimation procedure:

[25] If starting values are specified in the instruction file, the programme will skip the first two steps of the procedure.

1. The first step involves calculating the OLS estimators of β and σ_s^2. These are unbiased estimators of the parameters in equation (1.3), with the exception of the intercept, β_0 and σ_s^2;
2. In the second step, the likelihood function is evaluated for a number of γ between zero and one (this is referred to a 'grid search'). In these calculations, the OLS estimates of β_0 and σ_s^2 are adjusted according to the corrected ordinary least squares formula presented in Coelli (1995). The OLS estimates are used for the remaining parameters in β;
3. The final step uses the best estimates (that is, those corresponding to the largest log-likelihood value) from the second step as starting values in a Davidon-Flectcher-Powell iterative maximisation routine which obtains the ML estimates when the likelihood function attains its global maximum.

3.3.2 Alternative functional forms

The model presented dealt with the case of the half-normal distribution for the technical inefficiency effects, because it has been most frequently assumed in empirical applications (Coelli et al. 1998). Its simplicity is an attractive feature. A logarithmic transformation provides a model which is linear in the logarithms of the inputs and, hence, the Cobb-Douglas form is easy to estimate. However, a common criticism of the stochastic frontier method is that there is no *a priori* justification for the selection of any particular distributional form for the technical inefficiency effects, u_i. The half-normal distribution is an arbitrary selection. As this distribution has a mode at zero, it implies that there is the highest probability that the inefficiency effects are in the neighbourhood of zero. This, in turn, implies relatively high technical efficiency. In practice, it may be possible to have a few very efficient firms, but a lot of quite inefficient firms.

A few researchers have attempted to address this criticism by specifying a more general distribution form, such as the truncated-normal (Stevenson 1980) distribution for the technical inefficiency effects. The truncated normal distribution is a generalisation of the half-normal distribution. It is obtained by the truncation at zero of the normal distribution with mean, μ, and variance, σ^2. If μ is pre-assigned to be zero, then the distribution is the half-normal.

A translog[26] production frontier assuming a truncated normal distribution

$$\ln(y_i) = x_i\beta + v_i - u_i, i = 1, \ldots, n \tag{1.6}$$

where:

$\ln(y_i)$, x_i, β and v_i are as defined above, and u_i has truncated normal distribution.

3.3.3 Tests of hypotheses

For the frontier model defined by equation (1.3), the null hypothesis that there are no technical inefficiency effects in the model, can be conducted by testing the null and alternative, H_0: σ^2=0 versus H_1: σ^2>0.[27] This hypothesis can be tested using a number of different statistical tests. The Wald statistic involves the ratio of the ML estimator for σ^2 to its estimated standard error. This statistic, or a slight variant of it, has been explicitly or implicitly conducted in almost every empirical analysis involving the stochastic frontier model since the first application by Aigner et al. (1977) (Coelli et al. 1998). In many cases one of the equivalent sets of hypotheses, H_0: λ=0 versus H_1: λ>1, or H_0:γ=0 versus H_1:γ>0, is considered depending upon the parameterisation used in the

---

[26] The transcendental logarithmic function allows a wide range of non-linear models to be expressed in linear form. It includes the logarithm of every explanatory variable, as well as their products and cross-products (Coelli et al.1998).

[27] σ^2 is the variance of the normal distribution which is truncated at zero to obtain the distribution of u_i. If this variance is zero, then all the u_is are zero, implying that all firms are fully efficient.

estimation of the stochastic frontier model. In this thesis, the Battese and Corra (1977) parameterisation is adopted, thus the hypotheses involving γ are considered.

For the Wald test, the ratio of the estimates for γ to its estimated standard error is calculated. If $H_0{:}\gamma=0$ is true, this statistic is asymptotically distributed as a standard normal random variable. The test must be performed as a one-sided test because γ cannot take negative values. However Coelli (1995) suggested that the one-sided generalised likelihood-ratio test should be performed when ML estimation is involved. This is because the Wald test has very poor size (i.e. probability of a Type I error) properties, whereas the one-sided generalised likelihood-ratio test has the correct size properties (Coelli 1995).

The generalised likelihood-ratio test requires the estimation of the model under both the null and alternate hypotheses. Under the null hypothesis, $H_0{:}\gamma=0$, the model is equivalent to the traditional average response function, without the technical inefficiency effect, u_i. The test statistic is calculated as:

$$LR = -2\{\ln[l(H_0)/L(H_1)]\} = -2\{\ln[L(H_0)] - \ln[L(H_1)]\} \quad (1.5)$$

where $L(H_0)$ and $L(H_1)$ are the values of the likelihood function under the null and alternative hypotheses H_0 and H_1 respectively.

If H_0 is true, this test statistic is usually assumed to be asymptotically distributed as a chi-squared random variable with degrees of freedom equal to the number of restrictions involved (in this instance one). The calculation of the critical value of this one-sided generalised likelihood-ratio test of $H_0{:}\gamma=0$ versus $H_1{:}\gamma>0$ is as follows. The critical value of a test of size α is equal to the value χ_1^2 (2α), where this is the value which is exceeded by the χ_1^2 random variable with probability equal to 2α. Thus the one-sided generalised likelihood-ratio (LR) test of size α is: "Reject $H_0{:}\gamma=0$ in favour of $H_1{:}\gamma>0$ if LR exceeds χ_1^2 (2α)". Thus the critical value for a test of size $\alpha=0.05$ is 2.71 rather than 3.84 (Coelli et al. 1998)

The one-sided likelihood test to test the null hypothesis that there are no technical inefficiency effects in the half-normal model can be extended for use in the truncated-normal model. If the null hypothesis, that there are no technical inefficiency effects in the model, is true, then the generalised likelihood-ratio statistic is asymptotically distributed as a mixture of chi-square distributions. The crucial value for this mixed chi-square distribution is 5.138 for a 5% level of significance.

3.3.4 Adjusting for the environment

A number of empirical studies (e.g. Pitt and Lee 1981) have estimated stochastic frontiers and predicted firm-level efficiencies using these estimated functions, and then regressed the predicted efficiencies upon firm-specific variables (such as managerial experience and ownership characteristics) in an attempt to identify some of the reasons for differences in predicted efficiencies between firms in an industry. This has long been recognised as a useful exercise, but the two-stage estimation procedure has also been considered to be one which is inconsistent in it's assumptions regarding the independence of the inefficiency effects in the two estimation stages. The two-stage estimation procedure is unlikely to provide estimates which are as efficient as those that could be obtained using a single-stage estimation procedure.

3.4 Dimensionality and sample size considerations

An issue in the choice of approach is sample size. One of the most problematic issues in DEA is the specification of the inputs and outputs to be included in the analysis. Careful reflection on the production process is usually sufficient to identify the main dimensions. The omission of variables may have an adverse effect on the efficiency estimates, but the number of variables cannot be increased without constraint. Increasing the number of dimensions used in the characterisation of production reduces the discriminatory power of the analysis, increasing measured efficiency and the number of firms identified as fully efficient (Coelli et al. 1998). DEA is therefore subject to an effect analogous to the loss of degrees of freedom in econometric analysis.

The relationship between the number of dimensions in the DEA problem and the discriminatory power of the analysis arises because of the flexibility in the weights used in making efficiency comparisons. DEA adopts the weights for each firm that maximise each firm's relative performance. As the number of dimensions is increased, the opportunity to differentiate one firm from the others also increases. A firm may therefore be deemed efficient, because of the lack of comparator observations. Such differentiated firms will be judged efficient, but will rarely be identified as a peer observation of other firms in the sample

Thus, a simple method suggested for identifying the loss of discriminatory power in DEA is to count the number of firms for which each efficient observation is identified as a peer. Boussofiane et al. (1991) argue that the minimum number of firms identified as efficient increases with the number of dimensions and will be approximately equal to the product of the number of inputs and outputs included, identifying the minimum sample size required for the analysis.

Rules of thumb commonly used with DEA suggest that the number of observations in the data set should be at least two to three times the sum of the number of input and output variables (Drake and Howcroft 1994; Cooper et al. 2003). There is an alternative expression by Dyson et al (2001), which states that the number of observations should be at least twice the product of the number of inputs and outputs. Avkiran (2002) suggests a further rule of thumb, which states that a sample is large enough if the number of fully efficient firms does not exceed one third of the sample. Thus, it is important to note that where a sample is small, it is possible that a high proportion of firms will be classed as efficient, some of which would not otherwise be considered efficient if a larger sample was used. Reducing the sample size will tend to inflate the average efficiency score as it creates fewer comparable organisations and improves the likelihood of any entity being placed on the frontier 'by default'.

Nevertheless, non-parametric methods are preferable for studies with small sample sizes because parametric methods are based on econometric techniques. Employing econometric methods on a small sample may not correctly separate random noise from inefficiency, which is one of the main advantages of SFA over DEA. Because SFA is based on regression methods, it requires a minimum sample size (usually at least, 30) to get significant results (Altman 1991).

In conclusion, it is difficult to define a clear-cut sample size below which inferences become problematic as this will ultimately depend on the quality and nature of the data, the number of explanatory variables and the estimation procedure being followed.

3.5 Strengths and weaknesses of both approaches

The following are the main strengths of DEA (Charnes et al. 1995; Coelli et al. 1998; Cooper et al. 2003; Fried et al. 1993):

- it gives a measure of efficiency that is empirically obtainable in a given scenario (e.g. given available resources and institutional set-up), as firms are directly compared against a peer or combination of peers. Hence one can compare the efficiency of individual facilities against realistic benchmarks;
- DEA does not impose a specified functional form to model and calculate the efficiency of a firm. Therefore, unlike the stochastic frontier models, DEA has the advantage of having few assumptions about the shape or form of the production and cost frontiers (which can cause model mis-specification and hence misleading results), as well as the distribution of the error terms;
- DEA accommodates multiple inputs and outputs in a single measure of efficiency, and can address efficiency issues directly instead of using average relationships. Consequently, DEA can pinpoint inefficient health care facilities from large samples, and indicate the extent of cost savings and efficiency gains from a shift to efficient production.

However, there have also been criticisms levied against this technique. Coelli et al. (1998) note the following limitations or possible problems, of DEA (which paradoxically are often the same characteristics that make DEA a useful tool):

- measurement error, outliers and other 'noise' may influence the shape and position of the frontier, and therefore influence the results;
- when few observations and many inputs and / or outputs are included many of the firms will appear on the DEA frontier Equally, the exclusion of an important input or output can result in biased results;
- the addition of an extra firm in a DEA cannot result in an increase in the TE scores of the existing firms. Similarly, the addition of an extra input or output in a DEA model cannot result in a reduction in the TE scores;
- the efficiency scores obtained are only relative to the best firms in the sample. The inclusion of extra firms may reduce efficiency scores;
- comparing the mean efficiency scores from two studies only reflects the dispersion of efficiencies of one sample relative to the other;
- treating inputs and / or outputs as homogenous commodities when they are heterogeneous may bias results;
- not accounting for environmental differences may give misleading indications of relative managerial competence.

To this list can be added:

- DEA is good at estimating 'relative' efficiency of a firm but it converges very slowly to 'absolute' efficiency (Cooper et al. 2003). In other words, it identifies how well a firm is performing in comparison to its peers but not compared to a 'theoretical maximum';
- because it gives a relative measure of efficiency it has the potential of justifying inefficiency, i.e. even those that appear to be efficient in the sample might actually be

inefficient in absolute terms (in the engineering sense) (Hollingsworth et al. 1999). This problem can, however, be minimised by using a large sample and data set;

- it is not prescriptive in what to do about inefficiencies, it only suggests where costs can be saved without reducing output (Sherman 1984). This is because the measure of inefficiency is based on the most efficient health facility in the group, which may itself be inefficient;
- because DEA is a non-parametric technique, statistical hypothesis tests are difficult to perform Fried et al. (1993).

The stochastic frontier model also, however, has problems. The main criticism is that there is generally no *a priori* justification for the selection of any particular distributional form for the u_is. Of the list of possible pitfalls in DEA presented above, most are applicable, in varying degrees, to SFA. In addition, SFA has a few specific problems of its own (Coelli et al. 1998; Kumbhakar and Lovell 2000), namely:

- the selection of a distributional form for the inefficiency effects is arbitrary. Therefore general distributions, such as the truncated-normal, are considered best;
- the stochastic frontier approach is only well-developed for single-output technologies, unless it is possible to aggregate output into a single measure.

However, stochastic frontiers also have some advantages relative to DEA (Coelli et al. 1998; Kumbhakar and Lovell 2000):

- DEA assumes all deviations from the frontier are due to inefficiency. If any 'noise' is present, e.g. due to measurement error, weather, strikes or epidemics, then this may influence the placement of the DEA frontier (and hence the measurement of efficiencies) more than would be the case with the stochastic frontier approach;

- statistical tests of hypotheses regarding the existence of inefficiency and also regarding the structure of the production technology can be performed in a SFA.

Therefore, on balance, Coelli et al. (1998) believe SFA is likely to be more appropriate than DEA where data are heavily influenced by measurement error. However, in the non-profit service sector, where:

- random influences are less of an issue;
- multiple-output production is important;
- prices are difficult to define;
- behavioural assumptions, such as cost minimisation, are difficult to justify,

the DEA approach may often be the optimal choice. But ultimately, the selection of the appropriate method should be made on a case-by-case basis.

3.6 Issues in the measurement of efficiency in the production of health services

In this section, some of the methodological difficulties involved in efficiency analyses are reviewed including adjusting for case mix; allowing for variation in technical quality; and knowledge of input prices.

3.6.1 Efficiency and case mix

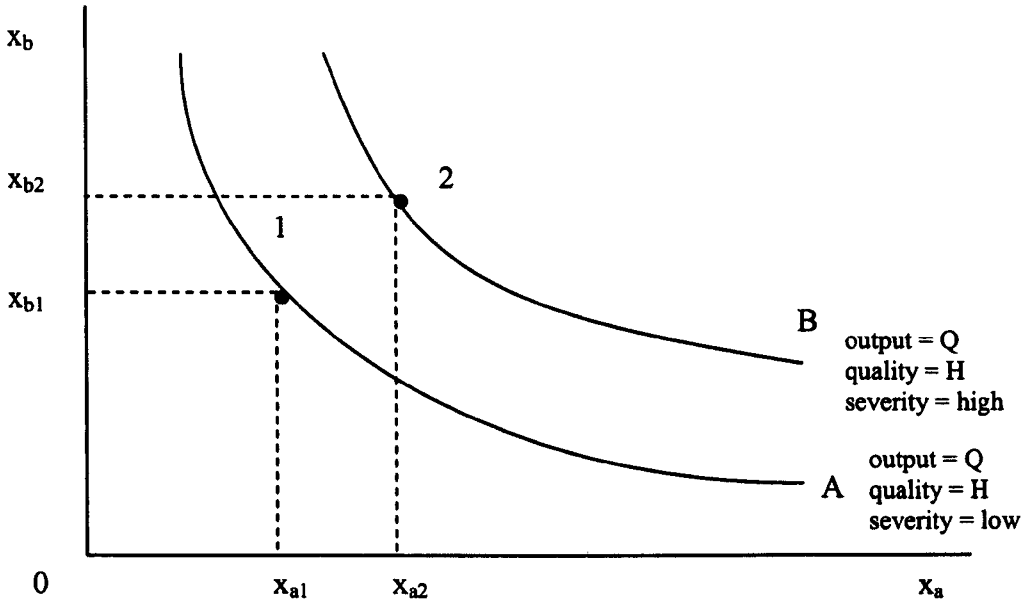
Case mix is an important, yet hard to define, concept through which researchers attempt to define output. Available definitions involve some or all of the following terms: facilities (or services) available; intermediate and final services provided; complexity of the cases treated; and patient characteristics (for example, age and gender)[28]. Everything else being constant, it would be expected that efficient providers dealing

[28] Health related groups (HRGs) and diagnostic related groups (DRGs) are examples of systems developed to better reflect case mix.

with different case mix to use different levels of inputs. For example, a facility with a greater proportion of complex cases should be expected to use more resources in providing health services to care for those cases, than an otherwise identical facility treating a set of patients with fewer severe cases.

Unless case mix is considered, comparative studies of technical and allocative efficiency among several providers are likely to be misleading and wrong. To illustrate this point, consider in Figure 7 the case of two providers, A and B, with B treating high severity patients, such as children with severe dehydration from diarrhoea, and A treating low severity patients, like children with mild dehydration from diarrhoea. Note, the number of children treated by providers A and B are assumed to be identical.

Figure 7: Case mix and efficiency



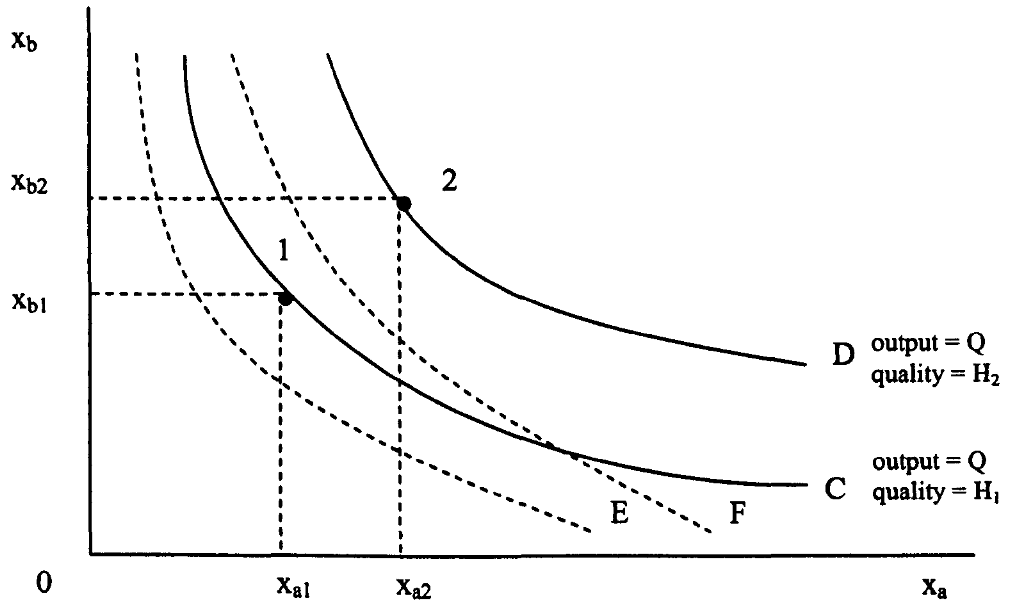
Highly dehydrated children may need to remain hospitalised for several days, often receive intravenous feeding and rehydration, and require close attention by the facility staff. Children with mild dehydration on the other hand, can be sent home with instructions to the parents on using oral rehydration salts and the appropriate treatment for diarrhoea.

Suppose that provider A performed at point 1 to treat high severity cases while provider B performed at point 2 to treat milder cases. If case severity was not considered, the uninformed researcher would wrongly conclude that provider A, the one with the lowest input use, is the more technically and allocatively efficient. If case mix were considered, however, the researcher would observe that the provider consuming the greatest amount of resources also happens to treat the most severe cases. Without further analysis, definitive statements about relative efficiency could not be made.

3.6.2 Efficiency and quality of care

Just as differences in case mix can obscure comparisons between technical and allocative efficiency among providers, so too can differences in the technical quality of care provided.

Figure 8: Technical quality of care and efficiency



Different levels of quality for example, often consume different levels of production inputs. Thus, failure to control for differences in quality may ascribe higher efficiency to lower-quality producers and *vice versa*. To illustrate this point, consider the two providers in Figure 8, C and D, each capable of producing the same volume of output (e.g. Q ambulatory visits) according to their respective isoquants. While both providers operate at the same output level, they provide different technical quality care: provider C is assumed to provide care of lower technical quality, H_1, while provider D provides care of a higher technical quality, H_2.

Suppose that provider C performs at point 1 and provider D performs at point 2. If an analyst attempting to compare technical and allocative efficiency between the two providers did not take into account their differences in technical quality, s/he would reach the conclusion that provider C is technically and allocatively more efficient than D. This would arise from the fact that provider C uses fewer production inputs than D (X_{a1} and X_{b1} versus X_{a2} and X_{b2}, respectively) and, as a consequence, provider C produces the level of output Q at a lower total cost than D. This conclusion however, would be incorrect. An appropriate comparison of efficiency is one which, at any given level of output, relates technical quality to input use. The analyst should therefore establish a relationship between H_1 and (X_{a1}, X_{b1}) for provider C and compare it with the equivalent relationship between H_2 and (X_{a2}, X_{b2}) for the provider D.

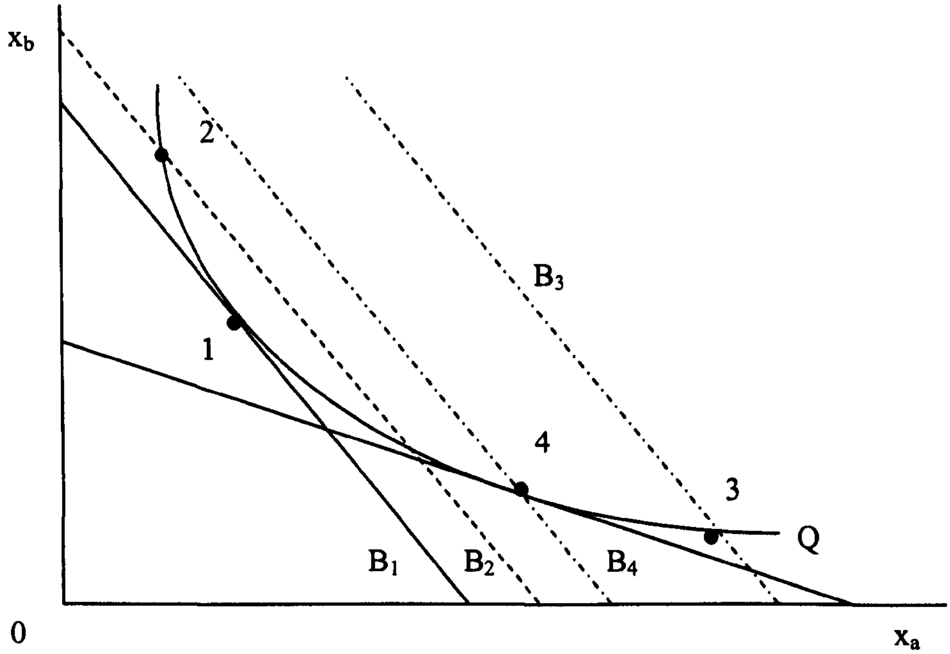
Contrary to what is suggested by isoquants C and D in Figure 8, higher technical quality does not necessarily imply greater use of inputs. Although it is assumed that technical quality is higher along the isoquant D than along C, and also that resource use is greater for D. This does not necessarily have to be the case for all situations, for example, consider production of quantity Q according to F. Provider F's technical quality could be higher than C's, with F using smaller quantities of inputs when both providers operate at the far right of their isoquants (that is, production that is intensive in resource X_a). Alternatively, the technical quality of provider E could be greater than that of C at all points, yet with E consuming fewer inputs than C and thus being technically and allocatively more efficient.

3.6.3 Allocative efficiency and input prices

Depending on a variety of circumstances such as the incentives, constraints, and information available to health facility managers, some providers may operate in a technically efficient, yet allocatively inefficient manner. For example, in the case of production input prices, allocative inefficiency arises when production occurs at a point that is not cost minimising. This can happen because facility managers either do not know their input prices or in spite of knowing the prices they fail to minimise their costs for a number of other reasons.

To distinguish between those two cases, consider the example of two providers operating at points 1 and 2 in Figure 9, each producing output level Q according to the same isoquant. Assume also that the two providers pay the same prices for their production inputs, X_a and X_b. Under those circumstances, provider 1 would be the most allocatively efficient of the two because the production cost would be B_1, lower than B_2. If an analyst wanting to study the allocative efficiency for these providers knew that both face the same input prices, s/he would not need to measure those prices at all to rightly conclude that 1 is more economically efficient than 2.

Figure 9: Allocative efficiency and input prices



Suppose, instead, that providers 1 and 2 face different input prices. Unless the analyst knew exactly what those sets of prices were, s/he would be unable to make any statements about the providers' relative allocative efficiency. For example, although both providers could be cost minimisers, given the different prices that they face, they could also operate at different points along the production frontier. Alternatively, both could operate at points that are not cost minimizing. Thus, in order to ascertain relative allocative efficiency, both knowledge and the use of price information would be essential.

3.7 Summary

- In 1957 Farrell defined a measure of firm efficiency which could account for multiple inputs. He suggested the use of either a non-parametric piece-wise linear convex isoquant constructed such that no observed point lies to the left or below it, or a parametric function fitted to the data, to estimate the efficient isoquants from sample data;

- Charnes et al. (1978) were the first to pursue the piece-wise-linear convex hull approach to frontier estimation proposed by Farrell (1957), resulting in the development of DEA;
- Aigner et al. (1977) pursued the use of a parametric function proposed by Farrell (1957), resulting in the development of SFA;
- Thus, there are two main alternative empirical approaches for estimating the production frontier that are distinguished by whether they are parametric (SFA) or not (DEA), and whether they are deterministic (DEA) or stochastic (SFA). Parametric methods assume a specific functional form for the frontier, whereas non-parametric methods do not; and deterministic methods assume that the distance of a unit from its frontier is a result of inefficiency whereas stochastic methods assume that this is also partially due to random error;
- Given the limitations of frontier techniques at present it may be that they are best employed in tandem, when possible, and if different methods suggest similar directions for results then the validity of such findings is enhanced.

The next chapter reviews the use of these two parametric and non-parametric frontier efficiency measurement techniques to obtain data on the efficiency of health care services / systems.

Chapter 4

A REVIEW OF FRONTIER EFFICIENCY MEASUREMENT TECHNIQUES IN HEALTH CARE SERVICES IN LOW- AND MIDDLE-INCOME COUNTRIES

In this Chapter, the evidence-base on the efficiency of health care services in low- and middle-income countries is reviewed. The review focuses exclusively on the use of parametric and non-parametric frontier efficiency measurement techniques in health care services / systems.

4.1 Introduction

Health care costs in most developed economies have grown dramatically over the last few decades, and it is widely believed that the inefficiency of health care institutions, has, at least in part, contributed to this phenomenon (e.g. Worthington 2004). In response, there has emerged, in recent years, a growing body of literature on the efficiency of health care services in industrialised countries, particularly in the US (Hollingsworth 2003).

Unfortunately, there has not been a similar focus on efficiency in the production of health care services in less-developed economies. This is particularly disappointing given the developing world's greater scarcity of financial resources such that the inefficient use of scarce resources exacting a much higher penalty in terms of foregone health benefits. Productivity and efficiency improvements are thus critical, given resource constraints faced by the health sector in many developing countries. Improving the efficiency of health services in developed and developing countries should be a major goal of public, private and non-profit providers alike. Knowledge of the levels and determinants of health services' efficiency can help policy-makers and health care managers take measures aimed at curtailing costs while maintaining acceptable levels of quality and access. However, there are methodological problems that make the measurement of both health services' productivity and efficiency challenging.

Methods for measuring the efficiency of providing health services were described in Chapter 3. Briefly, therefore, frontier methods entail the estimation of an efficiency frontier or envelopment surface from observed sample data, based upon best performance within the sample. Measurement of the deviation of individual production units from this frontier allows the calculation of relative efficiency scores, and the computation of potential efficiency gains if units could achieve best performance levels. There are two major features that distinguish alternative empirical approaches for forming the frontier and measuring efficiency and productivity: whether they are parametric or not, and whether they are deterministic or stochastic. Parametric methods assume a specific functional form for the frontier, whereas non-parametric methods do not. Deterministic methods assume that the distance of a unit from its frontier is a result of inefficiency whereas stochastic methods assume that some of this is due to random error. DEA is a non-parametric, deterministic method, while SFA is a parametric, stochastic method.

In order to assess whether, and the extent to which, productivity and / or efficiency has varied over time, the Malmquist index can be used. The Malmquist index is the mean of two indices, measuring the change in efficiency from one time period to the next, allowing a breakdown of efficiency changes over time[29].

4.2 Aim and objectives

The overall aim of this chapter is to identify applications of parametric and non-parametric measurement techniques of health service efficiency to low- and middle-income countries or regions. Studies were appraised with a view to determining: the methods and data used; models specified; and sensitivity analysis employed in order to better inform the subsequent parametric and non-parametric analyses in Chapters 7 and 9. The studies' results and policy implications are summarised in order to place the findings from Chapter 7 and 9 in a broader context.

4.3 Methods

4.3.1 Definition of low- and middle-income countries

This review has used the World Bank's classification of economies, which uses gross national income (GNI) per capita as its main criterion. Based on its GNI per capita, every economy is classified as low-income, middle-income (subdivided into lower-middle and upper-middle), or high-income. Low- and middle-income economies are sometimes referred to as developing economies (World Bank 2005a). Economies are divided according to 2004 GNI per capita, calculated using the World Bank Atlas method. The groups are: low-income, \$825 or less[30]; lower-middle-income, \$826-3,255; upper-middle-income, \$3,256-10,065; and high-income, \$10,066 or more.

[29] See Hollingsworth et al. (1999) for further details.
[30] Bangladesh, with a GNI per capita of \$430 in 2004 is a low-income country.

4.3.2 Search strategy

Studies were sought for the period 1983 (the year noted by Hollingsworth (2003) in which the first application of a frontier efficiency measurement technique was published) up to and including September 2005. The following databases were searched: EconLit, Medline and Web of Science (Science Citation Index and Social Science Index). Keywords and MeSH terms included were: "efficien*", "producti*", "health care", "data envelopment analysis", "DEA", "stochastic frontier analysis", "SFA" and "Malmquist". A free text search in Google using the same keywords and "similar pages" was also performed. The references of key studies were examined. Key journals were hand-searched (e.g. the Journal of Medical Systems and the International Journal of Operations and Production). A message was posted on the DEA and Health Economic list-servers requesting additional studies. Finally, researchers known to be active in this area were contacted (e.g. Emrouznejad, Hollingsworth, Kirigia, Ozcan and Sambo). Editorials and letters were excluded and the search was limited to English-language research covering developing countries and regions. Thus two Spanish-language studies were excluded (Penaloze Ramos 2003; Pinzon Martinez 2003).

4.4 Results

4.4.1 General characteristics

While Hollingsworth (2003) notes that the earliest application of a parametric or non-parametric frontier efficiency measurement technique was published in 1983 (Nunamaker 1983), a study from a low- and middle-income setting was not published until 1997 (Ersoy et al. 1997). However, since then, 23 additional studies have been published (Table 1). Of these 24 studies, 21 are intra-country analyses, of which 16 have used DEA, four SFA and one has compared both techniques. The study by Al-Shammari (1999) was re-visited by Sarkis and Talluri (2002). Evans et al. (2001) used econometric methods to analyse the efficiency of national health systems. These results were re-visited by Gravelle et al. (2003) and Hollingsworth and Wildman (2003) using parametric and non-parametric techniques. Thus, three studies compared inter-country variation in efficiency.

Five out of six World Bank regions were represented among the 21 intra-country studies, although only ten countries were represented (Table 1)[31]. Twenty of the studies were published in peer-reviewed journals, of which four were published in the Journal of Medical Systems and two in the International Journal of Operations and Production. Of the remaining studies, two were book chapters, one was a paper from the African Econometric Society's 10th Annual Conference of Econometric Modelling in Africa and one was a working paper. Six authors had co-authored two or more of the papers: Kirigia=4; Sambo=3; Valdmanis=3; Emrouznejad=2; Ozcan=2; and Walker=2.

[31] Including the Spanish language reports by Pinzon Martinez (2003) and Penaloza Ramos (2003), both of which were based in Colombia, would have ensured all six World Bank regions were represented.

Table 1: Summary of frontier measurement studies

Reference	Country / region	Topic	Number of units	Method(s)	Software used	Time period	Mean efficiency scores
Al-Shammari (1999)	Jordan	Public hospitals	15	DEA	LINDO	1991-93	91: 0.867 92: 0.937 93: 0.977
Bhat et al. (2001)	India	Public & grant-in-aid hospitals	20 and 21	DEA	NS	2000	Public: 0.85 Grant-in-aid: 0.89
Chakrabati & Rao (in press)	India	States	14	SFA	Frontier 4.1	1986-95	Mean: 0.692
Dervaux et al. (2003)	Bangladesh	Public and NGO vaccination delivery units	117	DEA	NS	1999	46 (40%) vaccination delivery units operated with non-optimal scheduling of sessions
Ersoy et al. (1997)	Turkey	Acute general hospitals	573	DEA	Integrated DEA System Version 5.1	1994	519 (91%) hospitals were inefficient
Evans et al. (2001)	Global	National health systems	191	SFA	NS	1993-97	Range: 0.08 – 0.992
Gravelle et al. (2003)	Global	National health systems	191	SFA	NS	1993-97	Country rankings and efficiency scores are sensitive to the

							definition of efficiency and choice of model specification
Hollingsworth & Wildman (2003)	Global	National health systems	191	Malmquist, DEA & SFA	NS	1993-97	DEA: 0.89 (0.49 – 1.00) SFA: 0.84 (min = 0.52)
Jacques & Koch (2005)	South Africa	Public hospitals	15	DEA	Frontier Analysis	1999-04	NS
Kathuria & Sankar (2005)	India	States	16	SFA	NS	1986-97	Fixed effects models: 0.69-1.00 Random effects model: 0.73-1.00 ML effects model: 0.72-1.00
Kirigia et al. (2000)	South Africa	Provincial hospitals	55	DEA	DEAP 2.1	1995-96	Mean TE: 0.906 Mean SE: 0.953
Kirigia et al. (2001)	South Africa	Public primary health care clinics	155	DEA	EMS Data Envelopment Software	1995-96	108 (70%) health centres were inefficient
Kirigia et al. (2002)	Kenya	Public hospitals	54	DEA	DEAP 2.1	NS	Mean TE: 0.84 Mean SE: 0.9
Kirigia et al. (2004)	Kenya	Public primary health care clinics	32	DEA	DEAP 2.1	NS	18 and 13 health centres were technically scale inefficient respectively
Masiye et al.	Zambia	Public hospitals	20	DEA	OnFront	1997	Mean: 0.64

(2002)							
Osei et al. (2005)	Ghana	Public district hospitals & health centres	17 and 17	DEA	DEAP 2.1	2000	8 (47%) hospitals were technically inefficient with a mean TE score of 0.61. 10 (59%) hospitals were scale inefficient with a mean SE score of 0.81. 3 (18%) health centres were technically inefficient with a mean TE score of 0.49. 8 (47%) were scale inefficient with a mean SE score of 0.84.
Owino & Korir (1997)	Kenya	Public provincial, district and sub-district hospitals	4, 22 and 10	SFA	Frontier 4.1	1994-96	Mean: 0.70
Pavananunt (2004)	Thailand	Public community hospitals	662	SFA	SPSS	1996-00	Mean: 0.55
Ramanathan et al. (2003)	Botswana	Health districts & district hospitals	22 and 13	DEA / SFA	DEAP 2.1 / Frontier 4.1	1997	3 (14%) districts and 1 (8%) district hospital were inefficient
Sahin & Ozcan (2000)	Turkey	Public hospitals	80	DEA	IDEAS	1996	55% of the public hospitals were inefficient

Sarkis & Talluri (2002)	Jordan	Public hospitals	15	DEA	NS	1991-93	1991 range: 0.34-1.00 1992 range: 0.42-1.00 1993 range: 0.52-1.00
Valdmanis et al. (2003)	Bangladesh	Vaccination delivery units	117	DEA	DEAP 2.1	1999	TE CRS: 0.33 TE VRS: 0.50 Scale: 0.64
Valdmanis et al. (2004)	Thailand	Public general hospitals	68	DEA	OnFront	1999	Possible increases in capacity utilisation amounted to 5% of capacity
Zere et al. (2001)	South Africa	Non-specific hospitals	86	DEA / Malmquist	DEAP 2.1	1992-93 1992-97	Mean: 0.74

NS: not stated; DEA: data envelopment analysis; SE: scale efficiency; SFA: stochastic frontier analysis; TE: technical efficiency

4.4.1 Methodological characteristics

Tables 2 and 3 summarise the methods used by the 17 DEA studies and 5 SFA applications respectively.

4.4.1.1 Software used[32]

For the 17 DEAs, the following software packages were used: DEAP 2.1 (n=7), OnFront (n=2), EMS Data Envelopment Software (n=1), Frontier Analysis (n=1), Integrated Data Envelopment Analysis System Version 5.1 (n=1), LINDO (n=1), IDEAS (n=1) and the package was not specified on five occasions. Frontier 4.1 (n=3) and SPSS (n=1) were used for the SFAs. One SFA failed to specify the software used.

4.4.1.2 Number and type of units

The unit of analysis ranged from vaccination delivery units to national health systems, encompassing every level of health care. More specifically, 14 applications are of hospitals, five of primary health care centres, of which two were of vaccination delivery units. In addition, two studies analysed administrative units, e.g. states and provinces. All the units of analysis were public, with the exception of one study which compared the efficiency of public and private-not-for-profit vaccination delivery units. Excluding the health system analyses, the mean number of health facility units examined in the papers reviewed was 94 with a range of 13-573.

4.4.1.3 Number and type of inputs

The typical inputs were: different types of personnel (e.g. doctors, nurses, other health staff and administrative staff), different types of capital items (e.g. size of the facility,

[32] See Appendix 4 for a review of the different parametric and non-parametric efficiency measurement software available.

beds, equipment and vehicles) and different types of recurrent items (e.g. drugs, vaccines and miscellaneous expenditure). The selection of the inputs was rarely justified. The mean number of inputs was six and ranged from 1-14.

4.4.1.4 Number and type of outputs

The typical outputs were admissions and OPD visits, which were broken down by type to varying degrees, e.g. general medical admissions, paediatric admissions, maternity admissions, dental care visits, OPD visits for 'poor' patients and OPD visits for 'rich' patients. While most studies simply used a count of hospital admissions, one study adjusted inpatient cases with an average DRG weighting. As illustrated in Chapter 3, SFA is only well-developed for single-output technologies, unless it is considered acceptable to aggregate output into a single summary measure. Interestingly, one study aggregated output by using the unit costs of the outputs as the weights, while another study ran 15 SFAs, i.e. an analysis for each of the outputs identified. With the exception of two of the SFAs, which used infant mortality rates, none of the intra-country studies used health outcomes as outputs. The selection of the outputs was rarely justified. The mean number of outputs was five and ranged from 1-14.

It is not good practice to rely on a single technique or model specification to test robustness of results which may influence policy. To guard against incorrect inferences being drawn, it is essential that models are subject to extensive sensitivity analysis (Gravelle et al. 2003). Unfortunately, many of the studies reviewed failed to subject their data to sensitivity analysis.

Table 2: Methodological summary of the DEA studies

Reference	Model(s)	Inputs	Outputs
Al-Shammari (1999)	1 input- and output-oriented VRS specification	3: bed days, physicians, other health personnel	3: patient days, minor operations, major operations
Bhat et al. (2001)	8 input- and output-oriented CRS specifications	14: physical infrastructure index, equipment index, beds, expenditure on drugs, maintenance expenditure, specialized infrastructure, specialized equipment, OPD hours per week, laboratory hours per week, doctors, nurses, paramedical staff, administrative staff, non-technical staff	5: medico legal cases, laboratory cases, inpatients cases, OPD cases, maternal and child health cases
Dervaux et al. (2003)	2 output-oriented VRS specifications	5: vaccine wastage, full-time equivalent staff, size of the facility, hours of operation and the number of sessions	5: DPT, TB, polio, measles and TT vaccines administered
Ersoy et al. (1997)	1 input-oriented CRS specification	3: beds, primary care physicians, specialists	3: inpatient discharges, outpatient visits and surgical operations
Jacques & Koch (2005)	NS	3: beds, nurses, doctors	3: inpatients days, surgeries, outpatient visits
Kirigia et al. (2000)	1 input-oriented VRS specification	9: doctors, nurses, paramedics, technicians, administrative staff, general staff, labour provisioning staff, other staff, beds	NS

Kirigia et al. (2001)	1 input- and output-oriented VRS specification	2: nursing staff, general staff	8: antenatal visits, births, child health, dental care visits, family planning visits, psychiatry visits, STI visits, TB visits
Kirigia et al. (2002)	1 input- and output-oriented VRS specification	12: medical officers / pharmacists / dentists, clinic officers, nurses, administrative staff, technicians / technologists, other staff, subordinate staff, pharmaceuticals, non-pharmaceutical supplies, maintenance of equipment, vehicles, and buildings, and food and rations	7: OPD casualty visits, special clinic visits, MCH/FP visits, dental care visits, general medical admissions, paediatric admissions, maternity admissions and amenity ward admissions
Kirigia et al. (2004)	1 input- and output-oriented VRS specification	6: clinical officers, other health staff, administrative staff, non-wage expenditure, beds	4: diarrhoea + malaria + STI + urinary tract infections + intestinal worms + respiratory disease visits, ANC + FP visits, immunisations, other general OPD visits
Masiye et al. (2002)	2 input-oriented VRS specifications	6: total expenditure, non-labour expenditure, doctors, other personnel, wages of doctors, wages of other personnel	7: child OPD visits, adult OPD visits, all OPD visits, child bed-days, adult bed-days, all bed-days, deliveries
Osei et al. (2005)	1 input-oriented and 1 output-oriented, both under VRS, specification used for hospitals and	4 for the hospital analysis: medical officers, technical officers, support staff and beds 2 for the health centre analysis: technical staff, support	3 for the hospital analysis: MCH care, deliveries, patients discharged 4 for the health centre: deliveries, FVCs, other

	health centres respectively	staff	MCH, OPD curative visits
Ramanathan et al. (2003)	1 input- and output-oriented CRS specification	7: hospitals in the district, clinics in the district, health posts in the hospital(s), beds, doctors, nurses, other health staff	14: 11 disease groups, new births discharged alive, inpatients discharged alive, patient days
Sahin & Ozcan (2000)	1 input-oriented VRS specification	6: patient beds, four levels of health labour, expenditure	3: mortality rate as quality measure, inpatient discharges and outpatient visits
Sarkis & Talluri (2002)	1 input- and output-oriented VRS specification	3: bed days, physicians, other health personnel	3: patient days, minor operations, major operations
Valdmanis et al. (2003)	1 input-oriented VRS specification	1: total costs	5: DPT, TB, polio, measles and TT vaccines administered
Valdmanis et al. (2004)	1 output-oriented CRS specification	7: beds, doctors, nurses, other staff, and allowance expenditure, drug expenditure and other operating expenditure	4: outpatient visits for poor patients, outpatient visits for non-poor patients, inpatient cases adjusted with average DRG weighting for poor patients, inpatient cases adjusted with average DRG weighting for non-poor patients
Zere et al. (2001)	1 input-oriented VRS specification	NS	NS

NS: not stated; ANC: antenatal care; CRS: constant returns to scale; DPT: diphtheria; DRG: diagnostic related group; FP: family planning; FVC: fully vaccinated child; MCH: maternal and child health; OPD: outpatient department; STI: sexually transmitted infection; TB: tuberculosis; TT: tetanus toxoid: VRS: variable returns to scale

Table 3: Methodological summary of the SFA studies

Reference	Model	Output(s)	Inputs
Chakrabarti & Rao (in press)	Technical efficiency effects model with panel data	A 'performance indicator' based on the infant mortality rate	6: Per capita PHC centres, per capita hospitals, health expenditure, births in institution, births in home by trained practitioners and per capita net state domestic product
Kathuria & Sankar (2005)	A Cobb-Douglas production function using the fixed-effects and random effects approaches with panel data	Infant mortality rate	5: Primary health centres, doctors, para-medical staff, hospital beds and % of institutional deliveries
Pavananunt (2004)	A Cobb-Douglas production function using the fixed-effects approach with panel data	IPD days, OPD visits and accident and emergency cases were combined into one aggregated output by using the unit costs of the outputs as the weights	3: Labour, capital and supplies / material
Owino & Korir (1997)	A Cobb-Douglas production function	cost	5: wages, admissions, outpatients, operations, beds
Ramanathan et al. (2003)	A half-normal distribution	15: 11 OPD disease groups, all outpatients, new births discharged alive, inpatients discharged alive and patient days	7: Hospitals in the district, clinics in the district, health posts in the hospital(s), beds, doctors, nurses and health staff

IPD: inpatient department; OPD: outpatient department ; PHC: primary health care

4.4.2 Summary of findings and recommendations

4.4.2.1 Main findings

Given that findings can differ for a number of reasons including differences in case mix and levels of technical quality, together with model specification issues, estimation techniques and data availability and quality (Hollingsworth 2003), results from different studies are not strictly comparable. Results are therefore strictly only valid for the units under investigation, and hence are not necessarily generalisable. For these reasons, no attempt was made to meta-analyse the findings by different types of units. A brief summary of each paper is provided below.

4.4.2.2 Analysis of explanatory variables

Institutional factors at the discretion of management as well as environmental factors beyond its control can affect the efficiency of a facility. Ten of the intra-country analyses failed to perform an analysis to explain variation in efficiency. However, of the remaining studies, Table 4 illustrates the range of variables tested, whether or not they were found to be significant predictors of efficiency and in which direction, and finally the methods used. It can be seen that 10 of the studies failed to perform an analysis of explanatory variables. Of those that did, location and type of ownership were the most commonly used variables. Two-stage regression was the method most often used to analyse the impact of environmental variables.

Table 4: Authors' analyses of explanatory variables

Reference	Explanatory analysis performed?	If yes, which variables considered?	Which, if any, of these, were significant?	Which method(s) were used?
Al-Shammari (1999)	No	-	-	-
Bhat et al. (2001)	Yes	Type, location	Grant-in-aid hospitals more efficient than public hospitals	Mann Whitney
Chakrabati & Rao (in press)	Yes	Literacy rate, proportion of rural population, revenue and capital expenditure on water supply and sanitation, year and trend	All except the trend variable. Capital expenditure on water supply and sanitation and literacy have a negative impact, proportion of rural population has a positive impact and revenue expenditure has a positive impact	Single-stage analysis with technical inefficiency effects specified in the model
Dervaux et al. (2003)	Yes	Ownership and type	Neither	Kruskal-Wallis
Ersoy et al. (1997)	No	-	-	-
Jacques & Koch (2005)	Yes	Average bed occupancy rate, size, bed utilisation, data availability and response time	Data availability has a positive impact	Two-stage regression
Kathuria & Sankar (2005)	Yes	Rural literacy rates, health expenditure as a share of GDP, per capita income,	% of population having a lavatory has a positive impact	Two-stage regression

		availability of water, % of population having a lavatory, % of mothers receiving after-birth care, % of children vaccinated		
Kirigia et al. (2000)	No	-	-	-
Kirigia et al. (2001)	Yes	Nursing staff, general staff, ANC visits, births, child health care visits, dental care visits, FP visits, psychiatry visits, STI visits, TB visits	Nursing staff has a negative impact, births and dental care visits have a positive impact	Two-stage regression
Kirigia et al. (2002)	No	-	-	-
Kirigia et al. (2004)	No	-	-	-
Masiye et al. (2002)	No	-	-	-
Osei et al. (2005)	No	-	-	-
Owino & Korir (1997)	Yes	Shortage of staff, poor combination of inputs, irregular or non-functional theatres and laboratories, transport problems, lack of, or poor distribution of drugs and medical supplies, frequent breakdown or poor servicing of equipment	all	Survey
Pavananunt (2004)	Yes	External factors: location, level of	External factors: level of competition	Multiple regression

		competition, i.e. no. of hospitals / clinics nearby and community demographic. Internal factors: age of hospital, size, technology, managing service, managing human resources, managing financial resources	and community demographic. Internal factors: age, size and managing financial resources	analysis
Ramanathan et al. (2003)	No	-	-	-
Sahin & Ozcan (2000)	No	-	-	-
Sarkis & Talluri (2002)	No	-	-	-
Valdmanis et al. (2003)	Yes	Ownership, type, length of time a unit had been in operation	Government units more efficient than NGO units, fixed units more efficient than outreach units, length of time a unit had been in operation positively correlated with efficiency	F-test, Median test, Kruskal-Wallis, correlation
Valdmanis et al. (2004)	Yes	Type and region	Region	Kruskal-Wallis
Zere et al. (2001)	Yes	Occupancy rate, average length of stay, outpatient visits	Higher occupancy rates are associated with level of technical efficiency, an increase in the number of outpatient	Censored Tobit model

		visits relative to inpatient days is likely to result in an increase in technical efficiency

ANC: antenatal care; FP: family planning; GDP: gross domestic product; NGO: non-government organisation; STI: sexually transmitted infection; TB: tuberculosis

4.4.2.3 Policy recommendations

Ten studies did not state any policy recommendations to improve efficiency in spite of finding high levels of inefficiency. Of the remaining studies, 15 strategies to improve efficiency were suggested. Most of these were supply-side strategies (see Table 5).

Table 5: Authors' suggested strategies to improve efficiency

Supply-driven	Demand-driven	Other
- close beds	- boost demand for services with unmet need	- use excess non-wage expenditure to improve the degree of responsiveness of dispensaries to patients' legitimate expectations
- contract with private clinic practitioners to use excess beds at a price		- use excess non-wage expenditure to improve health centres' quality of services
- identify characteristics of best performers and replicate in inefficient clinics		- use excess non-wage expenditure to support communities to start / sustain systematic risk and resource pooling and cost sharing mechanisms
- increase OPD activities		
- merge hospitals		
- reallocate surplus input to nearby or needy firms		
- replace jobs-till-old-age-retirement with fixed shorter term renewable contracts		
- sell excess beds		
- send excess general staff on early retirement		
- transfer excess beds / staff to more efficient health facilities		

4.5 Summaries of each study

Summaries of all the efficiency studies are presented in the following section organised by type of analysis and geographical region.

First though, Ramanathan et al. (2003) attempted to construct and present relative efficiency indices for the services provided by 22 health districts and 13 hospitals in Botswana, using SFA and DEA. The analysis indicated that three districts had efficiency scores below the optimum level. Among the 13 hospitals considered, only one was found to have an efficiency score of less than one. The authors stressed that because health services involve a number of factors, their findings ought to serve as indicators for further scrutiny of those units (health districts and hospitals) that fell below the optimum efficiency level.

4.5.1 DEA applications

4.5.1.1 Studies from Europe and Central Asia

Ersoy et al. (1997) used DEA to examine the technical efficiency of 573 Turkish acute general hospitals. Results illustrated that less than 10% of Turkish acute general hospitals operated efficiently compared to their counterparts. Inefficient, compared to efficient hospitals, on average utilised 32% more specialists, 47% more primary care physicians, and have 119% more staffed bed capacity. They also produced on average less output; specifically, 13% less outpatient visits, 16% inpatient episodes and 57% less surgical procedures. The authors note that hospital managers in Turkey generally have more control over inputs and therefore argue that they should devote more attention to the examination of inefficiencies generated by excessive input usage.

Again in Turkey, Sahin and Ozcan (2000) used DEA to examine public sector hospital efficiency in 80 provincial markets. Their results showed that 55% of the public hospitals operated inefficiently. An analysis of inefficient provinces suggested that the 44 of these were over-bedded and employed excessive number of specialists and other health workers. They also spent approximately $70,000,000 from their revolving funds in excess compared to efficient provinces.

4.5.1.2 Studies from sub-Saharan Africa

Kirigia et al. (2000) employ DEA to identify and measure efficiency among 55 public hospitals in Kwazulu-Natal Province, South Africa. The authors found that the overall average level of technical efficiency among these hospitals in 1995-96 was 90.6%. Twenty-two (40%) of the hospitals had some level of technical inefficiency and 32 (58%) were scale inefficient. In total, the authors estimated that the following inputs were not necessary in the production of hospital's stated output: 117 doctors, 2,709 nurses, 61 paramedics, 58 technicians, 295 administrative staff, 835 general staff and 1,752 beds. The authors provided a number of policy options that decision-makers might consider in order to realise these savings.

Again in Kwazulu-Natal Province, Kirigia et al. (2001) investigated the technical efficiency of 155 primary health care centres. Forty seven (30%) were found to be technically efficient. Among the 108 (70%) technically inefficient facilities, 16% had an efficiency score of 50% or less. To achieve technical efficiency, the authors estimated that Kwazulu-Natal centres would, in total, have to decrease inputs by 417 nurses and 457 general staff. Alternatively, outputs would have to be increased by 115,534 antenatal visits, 1,010 deliveries, 179,075 child care visits, 5,702 dental visits, 121,658 family planning visits, 36,032 psychiatric visits, 56,068 sexually transmitted disease visits and 34,270 tuberculosis visits.

Kirigia et al. (2002) examined the technical efficiency of 54 public hospitals in Kenya. Fourteen (26%) of them were found to be technically inefficient and 16 (30%) of the hospitals were scale inefficient. The authors provided the magnitudes of specific input reductions or output increases needed to attain efficiency. They also provided some suggestions for hospitals with excess inputs that policymakers might consider to improve efficiency, e.g. transferring excess staff to other health centres. With respect to increasing output, the authors suggested that the Ministry of Health could embark on a campaign to boost demand for under-utilised services.

In 2004, once again in Kenya, Kirigia et al. (2004) measured the technical efficiency of 32 public health centres. Their analysis suggested that 14 (44%) of these public health centres were technically inefficient. The inefficient health centres had an average technical efficiency score of 65%, which implied that on average, they could reduce their utilisation of inputs by about 35% without reducing output. In addition, 13 (41%) of the health centres were scale inefficient, and these centres had an average scale efficiency score of 70%. This implied that there was potential for increasing output by about 30% using the existing capacity / size.

Zere et al. (2001) assessed the efficiency of 86 hospitals in South Africa. The results suggested that a significant number of the hospitals included in their analysis operated well below the efficient frontier; the mean efficiency was 0.74. The authors argue that, given the tight fiscal constraints and resulting stagnant real per capita health budgets in

South Africa, extending and improving the quality of primary health care services has to be funded through health service efficiency gains and / or increased health service revenue from non-tax sources. At the time the paper was written (2001), the main source of such revenue was that of user fees at public sector hospitals. In 1992-93, fee revenue was equivalent to approximately 9% of public sector hospitals' recurrent expenditure and fee revenue is noted to have declined dramatically since then. Thus the potential efficiency savings estimated in this study amount to more than three times that of the fee revenue collected, which means that very high levels of user fees would be required to generate revenues that could match the potential efficiency savings. The authors suggest the following options might be worth exploring to achieve the efficiency savings: bed closures, particularly in those hospitals that exhibit decreasing returns to scale; mergers of hospitals that exhibit increasing returns to scale, particularly those that are in close proximity to one another.

4.5.1.3 Studies from Middle East and North Africa

Al-Shammari (1999) evaluated the productive efficiency of 15 hospitals for a three-year period In Jordan. In 1991, eight (53%) of the sample of 15 hospitals were found to be operating inefficiently. In 1992 and 1993, the number of inefficient hospitals had fallen to six and four respectively. The author estimated both the potential reduction in the usage of inputs and the potential increase in the production of outputs for the hospitals identified as inefficient. Al-Shammari (1999) considered that the results could help policy-makers by providing new insights on the distribution of health resources to hospitals that will have the highest potential to utilise additional resources.

Sarkis and Talluri (2002) addressed certain issues that were not addressed by Al-Shammari (1999). Specifically it considered: the simultaneous evaluation of all units across three years, ranking of efficient units and identification of 'global' benchmarks for improvement (the benchmarks are global because an inefficient unit in a particular year could have benchmark hospitals for the same as well as other years). The authors believe the identification of global benchmarks provide more complete information for the decision-maker about best practices necessary to improve the performance of the inefficient hospitals

4.5.1.4 Studies from South Asia

Bhat et al. (2001) used DEA to analyse efficiency of district and grant-in-aid hospitals in India. The findings suggest that the efficiency variations are more significant within district hospitals than within the grant-in-aid institutions. The overall efficiency levels of grant-in-aid institutions are higher than the district level hospitals. The grant-in-institutions are relatively more efficient than the public hospitals. These differences are statistically significant. The study made an attempt to find whether location determines the efficiency levels of hospitals. For example, it may be argued that hospitals in remote areas, less dense or less urbanised areas would be relatively serving lesser population and therefore would be relatively less efficient. The mean difference of urban population and density of districts between less efficient hospitals and relatively efficient hospitals were not significantly different statistically.

Valdmanis et al. (2003) examined whether and to what degree a sample of 118 vaccination delivery units located in Dhaka City, Bangladesh, exhibited CRS, based on data collected in 1999. The authors found that the units were, on average, relatively inefficient both in terms of technical inefficiency as well as scale inefficiency. In order to become technically efficient, the units would have had to decrease their costs by an average of 50%, and if they had been operating at the right size, costs could have been reduced by a further 36%. The authors also considered some of the environmental factors that affected scale efficiencies, because these factors may have been beyond managerial control, yet affected units' positions *vis-à-vis* the best practice frontier. Units that were relatively more inefficient, on average, were NGO outreach delivery units. Therefore, the government owned units, perhaps due to more centralised control, were better at long term planning. It was also found that units that had been practicing longer were relatively more scale efficient, which is perhaps attributable to a learning curve effect.

Using the same data, Dervaux et al. (2003) modelled the optimal number of clinic hours and sessions needed in order to maximise outputs, i.e. vaccines administered. This analysis required two models: one DEA model with possible reallocations between the number of hours and the number of sessions but with the total amount of time fixed; and one model without this kind of reallocation in scheduling. Comparing these two scores identified the 'gain' that would be possible were the scheduling of hours and sessions modified while controlling for all other types of inefficiency. The authors found that optimality of scheduling was, on average, around seven sessions, with each session lasting four hours, per month. If optimality had been met, gains (i.e. the increase in vaccines administered) of between 10-20% could have been achieved.

4.5.1.5 Studies from East Asia and the Pacific

Valdmanis et al. (2004) used DEA to assess the capacity of 68 Thai public hospitals to proportionately expand services to both the poor and the non-poor. The authors found that increases in the amount of services provided to poor patients did not reduce the amount of services to non-poor patients. Overall, hospitals appeared to be producing services relatively close to their capacity given fixed inputs. Possible increases in capacity utilisation amounted to 5% of capacity.

4.5.2 SFA applications

4.5.2.1 Studies from sub-Saharan Africa

Owino and Korir (1997) set out to investigate and determine levels, causes, and effects of inefficiency in the public health system in Kenya. This study revealed an average inefficiency level of 30%. The inefficiency is a primary attribute of shortages of professional staff; poor combinations of inputs; irregular or non-functioning operating theatres and laboratories; transport problems; poor distribution or lack of drugs and medical supplies; and frequent breakdowns and poor servicing of machines and equipment.

4.5.2.2 Studies from South Asia

Chakrabarti and Rao (in press) estimated a stochastic production frontier with inefficiency effects based on data drawn from the fourteen major states of India over the period 1986 to 1995. The output of the production frontier was generated on the basis of infant mortality rates of the respective states. Elasticity estimates of the inputs incorporated in the production frontier, computed on the basis of the obtained maximum likelihood estimates of the parameters, contradicted the general notion that expenditure on curative services does not generate a substantial impact on health. In fact, health expenditure as a percentage of net state domestic product and births in an institution, with relatively higher elasticity values were found to play a dominant role. Surprisingly, per capita real net state domestic product, which is a measure of an individual's command over privately supplied medical service, was found to have a relatively lower impact on output.

Kathuria and Sankar (2005) analysed the performance of rural public health systems of 16 major states in India using SFA and panel data for the period 1986-97. The results illustrated that the states differed not only in capacity-building in terms of health infrastructure created, but also in efficiency in using those inputs. It was found that not all states with better health indicators have efficient health systems. The authors concluded by noting that states should not only increase their investment in the health sector, but should also manage it more efficiently in order to achieve better health outcomes.

4.5.2.3 Studies from East Asia and the Pacific

Pavananunt (2004) analysed the technical efficiency of 662 public community hospitals in Thailand by using the fixed-effect production function model approach. The principal variables used for the analysis were service output indicators and inputs used for the provision of the services. The results indicate that larger size hospitals tend to be more efficient than the small size hospitals. The distribution of efficiency scores among the sampled hospitals clustered around 0.05-0.63 with a mean value of 0.55. Using the efficiency scores, hospitals were categorized into four groups. Among the sampled community hospitals, 11% were ranked in the most efficient group, 42% in moderate efficiency category, 38% in low efficiency and 9% in the least efficient group. The determinants of efficiency were also investigated by using a multiple regression model. The results indicate that internal factors, such as, age and size of community hospitals and aspects related to the management of human resources, were significantly associated with technical efficiency scores.

4.5.3 Malmquist productivity applications

The study by Zere et al. (2001) from South Africa described above also documented a decline in productivity by 12% among 86 hospitals over the period 1992 to 1993 due to a lack of technological advance.

4.5.4 Studies of health systems

It is important to note that the data presented above from the available DEA, SFA and Malmquist productivity applications were not collected for the purpose of cross-country comparisons, but rather, the studies were performed in isolation. Therefore, the current state of knowledge about cross-country differences in health service productivity and efficiency is limited. However, the keenly debated ranking of national health systems performed by the World Health Organization (2000) represents a useful starting point.

The World Health Report 2000 (2000) focused on the performance of health care systems around the globe. It sought to improve the evidence-base for health policy by devising a method to measure and monitor the performance of health systems. More specifically, the report describes the relationship between population levels of health and the inputs used to produce health in 191 countries. The report used efficiency measurement techniques to create a league table of health-care systems, highlighting 'good' and 'bad' performers. Evans et al. (2001) described the methods used in the report. Using econometric methods, the estimated efficiency varied from nearly fully efficient (in relative terms) to nearly fully inefficient. Countries with a history of civil conflict or high prevalence of HIV / AIDS were less efficient. Performance increased with health expenditure per capita. They concluded that increasing the resources for health systems is critical to improving health in poor countries, but important gains can be made in most countries by using existing resources more efficiently.

Hollingsworth and Wildman (2003) argued that WHO's estimation procedure was too narrow and that contextual information was hidden by the use of one method. They used and validated a range of parametric and non-parametric empirical methods to measure efficiency using the WHO data. The rankings obtained were compared to the WHO league table and demonstrated that there were trends and movements of interest within the league tables. The authors recommended that the WHO broaden its range of techniques in order to reveal this hidden information.

Gravelle et al. (2003) assessed the robustness of the WHO results to definitions of efficiency and statistical procedures. The paper used the data originally analysed by the WHO. The results show that the country rankings and efficiency scores are sensitive to the definition of efficiency and choice of model specification. The authors concluded that econometric methods can yield insights into complex socio-economic phenomena. However, the lack of robustness of results to reasonable alternative specifications suggests that it is premature to use the methods adopted by the WHO to construct league tables of health systems.

4.6 Summary

- There is a dearth of parametric and non-parametric efficiency measurement studies of health care in low- and middle-income settings. However, what there is suggests that resources are used inefficiently in the delivery of health care;
- There is an emphasis, albeit weak, on hospital efficiency research in developing countries, which coincides with that in the developed world (Hollingsworth 2003). This can partly be explained by the fact that: hospitals account for the largest share of health care costs; governments tend to keep information on utilisation and costs, however inaccurate, in a uniform way, whereas private providers generally do not; the search for health care financing and delivery reform has focused on gauging and improving the performance of the public sector;
- Given that findings can differ for a number of reasons including differences in case mix and levels of technical quality, together with model specification issues, estimation techniques and data availability and quality, results from different studies may not be strictly comparable. Results may therefore only be valid for the units under investigation, and hence are not necessarily generalisable;
- Very few studies have subjected their data to sensitivity analysis, nor compared and contrasted the application of DEA and SFA to the same data sets;
- A number of policy recommendation were touted by authors of the studies to improve efficiency;
- Therefore, analysis of the efficiency of primary health care in Bangladesh clearly fills many gaps in the literature

The next Chapter provides a brief overview of the Bangladeshi health system in advance of Chapters 6-9, which present cost and efficiency analyses of aspects of Bangladeshi primary health care services in urban and rural areas.

Chapter 5

BACKGROUND AND CONTEXT – BANGLADESH

In this Chapter, a description of the Bangladeshi health system is provided. The main aim of this chapter is to help place in context the case studies of vaccination services in Dhaka, and the delivery of primary health care in rural areas, which will be examined in Chapters 6-9.

5.1 Introduction

Bangladesh is a South East Asian republic bordering India, Myanmar and the Bay of Bengal (see Figure 10 below). Initially a part of Pakistan, known as East Pakistan, following partition from India in 1947, Bangladesh achieved full independence in 1971. In 1991 a parliamentary democracy replaced the military regime. It has nearly 600km of coast and is low-lying with many rivers, forming a fertile delta which experiences frequent and severe flooding. A tropical monsoon climate generates frequent cyclones[33]. Rivers and flooding inhibit the development of road and rail transport; waterways are therefore significant.

The estimated 2004 population of Bangladesh is 140.5 million (World Bank 2005b), living within an area of 147,520 sq. km making it one of the most densely populated countries in the world with 952 people per sq. km. The population is projected to double to around 250 million by the year 2035 before demographic growth stabilises (Vaughan et al. 2000). According to the 2001 population census, the urban population in Bangladesh is 29 million, and has increased by 38% during the last ten years, which

[33] In 1970 500,000 people were killed in one of worst ever recorded natural disasters.

is about four times the rural rate. The population is largely Bengali and there are small numbers of ethnic minorities. The majority (85%) are Sunni Muslim, others are mostly Hindu. Life expectancy is 62 years and the population is young, with few aged over 65 year (only 3% of the population)[34]. Illiteracy is widespread[35], gender inequality is pervasive at all levels and many children work and therefore receive little education. For example, studies have reported discrimination against female children in the provision of food (Chen et al. 1981) and in health care seeking behaviours (Hossain and Glass 1988). There is a strong preference for sons in both early and later stages of family formation in Bangladesh (Rahman and DaVanzo 1993). Bangladesh has few natural resources; its manufacturing base is small although it is now beginning to exploit natural gas. In 2004, GNI per capita was around $US440 (World Bank 2005b) of which 40% is generated by agriculture; fishing, tea, and jute are important products.

5.2 Health status indicators

Bangladesh has made considerable progress over the past two decades in improving the health status of the nation. For example, the infant mortality rate has declined from 129 to 46 deaths per 1,000 live births (BBS 1997; BBS 2003). The national coverage rate of immunisations increased from less than 2% in the 1980s to 69% in terms of children aged between 12-23 months completing recommended vaccines (Perry 1999). In addition, each of the twice-yearly National Immunisation Days (NIDs) for oral polio vaccination now reach more than 90% of the 25 million children under five years of age in Bangladesh (EPI 2001).

[34] Bangladesh is one of the few countries in the world (along with India and Pakistan) where the life expectancy at birth is lower for females than for males (BBS 2003).
[35] 59% of the population aged 15 years or older (BBS 2003).

Figure 10: Map of Bangladesh

Source: http://www.un.org/Depts/Cartographic/map/profile/banglade.pdf

More than 80% of children one to five years of age receive vitamin A supplementation every six months; once during one of the two NIDs and once during National Vitamin A week. As a result, it is estimated that the number of children developing nutritional blindness each year has fallen from 30,000 to 6,000 (Perry 1999).

Nevertheless, although as a result of large-scale government programmes there have been notable improvements in some health indicators, health status remains poor. For instance, in the 2001 national Health and Demographic Survey (Mitra et al. 2001), the incidence of low birth-weight babies (2,500g or less) was almost 50% and maternal mortality was estimated to be about 500 per 100,000 live births, one of the highest rates in Asia. Most (85%) deliveries still take place at home, and almost a third of Bangladeshi women report chronic or residual morbidities associated with childbirth. In addition, average nutritional calorie intake was estimated to be 88% of requirements and only 34% of the population had access to adequate sanitation. On the basis of a number of criteria, including a daily calorific intake of only 1,600 per person, it is estimated that approximately half of the rural population lived in absolute poverty in 1998, 44% of whom fell into the category of the very poor.

A high proportion of child deaths are caused by poverty-related infectious diseases and malnutrition, most of which are readily preventable or treatable. The main causes of death, particularly in children, remain diarrhoeal diseases, acute respiratory infections, malnutrition, neonatal conditions, and accidents and injuries.

5.3 Ministry of Health and Family Welfare

The Ministry of Health and Family Welfare (MOHFW) has overall responsibility for health sector policy and planning, and until recently there have been two separate directorates, also called the 'two wings', for health services and family planning (FP). This division of responsibilities between the two directorates was originally established in the early 1970s, since when there has been considerable independence and competition between them (Vaughan et al. 2000).

They are both largely organised into vertical programmes and each has developed separate services, particularly for primary health care at the district (*zila*), sub-district (*upazila*[36]), union and village levels. This separation of services has also led to the development of specialised cadres of health personnel and training institutions, together with separate health facilities, supporting services and information systems (Vaughan et al. 2000). However, in recent years considerable efforts have been made to achieve greater integration by organising more joint services at the upazila level and below.

5.4 Health services

Health care provision in Bangladesh is highly pluralistic with a plethora of treatment options exist. Non-government service provision predominates, which includes both for-profit and non-profit organisations, and traditional and non-formal practitioners. The site of first access for most services, other than maternal and child health (MCH) and FP, is non-governmental, with a wide choice of providers depending on the symptoms, gender, socio-economic standing and geographic location (urban or rural) of the individual. Although allopathic practitioners are consulted in about 80% of cases

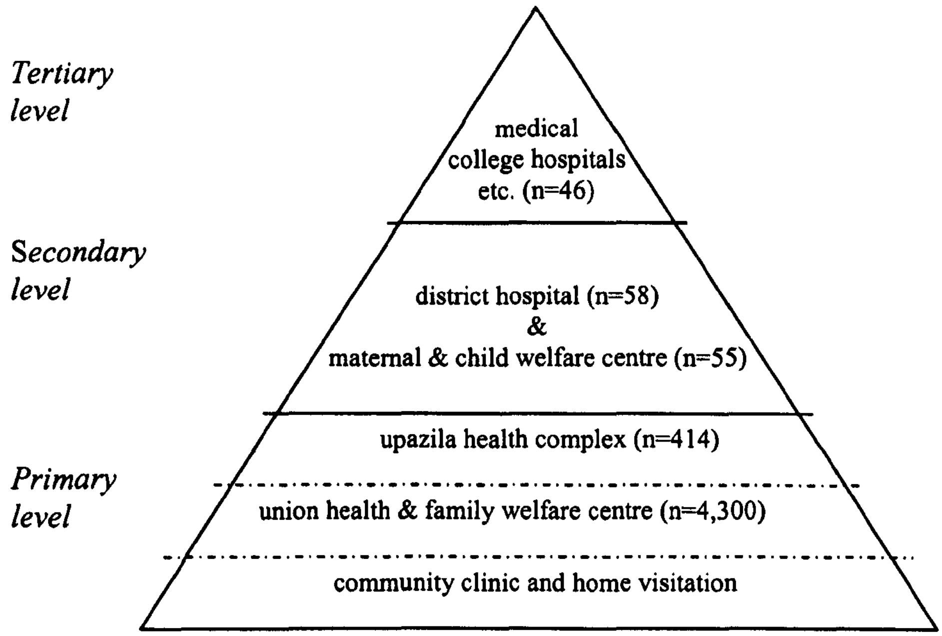
[36] Also known at *thana*.

when treatment is sought (Vaughan et al. 2000), the existence, length and quality of their training is as variable as the treatments they provide.

By contrast, the organisation of Government of Bangladesh (GoB) health services remains highly centralised in the MOHFW and the two directorates in Dhaka. The public health system is structured as a hierarchical pyramid with five layers (Figure 11): one at the tertiary level, one at the secondary level and three at primary level. Bangladesh consists of six Divisions, and tertiary care is provided at this level. Each of the Divisions has Divisional Director's offices of Health and of Family Planning, which manage health services at this level. There are 64 Districts in Bangladesh, each of which has a Civil Surgeon responsible for managing health services at the secondary level. And finally, there are 460 upazilas in Bangladesh each with an Upazila Health Officer and an Upazila Family Planning Officer in charge of respectively preventive and clinical health services, and FP and reproductive health services. These officers manage the delivery of health care at the primary level.

Bangladesh is served by medical colleges, each with around 650 beds, at the tertiary level, and district hospitals (50-200 beds) at the secondary level. The hospital system is over-used at both these levels, with high rates of self-referral, by-passing the sub-district hospitals, known as Upazila Health Complexes (UHCs). There is a high ratio of doctors to nurses in hospitals and potential to improve skill mix, but there is a notable lack of suitable nurses at all levels (Hossain and Begum 1998).

Figure 11: Organisation of health services



There has been a large GoB investment in the rural primary health care infrastructure, with the construction of more than 400 UHCs since the 1970s. The UHCs were established as the cornerstones of primary care. They were created to a standard design, including theatres, X-ray, pharmacy, basic laboratories, dental suites and delivery suites and each has a 31-50 bed ward. Physical facilities have deteriorated in most UHCs and poor staff practices exist in many (e.g. high levels of absenteeism (Chaudhury and Hammer 2004) and informal user-charging[37]). Skilled doctors are unwilling to work there, regarding postings as 'punishment' (Vaughan et al. 2000). As a result UHCs no longer enjoy public confidence and are under-used. The low state salaries earned by doctors have led to growth in private practice. Doctors are thereby diverted from their UHC duties and a vicious circle has evolved whereby their vested interests may wish to keep public sector service quality relatively low.

[37] A study undertaken by the MOHFW found that informal fees are common at all levels of the health system and they can amount to more than ten times the official charges (Killingsworth et al. 1999).

Below the upazila level, there is a network of about 4,300 union health and family welfare centres. Therefore, although geographical access is reasonable for health facilities, they are characteristically under-utilised, particularly at the union-level. Services delivered at this level are commonly perceived to be of poor quality, suffer from shortages of drug supplies and are inefficiently managed (Vaughan et al. 2000). Below the union level, the system depends heavily on community workers who provide FP supplies and provide health advice. Controversially, this door-step approach was to be phased out in favour of services delivered through newly built community clinics serving a population of around 6,000 people (Ensor et al. 2002).

The concept of an essential package of services (ESP) to be delivered in UHCs is well established, although delivery is patchy. The ESP consists broadly of (Ensor et al. 2002):

- reproductive health care, including safe motherhood (essential obstetric care, antenatal and post-natal care), FP, other reproductive services including sexually transmitted disease;
- child health care, including acute respiratory infections, diarrhoeal diseases, vaccine preventable disease and adolescent care implemented through an integrated management of sick child approach;
- communicable disease control, including TB, leprosy, malaria, filarial, kala-azar and emerging diseases;
- limited curative care, concentrating on first aid for trauma, medical and surgical emergencies, asthma, skin diseases, eye, dental and infectious ear disease;

- behaviour change communication is being implemented as a way of influencing health behaviours and health-care-seeking practices across all of the ESP components.

There are also a large number of NGOs that operate separately from the MOHFW (Abbassi 1999). However, there is an increasing tendency for the GoB to contract these NGOs to work in specific under-served areas and / or to carry out service programmes, particularly those for MCH-FP and disease control programmes (Loevinsohn and Harding 2005). Many of the NGOs working in FP have been directly supported by funds from bilateral donor agencies, particularly the United States Agency for International Development (USAID). A number of these NGOs also receive donor funding for primary health care and disease control programmes (Vaughan et al. 2000).

Private practitioners of all kinds, including many medical graduates, are numerous in both urban and rural areas. Drugs are widely available through the large number of private pharmacies and shop outlets. The number of private medical practices and hospitals in urban areas, together with numerous unqualified practitioners, is growing rapidly. These practitioners are poorly regulated and there is no adequate system for registering or licensing them by the GoB. Moreover, financial incentives often militate against medical practice as professional supervision and regulation is weak (Abbassi 1999).

5.5 The National Immunisation Programme

Following the eradication of smallpox, Bangladesh's Expanded Programme on Immunisation (EPI) was started in 1979 but little progress was made until the mid-

1980s when the coverage rate of children fully immunised went from 2% to 62% during the period 1985 to 1990, with higher rates in some divisions and 80% in one area by 1991. This accomplishment was considered so spectacular that it was hailed as the 'Near Miracle' (Huq 1991). Such an achievement was not easy and was brought about by high political commitment, national and international pressure, mobilisation of various stakeholders including civil society and NGOs, and strong donor support. Shortly after the declaration of this spectacular success in 1991, concerns emerged regarding the sustainability of the national immunisation programme (Walker et al. 2000). In fact, in May 1999, immunisation rates were reported to have dropped to 59-62%.

The national immunisation programme in Bangladesh aims to reduce morbidity and mortality from six vaccine-preventable diseases: diphtheria, measles, pertussis, poliomyelitis, tetanus and tuberculosis (TB)[38]. A fully vaccinated child receives six standard EPI antigens through eight vaccinations requiring, in theory, five contacts with health staff: three shots of DPT (which protects against diphtheria, pertussis and tetanus), three doses of OPV[39] (which protects against poliomyelitis), one shot of BCG (Bacillus Calmette-Guérin, which provides protection against TB) and one shot of the measles vaccine. The recommended schedule in Bangladesh is: one dose of BCG at birth[40], three doses of OPV together with three doses of DPT at ages 6, 10 and 14 weeks of age, and one dose of measles at age nine months of age. Pregnant women and those of childbearing age are given two shots of TT (tetanus toxoid) to prevent maternal and neonatal tetanus.

[38] In 2004, hepatitis b vaccination was introduced, which requires three doses be given alongside the DPT schedule.
[39] A neonatal dose, OPV0, is also recommended but rarely administered because most births do not take place in a health facility.
[40] In reality, BCG is given at six weeks of age alongside OPV1 and DPT1.

The programme is run under two different systems; one for the rural and one for the urban areas. In the rural areas, the EPI is the responsibility of the MOHFW. Services are provided at district hospitals, UHCs, union-level clinics (although only when they act as an outreach site) and NGO clinics. In addition to these fixed sites, the programme, unlike in many countries, relies heavily on outreach activities provided by two types of government-paid fieldworkers: health assistants, who provide a range of basic health services through house-to-house visits and vaccination through outreach sessions, and family welfare assistants who mainly deliver family planning services, but also assist in providing EPI. The current EPI strategy is based on a model of conducting monthly outreach sessions through eight outreach sites per ward (which has a total population of approximately 8,000 people). Porters deliver vaccines from the UHCs to distribution points where the field workers collect the vaccines in vaccine carriers, and sterilised needles and syringes in pre-sterilised drums[41], and take them to the designated outreach sites. Almost all people live within 15-20 minutes walking distance of an EPI site. Government EPI outreach sites delivered vaccines to around 80% of all vaccinated children in rural areas, according to the year 2000 National EPI Coverage Evaluation Survey (EPI 2000).

As noted above, a spectacular increase in national coverage was achieved in the 1980s and 1990s. However, because EPI was not a priority of municipal governments in the 1980s, vaccination coverage in the urban areas was found to lag considerably behind that of the rural areas. As a result, a number of donors, mainly USAID and the Asian Development Bank stepped in to fill the gap, developing EPI and child health projects in urban areas starting from around 1988.

[41] Auto-disable syringes have been introduced in a phased fashion alongside the introduction of hepatitis b vaccine.

The EPI programme in the urban areas, which consists of six city corporations and over 200 municipalities accounting for approximately 23% of the total population of Bangladesh, is a complex collaborative effort between municipal authorities, the MOHFW, the Ministry of Local Government, Rural Development and Co-operatives, NGOs and key donors (e.g. the World Bank, Swedish International Development Agency, United Nations Children's Fund and Japanese International Co-operation Agency). EPI services are provided at government clinics and outreach sites by a cadre of vaccinators, as well as by HAs and Vaccination Supervisors. Vaccines for the NGOs are provided by the MOHFW free of charge, through the local municipal authorities. Urban immunisation services are therefore provided by a complex combination of government, NGO, and private providers with little coordination between the various providers.

It is interesting to note that despite the widespread use of private health practitioners, in Bangladesh only a fraction of the population receives immunisations from private-for-profit providers (Levin et al. 1999). Unlike in many countries, even middle-class children receive their EPI immunisations in the public sector, because of the programme's good reputation and the fact that the vaccines are free of charge.

Despite the extensive infrastructure of EPI, especially in the rural areas, and the heavy use of outreach activities as the mainstay of the system, in 2000 only an estimated 53% of children were fully vaccinated by age one (Table 6). However, as illustrated by the high BCG rate (95%), access to the EPI programme is quite good. Unfortunately however, many children do not complete their series of vaccinations – drop-out rates have been rising since 1995 and in 2000 were estimated at 27% from BCG to DPT3 and 33% from BCG to measles (EPI 2001).

Table 6: Results of 2000 national vaccination coverage evaluation survey

Vaccine	Rural	Urban	National
BCG	95%	95%	95%
DPT3	66%	74%	68%
OPV3	66%	74%	68%
Measles	61%	64%	62%
Fully vaccinated child	52%	56%	53%

Source: 2000 National Coverage Evaluation Survey (EPI 2001)

5.6 Health care expenditure

While, total health expenditure in 2002 was US$1.54 billion, equivalent to 3.1% of GDP, total public expenditure on health was US$417 million, which equates to 0.88% of GDP. On a per capita basis, these amounts equate to total expenditure per capita on health of US$11 of which US$3 is public expenditure.

Table 7: Selected indicators of expenditure on health for the year 2002

Indicator	
Total expenditure on health as % of gross domestic product	3.1%
General government expenditure on health as % of total expenditure on health	25.2%
Private sector expenditure on health as % of total expenditure on health	74.8%
General government expenditure on health as % of general government expenditure	4.4%
Private households' out-of-pocket payment as % of private sector expenditure on health	85.9%
External resources on health* as % of total expenditure on health	13.5%
Total expenditure on health per capita	$11
General government expenditure on health per capita	$3

* External resources enter the system as a financing source, i.e. all external resources whether passing through governments or private entities are included under the public or private health expenditures.

The main bilateral donors to the health and population sector in Bangladesh are the governments of Australia, Belgium, Canada, Germany, Japan, Netherlands, Norway, Sweden, the United Kingdom and the United States (Vaughan et al. 2000). The World Bank, European Union, United Nations Children's Fund (UNICEF) and Asian Development Bank are also major donors (Buse and Gwin 1998). Approximately one-third of donor funding was channelled through the Fourth Family Planning and Health Project (FPHP4), which was supported by a consortium of the World Bank and nine bilateral donors, which the Bank had the responsibility for co-ordinating, and operated from 1993-98 (Buse and Gwin 1998). With donor support, SWAps, or sector-wide approaches (Cassells and Janovsky 1998), were adopted for the subsequent Fifth Health and Population Sector programme (HPSP), which ran for 1998-2003, and was thus in operation during the time data for this thesis were collected.

5.7 Recent and current health sector reform programmes

Although donors had periodically encouraged the GoB to adopt a national health policy, it was only in the 1990s that this became a condition of their support.

5.7.1 Health and Population Sector Programme 1998-2003

The HPSP was a five year sector-wide programme of the Ministry of Health and Family Welfare which ran from 1998 to 2003. The HPSP was supported by a consortium of donors, including the World Bank, which led the programme, the Swedish International Development Agency, the Netherlands, the UK Department for International Development, and the European Union. The aims of the strategy were to provide a sustainable universal package of essential services of health care for the people of Bangladesh, and to slow population growth, with an emphasis on client-centred, accessible services, particularly for children, women and the poor. The ESP grew out of recognition that it is not possible to provide all of the services needed by all segments of the population. It included a prioritised list of interventions to be delivered at upazila level and below, with referrals to secondary and tertiary levels also identified (see above). Unfortunately the GoB faces significant resource constraints in funding the ESP (Rannan-Eliya and Somanathan 2003), even though as much as two-thirds of HPSP financing was channelled into the ESP (Ensor et al. 2003b). It has been argued that the potential for additional resource mobilisation is limited and that improvements in the efficiency of health care services must be a critical component of efforts to provide the ESP to the whole population (Rannan-Eliya and Somanathan 2003).

The HPSP also emphasised the integration of the health and FP wings of MOH&FW and the decentralisation of management and financial responsibilities to the district and upazila (sub-district) level. At the central level, the Directorate General of Health Services was re-organised. Beginning in July 1999, EPI was changed from a vertical programme with its own director to one of several programmes in the ESP that is administered by the Director, Primary Health Care and Line Director, ESP. The Programme Manager, Child Health & Limited Curative Care, manages EPI, ARI (acute respiratory infections), CDD (control of diarrhoeal disease), School Health, and Limited Curative Care. A Deputy Programme Manager (EPI) assists the Programme Manager in managing EPI activities. Under HPSP, cold chain, logistics, training, surveillance, and communications are under the authority of various line directors responsible for each of the respective sector areas (e.g., Logistics, Training, Unified Management Information System, and Behavioural Change & Communication).

5.7.2 Health Nutrition and Population Sector Programme 2003-2006

Following on from the HPSP, the Health, Nutrition and Population Sector Programme (HNPSP) was initiated in 2004, including nutrition as a sub-sector. Also focusing on the vulnerable, including the elderly, the HNPSP emphasizes reducing malnutrition, mortality, and fertility, promoting healthy life styles, and reducing risk factors to human health from environmental, economic and social and behavioural causes.

5.8 Summary

- Bangladesh is one of the most densely populated and poorest countries in the developing world. Although there have been notable improvements in some health indicators since independence was achieved in 1971, health status remains poor, thus making health and population among the most important development issues;
- The Ministry of Health and Family Welfare has overall responsibility for health sector policy and planning;
- Health care provision is highly pluralistic and a plethora of treatment options exist with non-government service provision predominating;
- On a per capita basis, total health expenditure in the year 2002 was US$1 of which US$3 was public expenditure;
- At the time data collection took place for this thesis (1999-2003), the GoB was undertaking the HPSP, which focuses on the provision and utilisation of an essential package of services consisting of reproductive health care; communicable disease control; limited curative care; and child health care, under which the national immunisation programme falls;

- The EPI in Bangladesh was established in 1979 and became fully operational in 1985. Increase in coverage was achieved first in rural areas. USAID implemented a programme to strengthen vaccination services in urban areas of Bangladesh in 1988;
- The GoB faces significant resource constraints in funding the ESP. It has argued that the potential for additional resource mobilisation is limited, and that improvements in the efficiency of health care services must be a critical component of efforts to provide the ESP to the whole population.

The next chapter presents the costs of delivering routine vaccination services in Dhaka City.

CHAPTER 6

VARIATION IN THE COSTS OF DELIVERING ROUTINE VACCINATION SERVICES IN DHAKA, BANGLADESH

This chapter presents the costs of delivering routine vaccination services in Dhaka City. After a brief introduction, there are three parts to this chapter. The first focuses on describing the methods used, in particular, the study design, sampling, data collection and analysis. The second part describes the data and results, focussing on: the total and mean cost per delivery unit by type and ownership; number and type of doses administered and wasted by type and ownership of delivery unit; and finally the weighted mean cost per dose by type and ownership of delivery unit. The chapter concludes with a summary of the chapter.

6.1 Introduction

A comprehensive review of the Bangladeshi national immunisation programme in 1998 recommended the need for collecting cost information from urban areas (EPI 1998). While Khan and Yoder (1998) and Levin et al. (1999) both estimated the cost of the national immunization programme, neither reported the costs of the urban component of the programme, choosing rather to use a range of assumptions to extrapolate the costs of the rural component to urban areas.[42]

[42] Lacking data from urban settings, Khan and Yoder (1998) uused the costs of rural personnel as a proxy for those located in urban areas. They used a range of 14% to 25% of the cost of rural personnel as a proxy for the cost of urban personnel, which resulted in a range of $11.58 to $11.96 per FVC (the cost per dose varied from $0.69-$0.71); in total, EPI activities cost about $18 million in 1997-98 prices. A year later, Levin et al. (1999) estimated total cost of the routine national immunisation programme to be $28.9 million resulting in $0.84 per dose and $18.06 per FVC.

While there has been some reports of the cost of providing vaccination services (e.g. Brenzel and Claquin 1994), few studies to date have detailed intra-country variation of these costs (Walker et al. 2004)[43]. Indeed, in the application of CEA of health services, it is rare to see detailed cost analyses across units. Cost data can provide valuable information for national decision-makers and development partners. It can help EPI programme managers to: strengthen national budgeting and planning; identify inefficiencies (e.g. high wastage rates, 'high' cost providers which might indicate inefficiency); and, identify priorities as an input to CEAs. However, the representativeness of reported costs is frequently questionable as they are often based on national estimates of total expenditure or estimates from a few facilities. Hence variation in the expected costs (and benefits) at sub-national levels is often not addressed. Therefore, as noted by the Immunization Financing Database team, "Further work is needed to better understand the sources of variation we find in the cost of immunization programs. Understanding this variability will be extremely useful for future analyses ..." (WHO 2005). Systematic and significant variation in unit costs between production units, can present a powerful basis for benchmarking and for identifying relatively inefficient units.

In particular, the potential bias and inefficiencies involved in transferring data without resolving our understanding of variation could introduce inefficient interventions or halt the provision of efficient interventions. Alternatively variation within and between settings may not exist or may not significantly affect conclusions. It is therefore vital that research continues to assess how serious a problem this is and whether it leads to any systematic misallocation of resources. Different levels of efficiency in programme

[43] Both Khan and Yoder (1998) and Levin et al. (1999) used a top-down approach to costing.

operation within a particular setting would affect the unit costs of providing vaccines, which is the focus of the next chapter.

6.2 Aim and objectives

The overall aim of this chapter is to report and describe variation in the costs, from the perspective of providers, of delivering routine EPI in DCC, Bangladesh. The specific objectives are to:

- estimate the total cost of providing vaccination services and unit cost per antigen administered in DCC using standard costing methods;
- describe variation in these costs across providers;
- rank delivery units from the highest to lowest unit cost per dose administered.

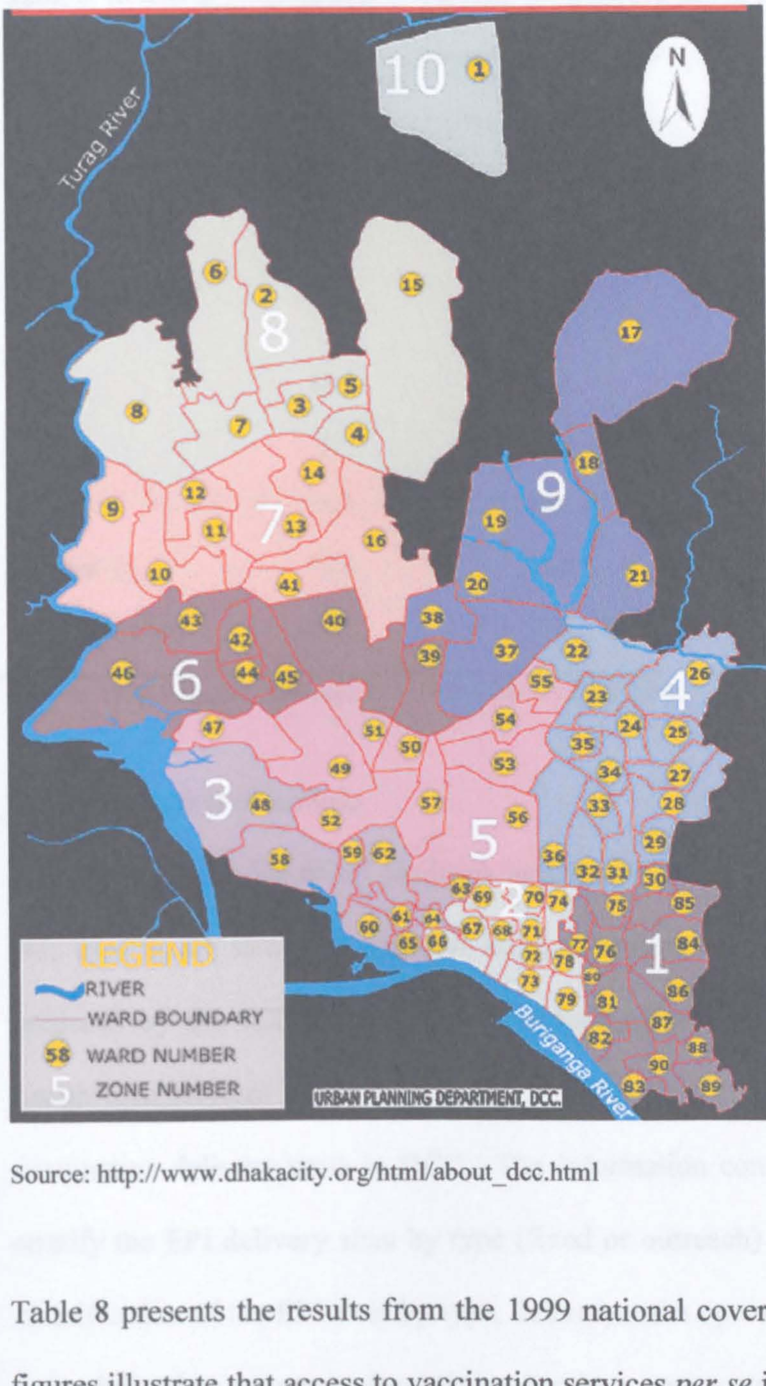
6.3 Methods

6.3.1 Dhaka City Corporation

Dhaka City Corporation is the largest of four city corporations in Bangladesh with an estimated population of 5,622,298 in the year 2000[44]. Rapid population growth rate of 6% has resulted in high population density peaking at 300-600 people per acre in the 'slum' areas of Dhaka. DCC area is divided into 10 administrative zones and 90 wards. Average population in each zone is about 562,229, with the largest population in zone 4 (843,489) and the smallest in zone 10 (325,189). See Figure 12 for a map of Dhaka City.

[44] The estimate of DCC area population for 2000 was calculated using the 1991 Census of Bangladesh. An annual growth rate of 6% was assumed.

Figure 12: Map of Dhaka City Corporation



Source: http://www.dhakacity.org/html/about_dcc.html

Table 8 presents the results from the 1999 national coverage survey for DCC. These figures illustrate that access to vaccination services *per se* is not a problem as evidenced by a coverage rate of 93% for BCG. There is a problem though of ensuring that mothers return with their children to complete the schedule at the appropriate time, as evidenced by the high drop-out rate between DPT1 and DPT3.

Table 8: Results of 1999 national vaccination coverage evaluation survey for Dhaka City Corporation

Antigen	Valid coverage	Crude coverage[45]
BCG	93.3	93.3
OPV1	89.9	93.8
OPV2	78.1	90.4
OPV3	73.7	87.5
DPT1	88.4	92.3
DPT2	79.0	89.9
DPT3	74.6	87.0
Measles	70.0	75.5
Fully vaccinated[46]	60.8	75.5

Source: 1999 National Coverage Evaluation Survey (EPI 2000)

6.3.2 Selection of sample

A comprehensive list of all facilities involved in the delivery of EPI services in DCC was used as the sampling frame to select a random sample of facilities. This list was prepared by the ICDDR,B to better understand the supply environment of primary health care services in Dhaka City (Mazumder et al. 1997). In 1998 there were 511 vaccination delivery units in DCC. The information contained in the list was used to stratify the EPI delivery sites by type (fixed or outreach) and location (zone). For the classification of the EPI sites by type, delivery units operating one day or less per week were defined as outreach sites while all others were categorised as fixed sites. From each of the defined stratum, 25% of facilities were chosen at random. This sampling procedure generated a sample of 132 EPI delivery sites.[47] The classification of health

[45] Percentage of children vaccinated irrespective of the validity of the vaccination or age at administration.
[46] Percentage of children receiving all eight doses.
[47] It should be noted that no power calculation was undertaken to guide this sample size. 132 vaccination delivery units, representing approximately 25% of all units operating at the time in Dhaka City, were selected simply because time and money allowed the project team to do so.

facilities by ownership (government or NGO) could not be carried out prior to sampling as the listing of facilities did not contain this information. However, since the study selected a large proportion of all EPI sites at random (25%), the results of the survey should provide a reasonable indication of the relative importance of GoB and NGO providers of EPI services in urban Dhaka.

Out of the 132 sites surveyed by the study, less than a quarter was GoB-run facilities. About 60% of all sites were NGO-run outreach centres. About 77% of the EPI delivery sites in Dhaka City were under NGO management and these sites organized 60% of all EPI sessions. The predominance of NGOs in the delivery of EPI in urban Bangladesh is in sharp contrast to the delivery structure in rural areas, where it is almost exclusively a publicly-run programme.

6.3.3 Cost analysis

Vaccination services have been costed by the 'ingredients' approach, in which the total quantities of goods and services actually employed in delivering the activities were estimated, and multiplied by their respective unit prices (Creese and Parker 1993). A structured questionnaire was used to collect information on resource use, including expenditure data, and the number of vaccinations administered for the calendar year 1999 (see Appendix 5). This was pre-tested at non-sampled EPI delivery sites. Relevant information was obtained from various sources, including administrative records, interviews and direct observation.

The first part of the instrument collected data on all capital and recurrent resources used in the process of delivering EPI services including donated items such as volunteer time and space provided by communities. More specifically, the resources reviewed included:

- capital items (resources typically with a unit cost greater than US$100 and / or a working life of greater than one year): equipment (e.g. refrigerator and cold boxes), furniture (e.g. tables and chairs) and vehicles;
- recurrent items: staff (e.g. salaries and benefits of staff providing and / or supporting EPI services), rent (including utilities, operating and maintenance), vaccines, supplies (e.g. syringes and ice-packs) and short-term training.

In lieu of obtaining the annualised value of land and buildings, the study collected information on the rent for facilities. If the facility was owned by the provider rather than rented (e.g. GoB facilities), the rental value for the facility was imputed on the basis of the average rent for similar sites in the same location.

Capital costs were annualised using a 3% discount rate and the working life of all EPI-related capital items was assumed to be five years. Joint (or shared) costs were apportioned to EPI on the basis of the proportion of time / space used for EPI activities. All figures are presented in 1999 US dollars using the average official exchange rate between January 1999 – December 1999[48].

The second part of the questionnaire collected information on other variables related to EPI services such as the number of sessions per month and year, the duration of these sessions and number of vaccines administered and wasted per session, month and year.

[48] 1US$ = 49.50 Bangladeshi Taka

The unit cost of providing each vaccine was calculated in the following way:

- the cost of the vaccine and the syringe (except for OPV which is administered orally) was assigned directly to each vaccine;
- personnel, remaining recurrent items and all capital items with the exception of the cold-chain were distributed on the basis of the number of visits;
- the costs of the cold-chain were distributed according to the vaccine doses administered.

The calculation of number of visits took the following into account:

- OPV and DPT vaccine doses were assumed to be administered at the same visits;
- other vaccines, i.e. BCG, measles and TT are administered at separate visits.

The weighted mean unit vaccine costs have been calculated using the number of vaccines administered as the weights. Also estimated and reported is the wastage rate, where the vaccine wastage is the proportion of vaccine supplied, but not administered to children, usually stated as a rate and calculated as:

vaccine wastage rate = ([doses supplied - doses administered] / doses supplied) x 100.

The cost per fully vaccinated child (FVC), as defined by the schedule, was also estimated, e.g. a child that received one dose of BCG, three doses of OPV, three doses of DPT and one dose of measles vaccine by their first birthday.

6.3.4 Missing data

Delivery units with missing or 'incorrect' values were excluded from the analysis. The final data set consisted of 110 out of a possible 132 delivery units. Hence, 83.3% of all delivery units in the sample were included. Table 9 presents the total number of delivery units included in the final sample, split by location (zone), ownership (GoB or NGO) and type of delivery unit (fixed or outreach).

Table 9: Final sample of vaccination delivery units

Zone	GoB		NGO		Total
	Fixed	Outreach	Fixed	Outreach	
1	3	0	2	6	11
2	5	1	0	0	6
3	2	0	0	2	4
4	1	0	4	11	16
5	4	0	2	4	10
6	0	1	2	6	9
7	0	0	2	14	16
8	1	1	2	11	15
9	0	2	1	16	19
10	1	0	0	3	4
Total	17	5	15	73	110

The type of data missing included: ownership form; duration of operation; and some of the inputs and outputs. 'Incorrect' values were identified after eye-balling the data. For example, where data indicated that a delivery unit had administered only one dose of each antigen it was excluded from the final sample on the grounds that this did not seem realistic, or plausible, given that the data were collected for the year 1999. This would suggest either that the data were entered incorrectly or that the delivery unit had only operated for a very brief period of time during 1999. On some occasions outreach delivery units recorded administering a single dose of BCG and / or measles (n=13). While these data are suspicious, they are not entirely implausible given that both vaccines require a single dose to give protection against their respective diseases. Furthermore, a significant drop-out rate between DPT3 (given at 14 weeks of age) and measles (given at nine months of age) has been documented in coverage surveys in Dhaka and Bangladesh more generally. One delivery unit was excluded because a value of one was recorded for OPV when the other vaccines had values of 840, 360, 360 and 360 for BCG, DPT, measles and TT respectively. As it was not possible to identify the cause of these 'incorrect' values, a decision to exclude them was taken. There were no systematic instances of missing data, i.e. missing data were evenly distributed across the zones and types of providers.

6.4 Results

6.4.1 Total and mean cost per delivery unit, by type and ownership

The total and mean cost per delivery unit by type and ownership is shown in Table 10. Total annual cost of routine EPI services in the surveyed EPI delivery sites was found to be \$197,583. The mean cost of running a vaccination delivery unit was \$1,796 per year. However, mean costs vary by ownership type, most markedly among outreach units where the annual mean cost for GoB outreach sites was \$2,867 compared to \$1,070 for NGO outreach sites. The annual mean cost of fixed sites was \$3,328 and \$3,228 for GoB and NGO sites respectively.

Table 10: Total and mean cost per vaccination delivery unit, by type and ownership, in 1999 US$

	Type of facility									
	GoB fixed (n=17)		GoB outreach (n=5)		NGO fixed (n=15)		NGO outreach (n=73)		Total (n=110)	
Item	Total cost	Mean cost per facility	Total cost	Mean cost per facility	Total cost	Mean cost per facility	Total cost	Mean cost per facility	Total cost	Mean cost per facility
Capital items										
Vehicles	0	0	0	0	0	0	300	4	300	3
Equipment	1,187	70	124	25	2,102	140	2,485	34	5,897	54
Furniture	631	37	102	20	1,221	81	881	12	2,836	26
Subtotal	1,817	107	222	44	3,323	222	3,366	46	8,728	79
Recurrent items										
Personnel	38,554	2,268	11,122	2,224	31,187	2,079	54,239	743	135,103	1,228
Rent	4,588	270	481	96	6,288	419	3,420	47	14,777	134
Vaccines	10,797	635	2,237	447	6,468	431	13,512	185	33,013	300
Supplies	629	37	88	18	543	36	1,496	20	2,756	25
Training	191	11	187	37	763	51	1,765	24	2,906	26
Subtotal	54,759	3,221	14,115	2,823	45,249	3,017	74,432	1,020	188,554	1,714
Total	56,576	3,328	14,337	2,867	48,572	3,238	78,098	1,070	197,583	1,796

6.4.2 Output of the delivery units

Table 11 reports a range of output measures. The surveyed delivery sites provided an average of 2,232 vaccinations during 76 sessions per year or 34 vaccinations per session. Each session lasted for an average 4.2 hours, thus providing an average of nine vaccinations per hour. OPV doses are the most common type of vaccine provided by all delivery units, followed by DPT, whereas doses of BCG and measles are the least regularly provided. The NGO and GoB outreach sites administer the least number of vaccinations, whilst the GoB fixed sites provide the most number of doses per annum.

Table 11: Mean number of vaccine doses administered by type and ownership

		Type of facility				
		GoB fixed (n=17)	GoB outreach (n=5)	NGO fixed (n=15)	NGO outreach (n=73)	Total (n=110)
Vaccines		4,462	3,437	3,493	1,370	2,232
	BCG	534	300	347	186	267
	DPT	1,232	1,085	1,008	350	609
	OPV	1,552	1,248	1,080	461	750
	Measles	491	334	210	118	198
	TT	654	470	847	256	408
Number of sessions		172	48	138	43	76
Vaccinations per session		29	72	30	33	34
Duration of session (hours)		4	3	6	4	4
Vaccinations per hour		8	21	7	9	9

Table 11 illustrates that mix of vaccines provided varied systematically across delivery units. There could be several reasons for this. First, there has been a worldwide mass campaign to eradicate polio for many years and therefore more people could be aware of the benefits of the polio vaccine and consequently demand is higher for this vaccine *vis-à-vis* the other vaccines available. In addition, the schedule requires three doses so it is perhaps not surprising that this is the most common vaccine administered. Similarly, it is not surprising to note that the BCG and measles vaccines are the least provided when they only require one dose each. However, it is interesting to note that DPT and OPV doses are supposed to be delivered together but the number of DPT doses delivered was about 19% lower than for OPV (Table 12).

Table 12: Drop-out rates between BCG and measles, and DPT3 and measles

	GoB fixed	GoB outreach	NGO fixed	NGO utreach	Total
BCG	534	300	347	186	267
DPT3	411	362	336	117	203
Measles	491	334	210	118	198
Measles/BCG	92%	111%	61%	63%	74%
Measles/DPT3	119%	92%	63%	101%	97%

6.4.3 Unit costs

Table 13 presents the mean number of vaccine doses administered, wastage, weighted cost per antigen and the cost per FVC by type and ownership of delivery unit. Vaccine wastage rates are highly variable across delivery units. For example, the BCG wastage rate among NGO outreach sites was 43% compared to 67% among GoB outreach sites. Wastage rates are highest for BCG (43 – 67%) and lowest for DPT (12 – 30%). The weighted mean cost per dose administered across type and ownership varied most for measles ($1.23 – $2.55) and least for DPT ($0.47 – $0.58). The cost per FVC ranged from $5.20 – $7.56.

Table 13: Mean number of doses administered, wastage and weighted mean cost per dose and fully vaccinated child, by type and ownership, in 1999 US$

	Type of facility											
	GoB fixed (n=17)			GoB outreach (n=5)			NGO fixed (n=15)			NGO outreach (n=73)		
Vaccine	No. of doses administered	Wastage (%)	Cost per dose	No. of doses administered	Wastage (%)	Cost per dose	No. of doses administered	Wastage (%)	Cost per dose	No. of doses administered	Wastage (%)	Cost per dose
BCG	534	55	1.08	300	66	1.54	347	60	1.52	186	43	0.96
DPT	1,232	27	0.51	1,085	12	0.47	1,008	20	0.58	350	30	0.57
OPV	1,552	33	0.46	1,248	43	0.49	1,080	34	0.59	461	33	0.49
Measles	491	46	1.23	334	43	1.47	210	63	2.55	118	48	1.58
FVC	-	-	*5.20*	-	-	*5.87*	-	-	*7.56*	-	-	*5.72*
TT	654	47	0.79	470	42	0.92	847	27	0.59	256	38	0.66

A second method was used for calculating the cost per FVC. Based on a crude birth rate of 20.4 per 1,000 population there were estimated to be 114,695 births in the year 2000.[49] The infant mortality rate of 51 per 1,000 live births was assumed, and most of these deaths occur early in the first year.[50] Thus, about 95% of live births (108,845) were assumed to survive to the recommended age of vaccination (9-12 months). The cost per fully vaccinated child was determined by dividing 60.8% (see Table 8) of the number of children under the age of one year (66,178) by the cost of the programme (4.65[51] x $197,583 = $917,862). This yielded a cost of $13.87 per fully vaccinated child.

Table 14 provides a breakdown of the weighted mean cost per dose administered by type and ownership of delivery unit. The mean fixed cost per dose (comprising capital items and salaries, which are fixed in the short-term) accounted for between 71 – 79% of the total mean cost per dose administered. The introduction of newer, more expensive vaccines such as those against hepatitis B, *Haemophilus influenze* type b and rotavirus would reduce the proportion of fixed costs.

---

[49] The estimate of the crude birth rate for 2000 was calculated using the Bangladesh Bureau of Statistics estimate for 1996 (BBS 1997). It was assumed that trends observed from 1993-97 continued.
[50] The estimate of the infant mortality rate for 2000 was calculated using the Bangladesh Bureau of Statistics estimate for 1996 (BBS 1997). It was assumed that trends observed from 1993-97 continued.
[51] The total number of vaccination delivery units in DCC (n=511) divided by the total number of vaccination delivery units in the sample (n=110).

Table 14: Weighted mean cost per vaccine delivery per dose, by type and ownership, in 1999 US$

	Type of facility									
	GoB fixed (n=17)		GoB outreach (n=5)		NGO fixed (n=15)		NGO outreach (n=73)		Total (n=110)	
Item	Cost per dose	% of total cost	Cost per dose	% of total cost	Cost per dose	% of total cost	Cost per dose	% of total cost	Cost per dose	% of total cost
Capital items										
Vehicles	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.38	0.00	0.15
Equipment	0.02	2.10	0.01	0.87	0.04	4.30	0.02	3.18	0.02	2.98
Furniture	0.01	1.12	0.01	0.71	0.02	2.15	0.01	1.13	0.01	1.44
Subtotal	0.03	3.21	0.01	1.58	0.06	6.45	0.04	4.69	0.04	4.42
Recurrent items										
Personnel	0.51	68.15	0.65	77.58	0.60	64.52	0.54	69.45	0.55	68.38
Rent	0.06	8.11	0.03	3.36	0.12	12.90	0.03	4.38	0.06	7.48
Vaccines	0.14	19.08	0.13	15.60	0.12	12.90	0.14	17.30	0.13	16.71
Supplies	0.01	1.11	0.01	0.61	0.01	1.07	0.01	1.92	0.01	1.39
Training	0.00	0.34	0.01	1.30	0.01	1.07	0.02	2.26	0.01	1.47
Subtotal	0.72	96.79	0.82	98.45	0.86	93.55	0.74	95.31	0.77	95.43
Total	0.75	100.00	0.83	100.00	0.94	0.03	0.78	100.00	0.80	100.00

Due to rounding some items may appear to account for zero cost.

Table 15 presents the unit cost per dose for all 110 vaccination delivery units. The weighted mean cost per dose administered was $0.80. The unit cost per dose ranged from $0.20 - $ 7.99; a 40-fold difference.

Table 15: Cost per dose of individual vaccination delivery units, in 1999 US$

ID#	Cost per dose	ID#	Cost per dose	ID#	Cost per dose	ID#	Cost per dose
1	0.48	29	0.60	57	**7.99**	85	1.77
2	2.23	30	0.41	58	1.82	86	0.54
3	0.77	31	0.27	59	1.16	87	1.65
4	0.80	32	1.53	60	2.29	88	1.03
5	0.24	33	1.51	61	3.38	89	1.43
6	0.23	34	0.97	62	1.03	90	1.76
7	0.76	35	0.45	63	1.30	91	0.37
8	1.01	36	0.48	64	0.44	92	0.64
9	0.22	37	4.05	65	0.60	93	0.86
10	0.42	38	2.21	66	0.87	94	1.46
11	3.33	39	4.51	67	0.77	95	0.30
12	0.66	40	1.03	68	1.59	96	0.82
13	1.06	41	0.34	69	0.79	97	2.03
14	1.04	42	1.00	70	0.27	98	0.77
15	0.52	43	0.60	71	0.69	99	2.10
16	0.45	44	1.26	72	1.98	100	0.40
17	0.81	45	0.92	73	0.48	101	0.67
18	**0.20**	46	0.65	74	1.94	102	1.77
19	0.48	47	3.27	75	1.37	103	1.35
20	1.74	48	4.37	76	5.45	104	3.67
21	0.39	49	2.12	77	2.76	105	0.37
22	2.43	50	4.66	78	1.74	106	1.48
23	1.04	51	3.11	79	1.32	107	0.76
24	0.47	52	1.57	80	1.88	108	6.02
25	0.49	53	2.20	81	0.69	109	3.99
26	0.43	54	1.80	82	1.38	110	0.71
27	2.04	55	0.65	83	0.72		
28	1.08	56	1.46	84	0.79		

Variation in the unit costs can be explained, in part, by the volume of output at each delivery unit and the wastage rates (Figures 13 and 14). The relationships suggest that as the number of vaccines administered increases the unit cost decreases, and that the unit costs are lower when wastage rates are low. In general there is a positive correlation between the cost per vaccine dose and wastage (0.49), and this relationship is significant at the 0.01 level. Conversely there is a negative correlation between the unit cost and output (-0.39), which is also significant at the 0.01 level.

Figure 13: Relationship between wastage and cost per dose in the vaccination delivery units

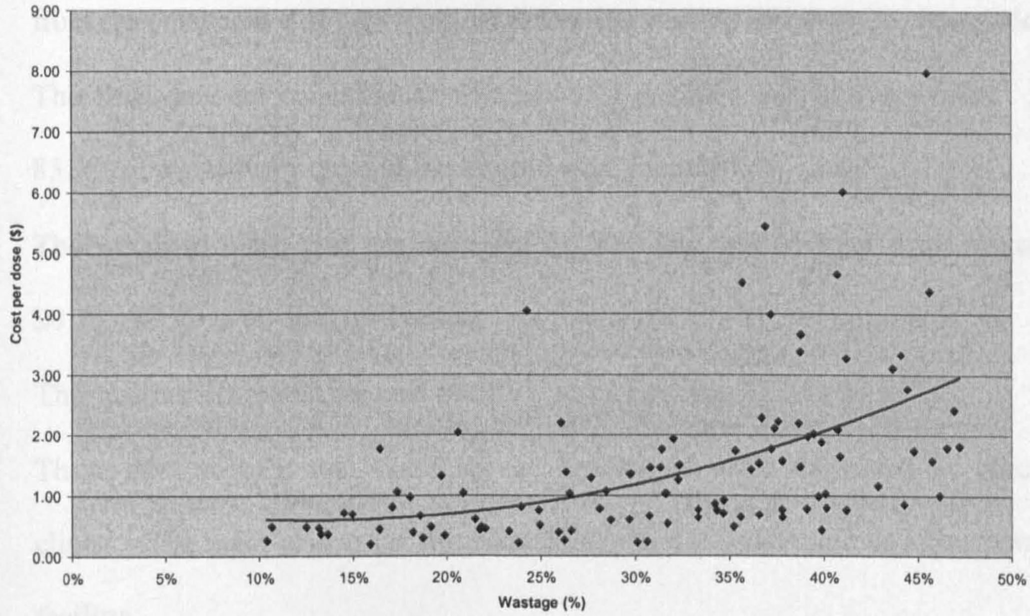
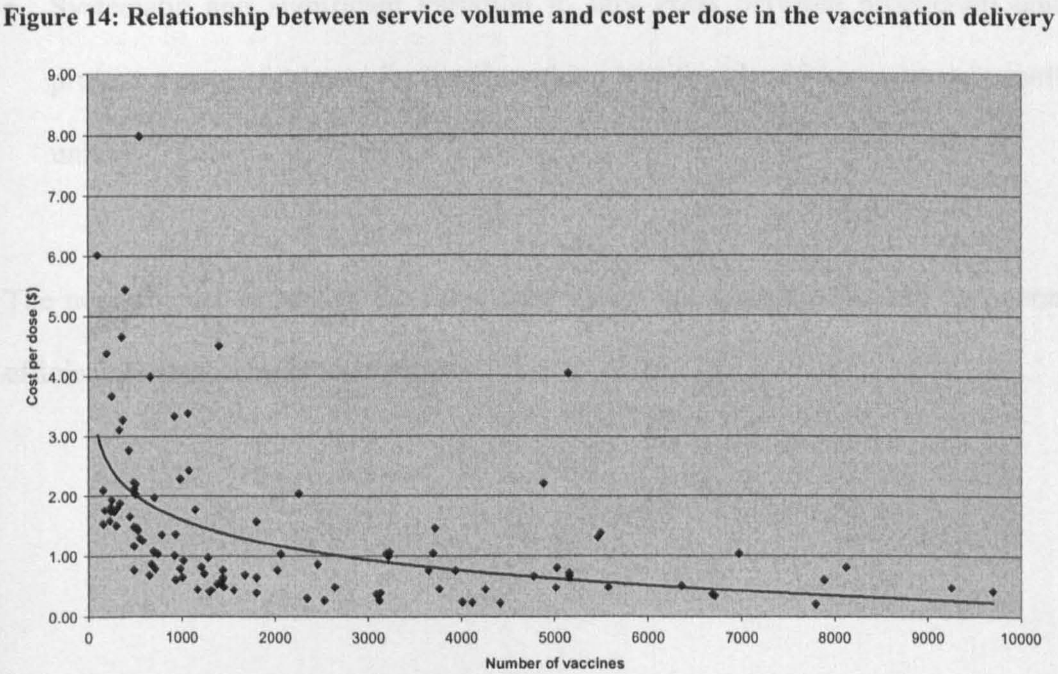


Figure 14: Relationship between service volume and cost per dose in the vaccination delivery units



6.5 Summary

- DCC is the largest of six city corporations in Bangladesh. Rapid population growth rate of 6% has resulted in high population density peaking at 300-600 people per acre in the 'slum' areas of Dhaka;
- This is the first study to report the costs of delivering vaccines in urban Bangladesh. The overall aim of this chapter was to report and describe variation in the costs, from the perspective of providers, of delivering routine EPI in DCC, Bangladesh;
- The final data set consisted of 110 out of a possible 132 delivery units. Hence, 83.3% of all delivery units in the sample were included;
- The weighted mean cost per dose was \$0.80. The unit cost per dose ranged from \$0.20 - \$7.99; a 40-fold difference;
- This chapter estimated the cost per FVC to be between \$5.20 - \$13.87;
- These data suggest that there are economies of scale attributed to vaccination clinics. The main reason for this relationship is the large fixed cost component per facility;
- Systematic and significant variation in unit costs between production units can present a powerful basis for benchmarking and for identifying relatively inefficient units.

The next chapter examines the same data by applying parametric and non-parametric efficiency measurement techniques.

Chapter 7

DATA ENVELOPMENT AND STOCHASTIC FRONTIER ANALYSIS OF VACCINATION SITES IN DHAKA CITY CORPORATION

This chapter examines the efficiency of the same vaccination delivery units presented in Chapter 6, derived by DEA and SFA. After a brief introduction, there are six parts to this chapter. The first describes the specific aim and objectives of this chapter. The second provides a description of the data used in the subsequent analyses. The third focuses on selecting the model specifications, in particular the inputs and outputs for the analyses. In total, nine specifications were included in this analysis; three for the DEAs and six for the SFAs. The fourth part provides an overview of the performance of vaccination delivery units based on the nine specifications, focussing on a summary of the efficiency scores, and for the DEA specifications, the number of efficient delivery units and number of times these efficient units act as peers for inefficient units. This fourth section of the chapter also includes a discussion of two issues: changing the number of inputs and outputs; and the stability of efficiency scores and rankings across specifications. The fifth section presents some policy implications of the results, in particular the level of potential savings and targets for improved performance. The chapter concludes with a summary.

7.1 Introduction

The idea of an 'efficient' health facility is derived from the neoclassical production model in which health care providers choose the mix of inputs that minimise cost with a given demand. Under certain circumstances this is a reasonable characterisation of the behaviour of some privately owned firms. However, cost-minimisation is only one among many possible objectives of the public sector. The existence of multiple goals may lead to compromises between, for example, improving access and minimising cost[52]. This may produce outcomes that are observationally equivalent to, but nonetheless different from, 'inefficient'. Furthermore, the specific incentives and constraints facing the public sector may lead to managerial behaviour that is actually inconsistent with cost-minimisation, for example, satisfaction[53]. Thus, in the context of the particular institutions within which public providers operate, 'efficient' production may not be a realistic policy goal. Rather the objective should be to improve efficiency. One way to do this is to identify those facilities that are performing relatively better than others. The factors that are associated with these performance differentials can then be identified, and interventions developed which can help bring the performance of the 'worst' facilities closer to that of the 'best' ones (Somanathan et al. 2000).

The previous chapter illustrated that the unit cost per dose delivered in a sample of 110 vaccination delivery units in Dhaka City ranged from \$0.20 - \$7.99; a 40-fold difference. Systematic and significant variation in unit costs between production units can present a powerful basis for benchmarking and identifying relatively inefficient units. It is useful to know the level of congruence between unit cost data and efficiency scores obtained through the use of parametric and non-parametric efficiency measurement techniques.

---

[52] Due to the randomness of demand, on any given day the administrator of a health facility cannot predict with perfect certainty the number of individuals who will demand services. Without an active appointment schedule process, there is no way for staff to control the stochastic demand of potential patients (Dervaux et al. 2003). Hence, in order to vaccinate a certain number of children, some amount of excess capacity is required.

[53] In 1966 Harvey Leibenstein published his seminal paper on X-efficiency, which allowed for non-maximising behaviour (Leibenstein 1966).

Asking whether vaccination services are technically and scale efficient is important for a number of reasons. First, identifying sources of inefficiency in a programme may yield helpful insights to potential cost reductions. For example, some programmes may utilise too many inputs to produce outputs. By identifying this technical inefficiency, the programme could reduce input levels while maintaining output production at a lower cost. This would by definition make the programme more cost-effective. Second, it offers the chance to question the scale efficiency of existing services, with a view to recommending the appropriate size of delivery units.

7.2 Aim and objectives

The overall aim of this chapter is to assess the efficiency of routine vaccination services in DCC. The specific objectives are to:

- estimate the efficiency of a sample of vaccination delivery units using DEA and SFA;
- compare and contrast the results of the DEAs, SFAs and unit costs obtained in the preceding chapter;
- explore the effects of different specifications, e.g. changing the number of inputs / outputs, on efficiency scores and ranks;
- for the DEA models, decompose technical efficiency into 'pure' technical and scale efficiency;
- for the SFA models, compare the impact of applying different weights to outputs when aggregating output into a single measure;
- investigate possible causes for differences in the efficiency scores among the sample of vaccination delivery units using a selection of environmental variables;

- consider the policy implications of the results, in particular the potential savings and related corollary, targets for improvement.

7.3 Data

Chapter 6 presented the results of a cost analysis of vaccination services in DCC. These data also provide the opportunity to assess the efficiency of EPI provision. As stated previously, the final data set consisted of 110 out of a possible 132 delivery units. Hence, 83.3% of all delivery units in the sample were included.

For each delivery unit, inputs are defined as the number of full time equivalent (FTE) medical staff, size of the facility dedicated to the delivery of vaccines services (in square feet), the annual total number of hours of operation and the total annual cost. Six outputs were defined for each delivery unit: the number of doses of BCG, DPT, OPV, measles and TT vaccines administered in 1999, and the total number of all types of vaccines administered during the same period, to children less than five years of age and pregnant women.

Table 16 contains the descriptive statistics of inputs and outputs. OPV vaccine was the most common type of vaccine provided, whereas the measles vaccine was the least regularly provided. The inputs and outputs are highly skewed. See Appendix 6 for information on the location, type of ownership and type of vaccination delivery unit for each of the 110 vaccination delivery units.

Table 16: Descriptive statistics of inputs and outputs for 110 vaccination delivery units

Variable	Mean	Std. Dev.	Min	Max
Inputs				
labour	2.98	3.14	1	20
size of facility dedicated to delivery of EPI services	1,805	2,745	35	13,068
total hours[54]	362	410	30	2,700
total cost	2,075	2,379	238	15,077
Outputs				
BCG	267	305	1	1,680
DPT	609	699	24	3,264
OPV	750	856	48	3,756
Measles	198	214	1	960
TT	408	452	12	2,208
total number of vaccines[55]	2,232	2,275	98	9,696

7.4 Methods

7.4.1 DEA models

Given the Government's stated objective to mobilise additional resources via improvements in the efficiency of health facilities (see Chapter 5), input-orientated specifications under VRS have been adopted for each model, which considers what reduction in inputs is possible given existing levels of outputs. However, it should be noted that this specification runs contrary to the Government's stated objective of increasing routine DPT3 vaccination coverage by 12 months of age to 90% in each district by 2005.

[54] Annual number of sessions x hours per session.
[55] BCG, OPV, DPT, measles and TT.

For each model, the technology was initially constructed under CRS and strong disposability of inputs, TE_{CRS} (as inputs increase, outputs must increase, *ceteris paribus*). Allowances were made in the constraints to allow for VRS technical efficiency (TE_{VRS}). Further, the type of scale inefficiencies was determined by employing a third model, TE_{NIRS}. In all these cases the definitions given by Färe et al. (1994) were followed, which were described in more detail in Chapter 3. The DEAP programme by Coelli (1996a) was used for the computations.

The linear programming problems are presented below[56].

Linear programming problem 1: CRS technology

$F_Q(TE_{CRS}) = \min \lambda$

s.t. $u^j \leq zM$

$\lambda q \geq zQ$

$z \in \Re^j_+$

where Q is total costs, u is the outputs of each vaccination delivery unit "j", M is the matrix of outputs, i.e. the vaccines, q is the input costs and z is the intensity variable applied to costs and the outputs.

In order to allow for VRS, a second linear programming problem is solved.

[56] The linear programming problems presented here are where total cost is the sole input. See Chapter 3 for the general specifications.

Linear programming problem 2: VRS technology

$$F_Q(TE_{VRS}) = \min \lambda$$

s.t. $u^j \le zM$

$$\lambda q \ge zQ$$

$$z \in \Re_+^j$$

$$\sum_{j=1}^{J} z_j = 1$$

The constraint on the z vector in the second linear programming problem allows the data to be enveloped more closely which in turn permits VRS to be exhibited. If the solutions to the two linear programming problems are equivalent then the technology is said to be operating at a cost, as well as a scale, efficient level. However, if they are not equal, to what extent inefficiency is caused due to operating at the wrong scale can be determined. Determining the type of scale inefficiency (either increasing returns to scale or decreasing returns to scale) requires the solution of a third linear programming problem, referred to as non-increasing returns to scale technology (NIRS).

Linear programming problem 3

$$F_Q(TE_{NIRS}) = \min \lambda$$

s.t. $u^j \le zM$

$$\lambda q \ge zQ$$

$$z \in \Re_+^j$$

$$\sum_{j=1}^{J} z_j \le 1$$

In order to define the type of scale inefficiency that is operating here, the solutions of the three linear programming problems are compared. If $\frac{TE_{CRS}}{TE_{VRS}} < 1$ and $TE_{CRS} = TE_{NIRS}$ then increasing returns to scale exist. If $\frac{TE_{CRS}}{TE_{VRS}} < 1$ but, $TE_{NIRS} > TE_{CRS}$, then decreasing returns to scale exist. If $TE_{CRS} = TE_{VRS}$ then by definition the vaccination delivery unit is operating under CRS. Using these models, the impact of scale effects on the delivery units can also be examined.

Three DEA specifications were chosen, with outputs ranging from one to five and inputs ranging from one to three (see Table 17). As stated previously, a rule of thumb commonly used with DEA suggests that the number of observations in the data set should be at least three times the sum of the number of input and output variables (Cooper et al. 2003), i.e. for model DEA2 which has the most inputs (n=3) and outputs (n=5), the data set should contain at least 24 observations (3 x [3 + 5] = 24). An alternative rule of thumb suggested by Dyson et al. (2001) states that the number of observations should be at least twice the product of the number of inputs and outputs, i.e. model DEA2 should be run with a data set containing at least 30 observations (2 x 3 x 5 = 30). According to either of these rules of thumb and the specifications chosen, the final sample size of 110 vaccination delivery units is acceptable.

Appendices 7-9 present the data sets used for models DEA1, DEA2 and DEA3.

Table 17: DEA specifications

Specification	DEA1	DEA2	DEA3
Inputs			
labour		✓	✓
facility size dedicated to EPI		✓	✓
total hours		✓	✓
total cost	✓		
Outputs			
BCG	✓	✓	
DPT	✓	✓	
OPV	✓	✓	
Measles	✓	✓	
TT	✓	✓	
total number of vaccines			✓

7.4.2 SFA models

Frontier Version 4.1 (Coelli 1996b) was used to estimate a Cobb-Douglas[57] production frontier assuming a half-normal distribution:

$$\ln(y_i) = x_i\beta + v_i - u_i,\ i = 1, \ldots, n$$

where:

$\ln(y_i)$ = logarithm of the production of the i th firm

x_i = a vector of the logarithm of the input quantities of the i th firm

β = a vector of unknown variables

v_i = assumed to be independent and identically distributed normal random (stochastic) variables with mean zero and constant variance, σ_v^2 ($N[0, \sigma_v^2]$), and independent of the u_i

[57] The frontier programme estimates models which are linear in parameters. Hence to estimate a Cobb-Douglas production frontier, the logarithms of the sample data were estimated.

u_i = non-negative random variables which are assumed to account for technical inefficiency in production and often assumed to be independent and identically distributed $|N(0, \sigma_v^2)|$

Frontier Version 4.1 (Coelli 1996b) was also used to estimate a translog[58] production frontier assuming a truncated normal distribution:

$\ln(y_i) = x_i\beta + v_i - u_i, i = 1, \ldots, n$

where:

$\ln(y_i)$, x_i, β and v_i are as defined above, and u_i has truncated normal distribution. The truncated normal distribution is a generalisation of the half-normal distribution. It is obtained by the truncation at zero of the normal distribution with mean, μ, and variance, σ^2. If μ is pre-assigned to be zero, then the distribution is the half-normal.

A general SFA specification was adopted in which the total number of vaccines was the output and three inputs (labour, size of the facility dedicated to the delivery of vaccination services and total hours) were considered. A limitation of SFA is that it's only well-developed for single-output technologies, or where it is acceptable to aggregate output into a single measure. Therefore, another objective of this chapter was to compare the impact of applying different weights to outputs when aggregating output into a single measure. Three different approaches for weighting outputs are compared: a unit weight applied to all outputs; weights inferred by the price of the vaccines; and weights inferred by the public health importance of the vaccine-preventable diseases in question (using DALYs as the indicator of importance).

[58] The transcendental logarithmic function allows a wide range of non-linear models to be expressed in linear form. It includes the logarithm of every explanatory variable, as well as their products and cross-products.

Therefore, as both a Cobb-Douglas production frontier assuming a half-normal distribution and a translog production frontier assuming a truncated normal distribution were estimated, a total of six different models were run. Appendix 10 presents the data sets used for models SFA1-SFA6.

Table 18: SFA specifications

Method of aggregating outputs	Cobb-Douglas production frontier	Translog production frontier
Unit weights	SFA1	SFA2
Weights defined by price	SFA3	SFA4
Weights defined by DALYs	SFA5	SFA6

7.4.3 Analysis of environmental variables

The ANOVA[59] test was conducted in order to test the null hypotheses that the mean technical, 'pure' technical and scale efficiencies of the delivery units are the same across the:

- 10 locations (zones 1 – 10);
- two types of ownership of the delivery units (GoB or NGO);
- two types of delivery unit (fixed or outreach); and
- type of ownership and delivery unit (fixed GoB, fixed NGO, outreach GoB, outreach NGO)

against the alternative hypotheses that they differ from one another. As the ANOVA test requires the population variances to be equal, the results derived from this test alone may not be valid. Therefore, the Kruskal-Wallis test, the non-parametric version of the

[59] The One-Way ANOVA procedure produces a one-way analysis of variance for a quantitative dependent variable by a single factor (independent) variable. Analysis of variance is used to test the hypothesis that several means are equal (Altman 1991).

ANOVA test (Altman 1991), was also performed, which does not require any assumptions regarding the normality or variances of the populations.

A correlation coefficient is an index that quantifies the linear relationship between a pair of variables (Altman 1991). The coefficient takes values between -1 and +1, with the sign indicating the direction of the relationship and the numerical magnitude of its strength. Values of -1 and +1 indicate that the sample values fall on a straight line. A value of zero indicates the lack of any linear relationship between the two variables. The Pearson correlation coefficient, and two non-parametric correlation coefficients were estimated: Spearman's rho, which is a rank correlation coefficient, and Kendall's tau statistic, which likewise measures the correlation between two sets of rankings. Correlation coefficients were estimated between the efficiency scores and the number of years the delivery unit had been open, population density, male / female ratio and literacy. Table 19 presents the descriptive statistics of these variables. In can be seen that the vaccination delivery units had been open for slightly more than five years on average. The mean population density was 51,572 people per square km, mean male-to-female ratio was 1.34 and the mean literacy rate was 62%.

Table 19: Descriptive statistics of the environmental variables[60]

Environmental variable	Mean	Std. Dev.	Min	Max
Years delivery site has been open since 2000	5.20	4.55	1	22
Population density	51,572	42,180	4,555	168,181
Male / female ratio	1.34	0.25	1.08	2.54
Literacy	62.0	9.3	39.2	74.5

[60] Statistics for population density, male / female ratio and literacy are based on the ward statistics where the sites are located.

7.5 Results

7.5.1 DEA models[61]

Table 20 presents the efficiency results and shows that mean technical efficiency of the vaccination delivery units was 0.35, 0.41 and 0.32 in specifications DEA1, DEA2 and DEA3 respectively. In other words, under DEA1, if the vaccination delivery units were technically efficient and operated at the correct scale, costs could be reduced by 65% without sacrificing the current level of outputs produced. By decomposing this technical measure into 'pure' technical efficiency (TE VRS) and scale efficiency, it can be shown that slightly more of the technical inefficiency is due to units incurring too much cost in providing vaccines rather than operating at the wrong size. However, both sources of this technical inefficiency would have to be addressed for these units to become less wasteful of scarce resources.

Table 20: Descriptive statistics of the DEA results (models DEA1-3)

Measure	Mean	Std. deviation	Min	Max
Technical efficiency				
DEA1	0.35	0.26	0.04	1.00
DEA2	0.41	0.34	0.02	1.00
DEA3	0.32	0.30	0.01	1.00
'Pure' Technical efficiency				
DEA1	0.52	0.28	0.05	1.00
DEA2	0.73	0.30	0.10	1.00
DEA3	0.69	0.31	0.10	1.00
Scale efficiency				
DEA1	0.66	0.25	0.11	1.00
DEA2	0.57	0.35	0.02	1.00
DEA3	0.49	0.35	0.01	1.00

[61] Appendices 11-13 show the unit-specific results.

A comparison between model DEA2, which has five outputs, and model DEA3, in which all the outputs are summed together (both models have the same three inputs), sheds light on the impact of aggregating outputs. It was found that the technical efficiency score drops from 0.41 to 0.32. A comparison between models DEA1 and DEA2, which have one and three inputs respectively, but the same number of outputs (n=5), sheds light on the impact of aggregating inputs. It was found that the technical efficiency score increased from 0.35 to 0.41. These findings are consistent with the dimensionality issue raised in Chapter 3, whereby increasing the number of dimensions used in the characterisation of production reduces the discriminatory power of the analysis, i.e. it increases measured efficiency and the number of units identified as fully efficient (see below).

Models DEA1 – DEA3 suggest that the majority of units in this sample exhibited VRS. Table 21 shows that 87, 80 and 94 vaccination delivery units under specifications DEA1, DEA2 and DEA3 respectively, exhibited increasing returns to scale (IRS) (implying that they are too small). 17, 9 and 7 of the units under specifications DEA1, DEA2 and DEA3 respectively exhibited decreasing returns to scale (DRS) (implying that they are too large). And only 6, 21 and 9 of the units under specifications DEA1, DEA2 and DEA3 respectively were the 'right' size, i.e. they were operating at CRS.

Table 21: Returns to scale in the vaccination delivery units (models DEA1-3)

Types of returns to scale	Number of delivery units		
	DEA1	DEA2	DEA3
Increasing returns to scale	87 (79%)	80 (73%)	94 (85%)
Constant returns to scale	6 (5%)	21 (19%)	9 (8%)
Decreasing returns to scale	17 (16%)	9 (8%)	7 (6%)

7.5.1.1 'Efficient' units

Table 22 presents the efficient vaccination delivery units by specification and the number of times that each of these vaccination delivery units act as peers.

Table 22: Efficient vaccination delivery units and the number of times they are a peer

	Specification			
Efficient delivery unit	DEA1	DEA2	DEA3	Summation
1	0	2	0	2
4	0	4	6	10
6	30	5	0	35
9	75	22	22	119
15	0	2	0	2
16	0	9	0	9
17	0	1	0	1
18	63	18	17	98
19	0	8	11	19
21	0	5	6	11
24	0	3	0	3
29	0	4	0	4
35	19	0	0	19
40	0	1	0	1
42	0	3	0	3
70	36	0	0	36
88	0	77	0	77
91	0	26	14	40
95	0	33	21	54
96	0	3	0	3
105	0	0[62]	0	0
Number of efficient delivery units	5	19	7	

[62] Delivery unit 105 is efficient but does not act as a peer for any other delivery units.

Among the 110 vaccination delivery units, 21 of them were efficient (although one of these, delivery unit 105 did not act as a peer for any other delivery units) and acted as peers between 1 and 119 times across all three specifications. Only delivery units 9 and 18 were efficient across all three models. Accordingly, these two units acted as peers the greatest number of times; 119 and 98 times each respectively. Delivery unit 9 is a fixed-NGO delivery unit, whereas delivery unit is a fixed-GoB unit. The former is located in zone 1, while the latter is located in zone 3. Therefore, these peers did not have similar characteristics (see Table 23). However, an additional six delivery units were efficient across two models; one of these is a fixed-GoB unit (4), while the remaining five are all outreach-NGO units (6, 19, 21, 91 and 95). Specification DEA2 had the highest number of efficient delivery units (n=19), which is consistent with the dimensionality issue raised above and earlier in Chapter 3.

A look at the other end of the efficiency scores shows that unit 108 is ranked third most inefficient and most inefficient in models DEA1 (0.045), DEA2 (0.021) and DEA3 (0.013) respectively (efficiency scores in brackets). Similarly, unit 57 is ranked the most inefficient, seventh most inefficient and fourth most inefficient in model DEA1 (0.036), DEA2 (0.054) and DEA (0.035) respectively. It is also interesting to note that unit 96 was ranked second most inefficient under model DEA1 (0.036), however it was deemed efficient under model DEA2 and had a rank of 13th most efficient (0.864) under model DEA3.

Units 18, 9, 96, 108 and 57 have a cost per dose of \$0.20, \$0.22, \$0.82, \$6.02 and \$7.88 respectively, which ranks them as, from lowest to highest cost, 1st, 2nd, 48th, 109th and

110th. Therefore, there appears to be a high level of congruence between the unit costs and the DEA efficiency results.

Examination of the outputs for these five delivery units illustrates that the efficient units delivered far greater numbers of vaccines (Table 23). And it is also apparent that the inefficient units use much larger areas for providing vaccination services.

Table 23: A comparison of the inputs and outputs for delivery units 9, 18, 57, 96 and 108

Variable	9	18	96	108	57
Ownership	NGO	GoB	NGO	NGO	NGO
Type	Fixed	Fixed	Outreach	Outreach	Fixed
Location	Zone 1	Zone 3	Zone 9	Zone 9	Zone 7
Inputs					
labour	1	1	6	1	10
facility size dedicated to EPI	270	600	150	1,600	1,000
total hours	270	200	30	337	405
total cost	970	1,561	9,785	582	4,347
Outputs					
BCG	240	1,680	156	1	36
DPT	1,164	1,800	420	24	180
OPV	1,164	2,400	420	48	240
Measles	300	960	36	1	24
TT	1,548	960	180	24	72
total number of vaccines	4,416	7,800	1,212	98	552

7.5.2 SFA models[63]

The use of unit weights meant that vaccines were given the following order of importance (from most to least): OPV, DPT, TT, BCG and measles. The use of DALYs

[63] Appendix 14 shows the unit-specific results.

as weights meant that vaccines were given the following order of importance: DPT, measles, TT, BCG and OPV. And finally, the use of vaccine prices as weights infers the following order of importance: measles, OPV, DPT, BCG and TT.

Table 24 presents the descriptive statistics of the SFA models. The mean efficiency of models SFA1, SFA3 and SFA5 was 0.487, 0.524 and 0.476 respectively, for the Cobb-Douglas production frontier assuming a half-normal distribution. For the translog production frontier assuming a truncated normal distribution, the mean efficiency of models SFA2, SFA4 and SFA6 was 0.275, 0.247 and 0.289 respectively.

Table 24: Descriptive statistics of the SFA results (models SFA1-6)

	Mean	Std. Deviation	Min	Max
SFA1	0.487	0.168	0.12	0.77
SFA2	0.275	0.246	0.01	0.99
SFA3	0.524	0.141	0.18	0.76
SFA4	0.247	0.233	0.01	1.00
SFA5	0.476	0.174	0.09	0.77
SFA6	0.289	0.258	0.01	1.00

It is interesting to note that the mean efficiency scores of the Cobb-Douglas production frontier models are higher than those observed for the translog models, although the maximum efficiency observed never reaches one. Across the six models, unit 108 was consistently the least efficient and unit 18 the most efficient. Therefore, there is a high level of congruence between the unit costs, DEA efficiency results and the SFA efficiency findings.

Table 25 shows the estimated coefficients of the stochastic Cobb-Douglas production frontier assuming a half-normal distribution (model SFA1) and the statistics for noise (γ) and the inefficiency component (LR test of the one-side error). A value of γ of zero indicates that the deviations from the frontier are due entirely to noise, i.e. the model is equivalent to the traditional average response function, without the technical inefficiency effect u_i. On the other hand, a value of one indicates that all deviations are due to technical inefficiency. The null hypothesis that there are no technical inefficiency effects in the model can be conducted by testing the null and alternative hypotheses, H_0: $\gamma = 0$ versus H_1: $\gamma > 0$.

Table 25: Estimated results of the stochastic Cobb-Douglas production frontier model (SFA1)

Parameters	β	Standard error	t-ratio
Intercept (β_0)	5.716	0.833	6.860
ln (labour) (β_1)	0.390	0.138	2.826
ln (size of EPI) (β_2)	0.150	0.082	1.834
ln (total hours) (β_3)	0.193	0.128	1.509
σ_s^2	1.844	0.897	2.054
γ	0.693	0.372	1.861
Log likelihood	-156.929		
LR test of the one-sided error	0.608		

The results suggest that the null hypothesis can be rejected, that is the γ-estimate is greater than zero, which therefore implies that there are technical efficiency effects in the model and that the model is not equivalent to the traditional average response function.

Table 26 shows the estimated coefficients of the stochastic translog production frontier assuming a truncated normal distribution (model SFA2). In this model, the γ-estimate is equal to one, indicating that all deviations from the frontier are due entirely to technical inefficiency.

Table 26: Estimated results of the stochastic translog production frontier (SFA2)

Parameters	β	Standard error	t-ratio
Intercept (β_0)	6.037	3.479	1.735
ln (labour) (β_1)	-1.793	1.461	-1.227
ln (size of EPI) (β_2)	-0.320	1.696	-0.188
ln (total hours) (β_3)	1.364	1.088	1.253
(ln labour)^2 (β_4)	0.130	0.154	0.848
(ln size of EPI)^2 (β_5)	-0.107	0.099	-1.075
(ln total hours)^2 (β_6)	-0.360	0.160	-2.245
ln (labour) * ln (size of EPI) (β_7)	0.076	0.101	0.750
ln (labour) * ln (total hours) (β_8)	0.237	0.255	0.928
ln (size of EPI) * ln (total hours) (β_9)	0.361	0.129	0.279
σ_s^2	1.427	0.679	2.102
γ	0.999	0.000	160,797.360
μ	1.722	0.458	3.757
Log likelihood	-148.708		
LR test of the one-sided error	4.547		

7.5.3 Comparison of the DEA and SFA efficiency scores and ranks

Table 27 shows high correlation between the efficiency scores and ranks from the three DEA and six SFA specifications[64]. Correlation between the scores and ranks of three DEA specifications is upwards of 0.569 and 0.604 respectively.

[64] All correlations are significant at the 0.01 level (2-tailed)

Table 27: Correlations of unit cost, DEA and SFA results

	Unit costs	DEA1	DEA2	DEA3	SFA1	SFA2	SFA3	SFA4	SFA5	SFA6
Scores										
Unit costs	1.000									
DEA1	-0.655	1.000								
DEA2	-0.517	0.636	1.000							
DEA3	-0.473	0.569	0.952	1.000						
SFA1	-0.585	0.616	0.830	0.802	1.000					
SFA2	-0.475	0.614	0.834	0.855	0.810	1.000				
SFA3	-0.572	0.587	0.802	0.787	0.992	0.803	1.000			
SFA4	-0.461	0.627	0.798	0.825	0.790	0.975	0.786	1.000		
SFA5	-0.567	0.575	0.808	0.796	0.990	0.798	0.992	0.778	1.000	
SFA6	-0.456	0.578	0.784	0.827	0.804	0.943	0.799	0.951	0.814	1.000
Ranks										
Unit costs	1.000									
DEA1	-0.951	1.000								
DEA2	-0.705	0.650	1.000							
DEA3	-0.680	0.604	0.964	1.000						
SFA1	-0.721	0.658	0.905	0.923	1.000					
SFA2	-0.688	0.622	0.873	0.907	0.919	1.000				
SFA3	-0.708	0.643	0.889	0.921	0.993	0.921	1.000			
SFA4	-0.685	0.630	0.853	0.898	0.921	0.986	0.931	1.000		
SFA5	-0.690	0.626	0.883	0.913	0.988	0.908	0.992	0.917	1.000	
SFA6	-0.666	0.625	0.853	0.899	0.924	0.972	0.932	0.984	0.931	1.000

There is a high degree of correlation between the efficiency scores of the six SFAs, although a clear distinction between the Cobb-Douglas production frontier (models SFA

1, 3 and 5 have a correlation of upwards of 0.990) and the translog production frontier (models SFA 2, 4 and 6 have a correlation of upwards 0.943) specifications can be seen. Taken together, the scores of SFA models 1-6 have a correlation of upwards 0.778. Similarly, there is a high degree of correlation between the ranks of the six SFA of upwards of 0.908. Correlation between unit costs and DEA scores (upwards of -0.473), DEA ranks (upwards of -0.680), SFA scores (upwards of -0.456) and SFA ranks (upwards of -0.666) illustrates that unit costs increase as efficiency decreases.

It is interesting to note that the correlation between DEA1 and the other two DEA models (upwards of 0.569) and the six SFA models (upwards of 0.575) is lower than between models DEA2 and DEA3 (0.952), and between models DEA2 and DEA3 and the six SFA models (upwards of 0.784). This suggests that the manner in which the EPI services has been costed merits closer examination and / or the appropriateness of labour, facility size and total hours as proxies for total cost.

7.5.4 Stability of efficiency assessment between the specifications

While correlations describe overall relationships, they are not a satisfactory way to examine the changes in efficiency scores across different methods and specifications, as they do not show what happens to individual vaccination sites' scores (Jacobs 2001). Therefore, it is worth considering the effect of alternative specifications on the efficiency estimates for individual delivery units (Street 2003).

It was found that of 110 vaccination delivery units, there were two whose efficiency did not vary between DEA1, DEA2 and DEA3 (they were efficient under all three specifications). Among the remaining 108 vaccination delivery units, the difference between the maximum and the minimum score which a vaccination delivery unit obtained ranged from 0.036 to 1.000 (a difference in rank of 107 places). The difference between the maximum and minimum scores and ranks across the Cobb-Douglas production models (SFA 1, 3 and 5) was 0.103 (0.322 vs. 0.425) and 15 places respectively. The difference between the maximum and minimum score and rank across the translog production function models (SFA 2, 4 and 6) was 0.334 (0.395 vs. 0.730) and 22 places respectively.

7.5.5 Analysis of environmental variables

7.5.4.1 DEA models

Tables 28 and 29 show the mean technical, 'pure' technical and scale efficiency scores of the vaccination delivery units by location, type and ownership by specification.

Table 28: Technical, 'pure' technical and scale efficiency of vaccination delivery units by location (models DEA1-3)

Zone	Technical efficiency			'Pure' technical efficiency			Scale efficiency		
	DEA1	DEA2	DEA3	DEA1	DEA2	DEA3	DEA1	DEA2	DEA3
1	0.50	0.70	0.53	0.58	0.81	0.70	0.81	0.84	0.75
2	0.38	0.58	0.44	0.47	0.76	0.74	0.83	0.72	0.57
3	0.63	1.00	0.95	0.79	1.00	0.97	0.80	1.00	0.97
4	0.46	0.50	0.33	0.67	0.64	0.55	0.69	0.72	0.54
5	0.30	0.45	0.36	0.40	0.75	0.66	0.79	0.67	0.62
6	0.11	0.18	0.14	0.25	0.91	0.91	0.46	0.24	0.21
7	0.30	0.20	0.15	0.54	0.53	0.52	0.57	0.42	0.34
8	0.24	0.22	0.15	0.44	0.66	0.61	0.58	0.45	0.43
9	0.34	0.45	0.36	0.53	0.92	0.92	0.62	0.50	0.40
10	0.31	0.34	0.28	0.50	0.48	0.45	0.58	0.57	0.45
Total	0.35	0.42	0.32	0.52	0.73	0.69	0.66	0.58	0.49

Table 29: Efficiency of the DEA models by type and ownership of the vaccination delivery units

Type of unit	Technical efficiency			'Pure' technical efficiency			Scale efficiency		
	DEA1	DEA2	DEA3	DEA1	DEA2	DEA3	DEA1	DEA2	DEA3
GoB (fixed)	0.35	0.64	0.54	0.39	0.71	0.68	0.89	0.87	0.77
GoB (outreach)	0.35	0.39	0.27	0.47	0.58	0.57	0.70	0.62	0.44
GoB	0.35	0.58	0.48	0.41	0.68	0.65	0.84	0.81	0.70
NGO (fixed)	0.36	0.55	0.42	0.46	0.70	0.57	0.77	0.78	0.76
NGO (outreach)	0.34	0.34	0.25	0.56	0.76	0.73	0.57	0.46	0.37
NGO	0.35	0.37	0.28	0.55	0.75	0.70	0.61	0.52	0.44
Fixed	0.35	0.60	0.49	0.42	0.71	0.62	0.83	0.83	0.77
Outreach	0.34	0.34	0.25	0.56	0.75	0.72	0.58	0.47	0.37
All	0.36	0.43	0.32	0.52	0.73	0.69	0.66	0.58	0.49

The ANOVA and Kruskal-Wallis test, illustrate that the technical, 'pure' technical and scale efficiencies of the delivery units varied systematically by location (Table 30). With respect to scale efficiency the tests illustrated that GoB vaccination delivery units were, on average, relatively more efficient than NGO units, and that fixed vaccination delivery units were, on average, relatively more efficient than outreach units. Taken in combination, the results indicate that GoB fixed vaccination delivery units were, on average, relatively the most scale efficient type, whilst, NGO outreach units were the least efficient type. With respect to technical efficiency the tests indicate that NGO vaccination delivery units were, on average, relatively more efficient than GoB units, and that outreach vaccination delivery units were, on average, relatively more efficient than fixed units. Taken together, the results indicate that NGO outreach vaccination delivery units were, on average, relatively the most technically efficient type, whilst, GoB fixed units were the least efficient type. Finally, with respect to technical efficiency, differences in ownership and type of vaccination delivery units made no difference.

Table 30: Significance of selected environmental variables and technical, 'pure' technical and scale efficiency scores from model DEA1

	Technical efficiency		'Pure' technical efficiency		Scale efficiency	
Environmental variable	F-test	Kruskal-Wallis	F-test	Kruskal-Wallis	F-test	Kruskal-Wallis
Zone (1 – 10)	2.827	25.182	2.597	20.759	2.570	21.286
	(0.005)	(0.003)	(0.010)	(0.014)	(0.011)	(0.011)
Ownership	0.001	0.099	4.433	5.500	17.627	14.813
(GoB or NGO)	(0.974)	(0.754)	(0.038)	(0.019)	(0.000)	(0.000)
Type	0.032	0.063	5.644	7.914	27.816	21.654
(fixed or outreach)	(0.858)	(0.803)	(0.019)	(0.005)	(0.000)	(0.000)
Type and ownership	0.015	0.135	2.215	9.098	10.554	24.308
	(0.998)	(0.987)	(0.091)	(0.028)	(0.000)	(0.000)

Using specification DEA1, Table 31 shows that the length of time a programme site has been in operation is positively correlated with scale efficiency (significant at the 0.01 level).

Table 31: Correlation coefficients for selected environmental variables and technical, 'pure' technical and scale efficiency for specification DEA1

	Technical efficiency			'Pure' technical efficiency			Scale efficiency		
Environmental variable	Pearson	Kendall's tau	Spearman's rho	Pearson	Kendall's tau	Spearman's rho	Pearson	Kendall's tau	Spearman's rho
Years delivery site	0.051	0.128	0.190	-0.090	0.059	0.072	0.348	0.168	0.253
has been open	(0.599)	(0.047)	(0.048)	(0.351)	(0.397)	(0.456)	(0.000)	(0.016)	(0.008)
Population density	0.139	-0.032	-0.044	0.045	-0.030	-0.045	0.034	0.002	0.006
	(0.146)	(0.629)	(0.647)	(0.642)	(0.645)	(0.638)	(0.727)	(0.973)	(0.947)
Male / female ratio	0.222	0.155	0.223	0.117	0.057	0.091	0.220	0.161	0.238
	(0.020)	(0.18)	(0.019)	(0.222)	(0.384)	(0.347)	(0.021)	(0.014)	(0.012)
Literacy	-0.039	0.019	0.030	-0.027	0.001	0.004	0.044	0.005	0.009
	(0.685)	(0.772)	(0.759)	(0.777)	(0.994)	(0.966)	(0.647)	(0.934)	(0.923)

7.5.4.2 SFA models

Table 32 shows the mean efficiency scores of the vaccination delivery units by location, type and ownership using models SFA1 – SFA6.

Table 32: Efficiency of the SFA models by location, ownership and type

Variable	SFA1	SFA2	SFA3	SFA4	SFA5	SFA6
Zone 1	0.61	0.50	0.62	0.48	0.59	0.47
Zone 2	0.61	0.42	0.63	0.33	0.61	0.46
Zone 3	0.72	0.76	0.72	0.75	0.72	0.80
Zone 4	0.46	0.27	0.49	0.24	0.44	0.31
Zone 5	0.60	0.34	0.61	0.28	0.59	0.37
Zone 6	0.40	0.10	0.46	0.10	0.39	0.12
Zone 7	0.41	0.17	0.47	0.16	0.39	0.18
Zone 8	0.42	0.15	0.47	0.13	0.41	0.17
Zone 9	0.47	0.24	0.50	0.20	0.45	0.23
Zone 10	0.43	0.26	0.50	0.25	0.42	0.25
GoB (fixed)	0.64	0.54	0.65	0.49	0.64	0.59
GoB (outreach)	0.52	0.35	0.56	0.28	0.53	0.36
GoB (all)	0.61	0.49	0.63	0.44	0.62	0.54
NGO (fixed)	0.55	0.33	0.57	0.30	0.54	0.37
NGO (outreach)	0.44	0.20	0.48	0.18	0.42	0.20
NGO (all)	0.46	0.22	0.50	0.20	0.44	0.23
Fixed (all)	0.60	0.44	0.62	0.40	0.59	0.48
Outreach (all)	0.44	0.21	0.49	0.18	0.43	0.21
All	0.49	0.27	0.52	0.25	0.48	0.29

The ANOVA and Kruskal-Wallis tests illustrate that the efficiency varied systematically by location, ownership and type (Table 33). Table 34 illustrates that efficiency is positively related to the number of years the unit has been open for (significant at the 0.01 level).

Table 33: Significance of selected environmental variables and efficiency scores from models SFA1 and SFA2

	SFA1		SFA2	
Environmental variable	F-test	Kruskal-Wallis	F-test	Kruskal-Wallis
Zone (1 – 10)	4.293	33.734	6.306	30.580
	(0.000)	(0.000)	(0.000)	(0.000)
Ownership	18.148	17.482	25.097	16.557
(GoB or NGO)	(0.000)	(0.000)	(0.000)	(0.000)
Type	23.743	20.832	23.177	19.677
(fixed or outreach)	(0.000)	(0.000)	(0.000)	(0.000)
Type and ownership	9.579	25.144	11.149	23.779
	(0.000)	(0.000)	(0.000)	(0.000)

Table 34: Correlation coefficients for selected environmental variables and models SFA1 and SFA2

	SFA1			SFA2		
Environmental variable	Pearson	Kendall's tau	Spearman's rho	Pearson	Kendall's tau	Spearman's Rho
Years delivery site	0.308	0.132	0.193	0.361	0.176	0.259
has been open	(0.01)	(0.058)	(0.044)	(0.000)	(0.011)	(0.007)
Population density	-0.028	-0.018	-0.018	-0.073	-0.051	-0.071
	(0.771)	(0.780)	(0.854)	(0.450)	(0.435)	(0.463)
Male / female ratio	0.287	0.211	0.322	0.221	0.206	0.309
	(0.000)	(0.001)	(0.001)	(0.020)	(0.002)	(0.001)
Literacy	0.038	-0.003	0.000	0.027	0.018	0.027
	(0.693)	(0.961)	(0.996)	(0.782)	(0.782)	(0.776)

Although there is no diagnostic tool with which to choose the best model specification, some general rules of thumb can be applied (Jacobs 2001). The most important criterion for selecting one specification over another is whether the model is consistent with theory and in some way theoretically justifiable. Another useful criterion is the number of efficient units. *Ceteris paribus*, the fewer the better, although there should be enough peers available to make useful comparisons. The distribution of efficiency scores makes another useful criterion. The wider the better, *ceteris paribus*.

Figure 15 shows the frequency distribution of efficiency scores for the three DEA specifications and highlights that specification DEA2 produces the higher efficiency scores, while specification DEA1 produces a spread of efficiency scores that are more average.

Figure 15: Distribution of efficiency scores for the three DEA specifications

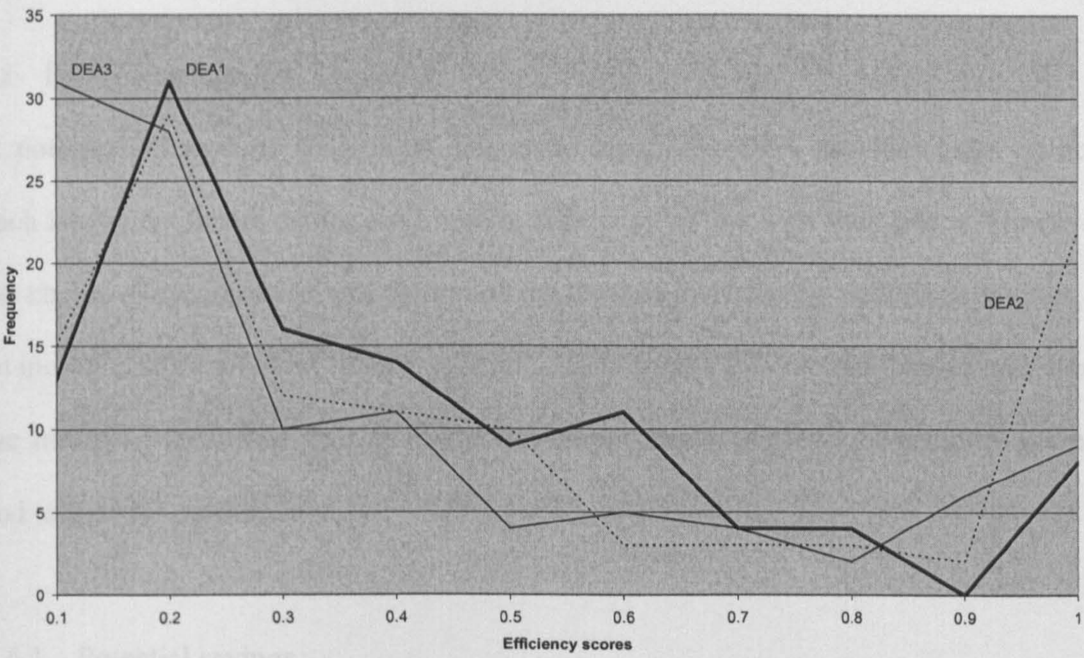
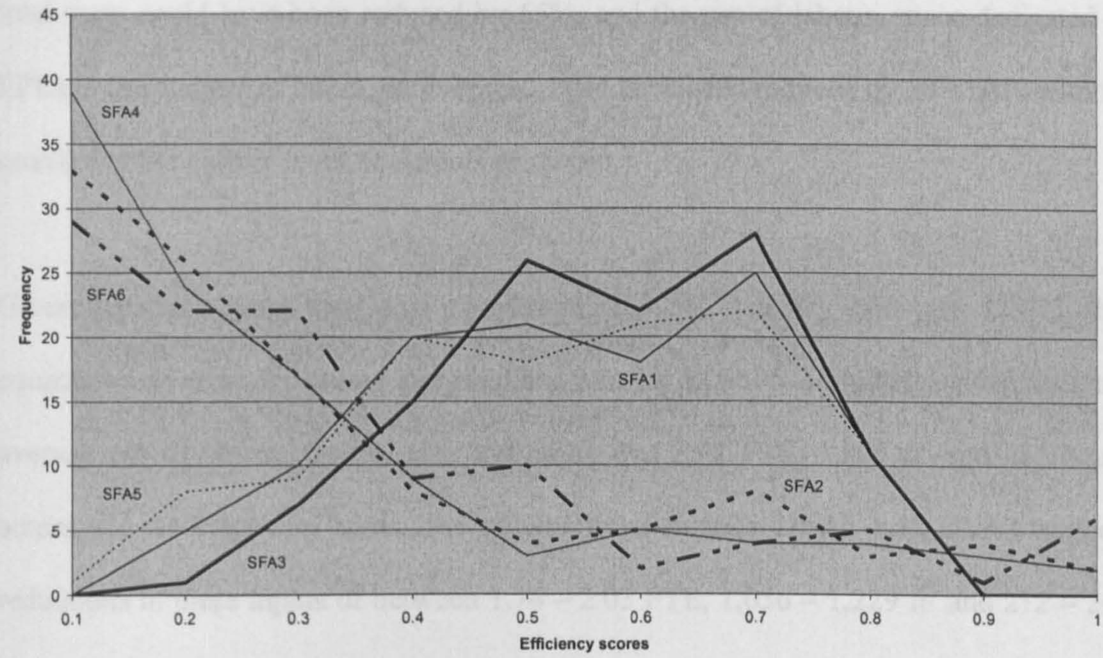


Figure 16 shows the frequency distribution of efficiency scores for the six SFA specifications and highlights that specification SFA6 produces the higher efficiency scores, while specification SFA3 produces a spread of efficiency scores that are more average.

Figure 16: Distribution of efficiency scores for the six SFA specifications



7.6 Policy implications

A potential strength of DEA is its diagnostic capability; DEA provides clues on how each inefficient health centre can improve efficiency in line with their peers. However, the choice of specification will determine the level of inefficiency, and hence savings, at an individual as well as an 'industry' level. Accordingly, this section focuses on linking the results of DEA with two related policy implications: the level of potential savings and targets for performance.

7.6.1 Potential savings

Model DEA1, and models DEA2 and DEA3, should be examined in turn, because the former uses total costs as the sole input, while specifications DEA2 and DEA3 use the amount of labour, amount of space used and total hours as inputs; the interpretation of the data is thus quite different. Average efficiency scores ranged from 35% for specification DEA1, to 32 – 42% for specifications DEA3 and DEA2 respectively. In other words, if delivery units were technically efficient and operated at the correct scale, total costs could have been reduced by 65%, and the use of labour, space dedicated to EPI and the number of hours, on average, could have been reduced by 58 – 68% without sacrificing the current level of outputs produced.

Given that the average total cost of operating the 110 delivery units was $2,075, this equates to an average saving per delivery unit of $1,359. Similarly, given that the average use of labour, facility size and hours was 2.98 FTE, 1,805 ft^2 and 362 hours across the 110 delivery units, the results from models DEA2 and DEA3 suggest reductions in these inputs of between 1.74 – 2.03 FTE, 1,056 – 1,229 ft^2 and 212 – 247 hours respectively.

7.6.2 Targets for efficiency improvement

DEA results can also be used as a managerial tool to improve efficiency of delivery units as it provides targets to achieve efficiency for each delivery unit. In light of the inevitable transaction costs of implementing efficiency improvement programmes, it is reasonable to target improvement to those delivery units which have most to gain, i.e. currently the least efficient. Table 35 provides an example of identifying target delivery units for possible improvement using specifications DEA1 and DEA2. In these specifications, there were respectively 105 and 91 inefficient delivery units of which 32 and 28 accounted for almost 40% of the technical inefficiency of the whole sample.

Table 35: Thirty-two and 28 most inefficient delivery units using specifications DEA1 and DEA2

DEA1			DEA2		
Delivery unit	Share of inefficiency %	Cumulative inefficiency %	Delivery unit	Share of inefficiency %	Cumulative inefficiency %
57	1.34%	1.34%	108	1.52%	1.52%
96	1.34%	2.68%	33	1.52%	3.04%
108	1.33%	4.01%	90	1.49%	4.53%
39	1.31%	5.32%	99	1.48%	6.01%
50	1.31%	6.63%	78	1.47%	7.48%
104	1.30%	7.93%	58	1.47%	8.95%
109	1.30%	9.23%	57	1.47%	10.42%
76	1.30%	10.53%	104	1.45%	11.87%
61	1.30%	11.83%	85	1.45%	13.32%
38	1.28%	13.11%	66	1.43%	14.75%
47	1.26%	14.37%	74	1.43%	16.18%
51	1.26%	15.63%	76	1.42%	17.60%
22	1.25%	16.88%	32	1.41%	19.01%
77	1.25%	18.13%	62	1.40%	20.41%
53	1.24%	19.37%	61	1.40%	21.81%
37	1.23%	20.60%	77	1.39%	23.20%
48	1.23%	21.83%	80	1.39%	24.59%
97	1.22%	23.05%	48	1.39%	25.98%
49	1.22%	24.27%	47	1.38%	27.36%
74	1.22%	25.49%	72	1.38%	28.74%
99	1.21%	26.70%	59	1.38%	30.12%
72	1.21%	27.91%	51	1.37%	31.49%
11	1.21%	29.12%	11	1.36%	32.85%
60	1.21%	30.33%	109	1.36%	34.21%
2	1.21%	31.54%	28	1.36%	35.57%
54	1.20%	32.74%	63	1.35%	36.92%
85	1.19%	33.93%	79	1.34%	38.26%
58	1.19%	35.12%	89	1.34%	39.60%
102	1.17%	36.29%			
52	1.17%	37.46%			
46	1.17%	38.63%			
33	1.16%	39.79%			

After identifying these target delivery units, information from the DEA results can be used to set unique target levels for each type of input that inefficient delivery units need to meet in order to become more efficient. Table 36 presents actual and target resource use for these delivery units. For example, it cost delivery unit 108 $582 to administer one dose of BCG dose, 24 doses of DPT, 48 doses of OPV, one dose of measles and 24 doses of TT (a total of 98 vaccinations). To become an efficient delivery unit, it should cost $242. Alternatively, to produce that level of output, the unit used one member of staff, 1,600 square feet of space and 337 of hours. In order to become an efficient delivery unit it would not need to reduce use of staff, however it would need to use 166 square feet and 35 hours.

Table 36: Actual and target resource use for the 32 and 28 most inefficient delivery units, using specifications DEA1 and DEA2

DEA1			DEA2						
	Actual	Target			Actual			Target	
Delivery unit	Total cost	Target Total costs	Delivery unit	Staff	Size	Hours	Staff	Size	Hours
57	4,347	338	108	1	1,600	337	1	166	35
96	9,785	497	33	5	900	900	1	37	35
108	582	242	90	1	2,700	506	1	186	35
39	6,313	493	99	1	9,801	100	1	3,279	33
50	1,651	284	78	2	450	200	1	79	35
104	911	260	58	2	324	324	1	35	35
109	2,532	341	57	10	1,000	405	1	45	37
76	3,599	403	104	3	200	200	1	35	35
61	2,135	324	85	1	300	300	1	35	35
38	15,077	1,255	66	2	843	506	1	66	41
47	1,204	301	74	3	160	160	1	35	35
51	1,030	286	76	2	490	270	1	63	35
22	2,588	423	32	3	2,613	70	1	1,022	34
77	1,177	311	62	8	2,700	2,700	1	240	240
53	1,094	300	61	2	980	504	1	80	44
37	12,022	1,383	77	1	375	375	1	35	35
48	833	270	80	1	288	288	1	39	36
97	1,010	302	48	1	320	320	1	37	35
49	1,026	316	47	1	300	300	1	35	35
74	483	257	72	1	3,267	2,400	1	52	38
99	1,382	354	59	2	252	252	1	35	35
72	325	255	51	1	225	225	1	35	35
11	1,668	399	11	2	13,068	225	1	61	41
60	2,217	450	109	2	1125	337	1	37	37
2	1,085	330	28	6	864	216	1	53	37
54	540	277	63	1	535	248	1	38	36
85	485	275	79	9	3,600	1,800	1	527	224
58	434	272	89						
102	2,008	460							
52	2,793	585							
46	2,793	585							
33	448	277							

7.7 Summary

- In this chapter, the efficiency of vaccination delivery units in Dhaka City, Bangladesh was examined. This was achieved through the use of DEA and SFA, in which best practice frontiers from 110 units were constructed;
- Given the Government's stated objective to mobilise additional resources via improvements in the efficiency of health facilities, input-orientated specifications under VRS were adopted;
- The mean technical efficiency of the vaccination delivery units was 0.35, 0.41 and 0.32 in specifications DEA1, DEA2 and DEA3 respectively. These findings are consistent with the dimensionality issue. The DEA specifications indicate that the majority of units in this sample exhibited IRS;
- The mean efficiency of the Cobb-Douglas production frontier models SFA1, SFA3 and SFA5 was respectively 0.49, 0.52 and 0.48. For the translog production frontier, the mean efficiency of models SFA2, SFA4 and SFA6 was respectively 0.28, 0.25 and 0.29;
- There appears to be a high level of congruence between the unit costs, the DEA and SFA efficiency scores. However, the maximum difference in score and rank between the DEA and SFA models was large, particularly for the DEA models;
- Efficiency varied systematically by location, type and ownership. Length of time a unit had been in operation was positively correlated with scale efficiency;
- After identifying inefficient delivery units, unique target levels for each type of input can be identified.

The next Chapter presents the costs of providing health care among a sample of health centres in rural Bangladesh.

Chapter 8

VARIATION IN THE COST OF DELIVERING PRIMARY HEALTH CARE IN RURAL BANGLADESH

This chapter presents the costs of providing health care among a sample of health centres in rural Bangladesh. After a brief introduction, there are three sections to this chapter. The first focuses on describing the methods used, in particular, the sampling and data collection. The second part describes the results, focussing on the mean cost per health centre by district, and the weighted mean cost per visit. The chapter concludes with a summary of the chapter.

8.1 Introduction

Chapters 6 and 7 presented data on the cost and efficiency of vaccination services in urban Bangladesh. Chapters 8 and 9 focus on rural Bangladesh. Ideally the same services would have been analysed, but as stated in Chapter 5, vaccines are not routinely provided at union-level health centres (although they regularly act as an outreach site once per month). While the DFID-funded study from which the health centre data come from (see Chapter 1), did collect data from nine UHCs, which co-ordinate EPI activities in rural areas, this sample was clearly too small to conduct the subsequent parametric and non-parametric in the following chapter. Therefore, the data collected from 36 health centres are used.

As previously noted in Chapter 5, in 1998 Bangladesh began a sector-wide approach to extend health care to vulnerable populations, especially through a package of essential services emphasising maternal care, certain communicable diseases and child health.

The package was designed to improve population health status through a targeting approach which singled out facilities used more by the poor, effective services for diseases borne proportionately more by the poor, and rural areas where population health is the lowest. As such, the assumption was that improvements in health status could be supply-led (Ensor et al. 2002).

First-level government health centres in rural Bangladesh are usually staffed by a paramedic (medical assistant or sub-assistant community medical officer), who is usually male with at least four years of clinical training. In addition, there is a female reproductive health worker (family welfare visitor) who has had 18 months training in the MCH and FP services. In some facilities, there is a position for a doctor, but in most cases these positions remain vacant (Arifeen et al. 2005).

On average, the government funds one health centre for every three unions, at a cost of around $4,000 per health centre, which is mostly for the salaries of the health and support staff (Ensor et al. 2003a). In addition, each health centre receives a medical and surgical requisite allocation of around $1,250, most of which is for drugs. As Ensor et al. (2003a) note, "... the current allocation process ... bears[s] little relation to either the size of the population or the number of patients treated". Therefore, the overall aim of this chapter is to estimate the cost, from the provider perspective, of delivering the ESP, excluding reproductive health services[65], at union-level health centres. The specific objectives are to:

- estimate the total and unit costs of delivering health care services;
- describe the variation in these costs.

[65] FP services were excluded for the purpose of the DFID-funded project, which required the cost per visit for general health services.

8.2 Methods

8.2.1 Selection of sample

Data were collected from 36 health centres from three districts in Bangladesh (Brahmanbaria, Chandpur and Moulvi Bazar). These districts were selected to represent high, medium and low performing districts using a variety of indicators of disease, vaccination coverage, health service provision and access to health services (BBS 1997; UNICEF 1999). The following variables were included for each of the 64 districts in the country:

- To reflect major childhood diseases:
  - number of episodes of diarrhoea per 1,000 population;
  - number of deaths from diarrhoea per 1,000 population;
  - number of episodes of pneumonia per 1,000 population;
  - number of deaths from pneumonia per 1000 population;
  - number of cases of measles per 1,000 population;
  - number of deaths from measles per 1,000 population.
- To reflect access to health care:
  - number of health centres;
  - number of beds per 1,000 population;
  - percentage of women delivering with an untrained midwife;
  - percentage of children never vaccinated with DPT vaccine;
  - percentage of children vaccinated with two or more doses of DPT vaccine;
  - percentage of children never vaccinated with OPV vaccine;
  - percentage of children vaccinated with two or more doses of OPV vaccine;
  - percentage of children vaccinated with measles vaccine.
- To reflect socio-economic variables:

- percentage of boys aged 6-10 years old in primary school;
- percentage of girls aged 6-10 years old in primary school;

Each district was ranked from best to worst for each of the above variables. These rankings for all the indicators were summed together for each district, and a rank established. These figures were compared with the ranked position of each district using two other approaches; the poverty index and the Human Development Index (Khatun 2001). Because the development of such indices is often controversial, only those districts that were consistently placed using all three approaches were included. For example, for a district to be chosen as a top third district it had to be in the top third of the Human Development and poverty indices and it also had to appear the greatest number of times in the top third of the index specifically constructed. This happened for the districts shown in Table 37.

Table 37: Ranking of districts using three different indices

Top third for all indices	Mid third for all indices	Bottom third for all indices
Chandpur	Dinajpur	Brahmanbaria
Chittagong	Gazipur	Kurigram
Dhaka	Kishoreganj	Netrakona
Jhalakati	Madaripur	Nilpharmari
Khulna	Manikganj	Rangpur
Narayanganj	Masura	Sherpur
Comilla	Moulvi Bazaar	
Feni		

As any of these districts could have been selected using this process, availability of data was considered, particularly whether the local-level planning system had been introduced[66], and whether the vaccine against hepatitis B was planned for introduction (which at the time included: Faripur, Feni, Jessore, Joypurhat, Moulvi Bazar and Pirojpur districts)[67]. It was also considered whether ICDDR,B was known and working within the area to be a potential benefit and a reason for selection. Therefore the final selection was: Chandpur, Moulvi Bazar and Brahmanbaria (Figure 17).

Within each district, two or three Upazilas were selected at random, from which three to eight health centres were selected at random, such that 12 health centres from each district were selected. This process meant that the sample selected in Chandpur, Brahmanbaria and Moulvi Bazar districts represented 13.5% (12 / 89), 12.4% (12 / 97) and 17.9% (12 / 67) of all health centres. Therefore, across the three districts 14.2% of all health centres were selected (36 / 253) (see Table 38). However, two health centres were excluded for missing output data. Therefore the final sample was 34 health centres. Table 40 provides some background details on each of the chosen upazilas.

Table 38: Sample of health centres

District	Upazilas	Number of sites	Final sample
Brahmanbaria	Akhaura	4	4
	Kasba	4	4
	Sarail	4	3
Chandpur	Haziganj	4	3
	Shahrasti	8	8
Moulvi Bazar	Borolekha	5	5
	Kulaura	4	4
	Srimongal	3	3
Total		36	34

66 The local-level planning system was being introduced in five districts at the time of data collection (Dhaka, Mymensingh, Radshahai, Chandpur and Gopelganj).
67 Again, this was motivated by the DFID-funded project.

Figure 17: Map of Bangladesh with Brahmanbaria, Chandpur and Mouli Bazar Districts indicated

Table 39: Background statistics of the selected Upazilas

Indicator	Brahmanbaria			Chandpur		Moulvi Bazar		
	Kasba	Akhaura	Sarail	Hajigonj	Shaharasti	Borolekha	Srimongal	Kulaura
Area (sq. km)	207.76	72	239.52	190	168	458	4.15	678
Number of *Unions*	10	5	10	11	9	12	10	17
Number of *Wards*	30	15	30	36	27	36	30	51
Number of Villages	258	125	146	157	169	325	204	575
Number of households	55,295	23,616	49,785	51,855	37,958	33,006	52,275	66,771
Total population	319,309	144,510	327,533	331,511	242,092	253,435	268,358	417,683
Number of 0-11 months old	11,222	4,448	10,490	10,742	6,152	8,123	8,322	11,825
Number of under 5 years old	66,092	23,797	54,330	41,902	38,733	44,447	46,644	68,460
Number of adolescents (10-19 years)	73,441	33,235	75,332	76,248	55,681	58,291	61,640	96,067
Number of 15-49 years old women	48,355	29,039	70,280	74,988	53,494	42,466	53,600	93,822
Number of births registered (July 00 - June 01)	11,249	5,277	5,556	8,161	6,152	2,301	500	87
Number of deaths registered (July 00 - June 01)	315	1,445	1,650	724	602	546	20	7
Number of GoB health centres	5	4	9	11	11	7	8	13
Number of NGOs (H&FP)	2	0	2	NS	NS	3	3	4

NS: not stated

8.2.2 Cost analysis

The cost analysis was performed using standard costing guidelines, and adopted a provider perspective (Creese and Parker 1993). The health centres were costed by the 'ingredients' approach, in which the total quantities of goods and services actually employed in delivering the activities were estimated, and multiplied by their respective unit prices[68]. Cost information was obtained from various sources, including administrative records, interviews and direct observation. Price data were collected from surveys undertaken in the upazila sadars[69]. The prices of drugs were obtained from the Bangladesh National Formulary (MOHFW 2003). Cost and output indicators were collected for the financial year July 2001 – June 2002. All figures are presented in 2002 US dollars using the average official exchange rate between July 2001 – June 2002[70].

Resources, and hence costs, have been categorised according to whether they are capital (land, buildings, transport, equipment and furniture) or recurrent (drugs, supplies, personnel, logistics and miscellaneous) items. To estimate the cost of buildings used to deliver general health services, the space used for reproductive health services was excluded. Area in square feet has been costed on the basis of the construction cost stated in the Public Works Department in 2003. A notional 25 year working life for buildings was used. Furniture and equipment for reproductive health services were excluded. Estimates of the working life of different pieces of furniture and equipment were obtained from the study by Barkat et al. (1999). Finally, a 3% discount rate was used in conjunction with the resource-specific working life estimates in order to obtain estimates of the annual economic equivalent costs of capital.

---

[68] See Appendix 16 for a copy of the health centre facility survey form used to collect the data.
[69] The sadar is the district capital.
[70] 1US$ = 59.63 Bangladeshi Taka.

In order to allocate the time and thus cost of staff for general health services, excluding reproductive health services, the estimate from the work by Barkat et al (1999) was used, in which 76.30% of staff time was devoted to general health services.[71]

The disease profiles routinely compiled by the health centres from July 2001 to June 2002 were used. This enabled four measures of output to be used: visits of patients aged under one year, visits of patients aged from one to four years, visits of patients aged greater than four years and visits for all age groups. Unfortunately, these data were not available broken down by ESP line items, e.g. child health care, communicable disease control and limited curative care.

8.3 Results

8.3.1 Total cost, output and unit cost per health centre

Table 40 presents the total costs, unit costs and cost profile of the health centres. The average annual cost per health centre was $8,873 and the annual number of visits was 11,582, resulting in a cost per visit of $0.77.

The cost per visit ranges from $0.31 to $1.77 per visit (a 5.5-fold difference), among the sample of 34 health centres (Table 41). The number of visits at the health centres largely determines the unit cost of delivering care. The unit cost of treating patients declines as the number of visits increases (Figure 18). A negative correlation was found between the cost per visit and the number of visits, and this relationship was significant at the 1% level. This implies that marginal cost is lower than average cost.

[71] It should be noted that the Barkat et al. (1999) study is now rather dated and used a relatively small sample. With more resources, alternative approaches to allocate staff time would have been explored such as time-and-motion methods. Given the importance of staff costs, it is important to bear in mind the potential sensitivity of results presented here and in the proceeding chapter to this assumption.

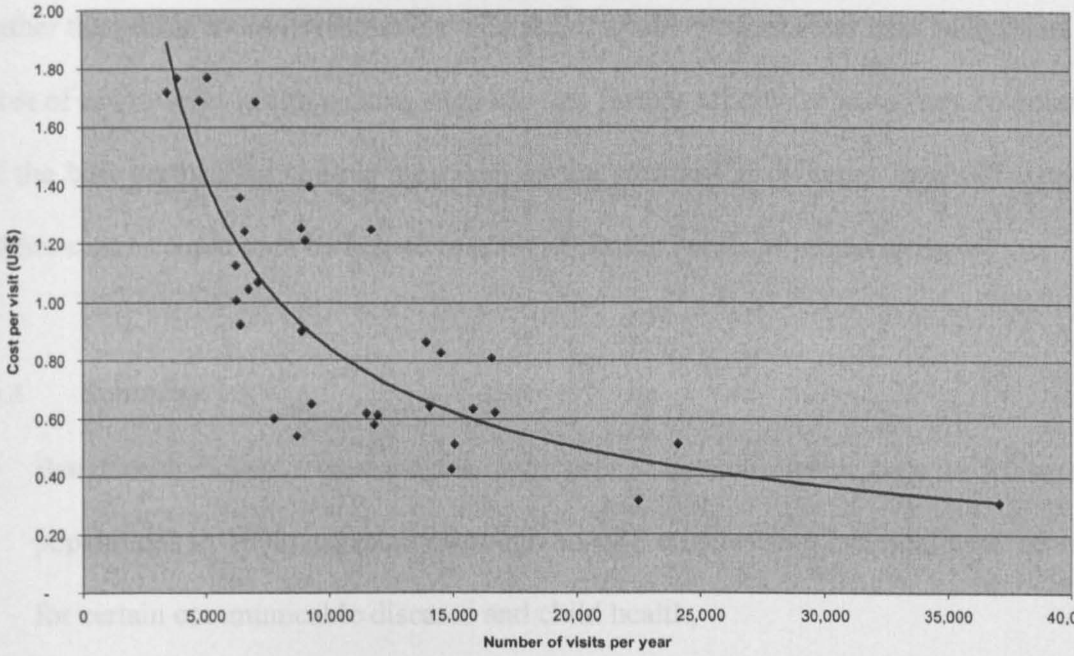
Table 40: Total cost, output and unit cost per visit, by District, in 2002 US$

Cost category	Brahmanbaria		Chandpur		Moulvi Bazar		Total	
	Cost	% of total cost	Cost	% of total cost	Cost	% of total cost	Cost	% of total cost
Capital								
furniture	3,445	2.94	3,422	3.40	4,215	4.54	11,082	3.57
equipment	4,795	4.09	4,333	4.31	3,820	4.12	12,949	4.17
transport	-	0.00	-	0.00	189	0.20	189	0.06
land & building	5,284	4.51	3,008	2.99	4,082	4.40	12,374	3.98
Sub-total	13,523	11.55	10,763	10.70	12,306	13.26	36,593	11.78
Recurrent								
salary	54,978	46.94	41,855	41.59	46,491	50.10	143,325	46.15
drugs	47,901	40.90	46,790	46.50	32,583	35.11	127,274	40.98
supply	57	0.05	195	0.19	355	0.38	606.58	0.20
miscellaneous	657	0.56	1,030	1.02	1,070	1.15	2,757	0.89
logistics	-	0.00	-	0.00	-	0.00	-	0.00
Sub-total	103,594	88.45	89,870	89.30	80,499	86.74	273,963	88.22
Total	117,117	100.00	100,633	100.00	92,806	100.00	310,556	100.00
Total number of visits	169,025		78,682		157,668		405,375	
Cost per visit	0.69		1.28		0.59		0.77	

Table 41: Total cost, output and unit cost per visit, by health centre, in 2002 US$

		total cost	Number of patients	unit cost
1	FWC 1	7,651	7,172	1.07
2	FWC 2	8,136	4,807	1.69
3	FWC 3	6,317	6,274	1.01
4	FWC 4	9,032	5,106	**1.77**
5	FWC 5	5,983	3,478	1.72
6	FWC 7	12,959	9,261	1.40
7	FWC 8	12,048	13,913	0.87
8	FWC 9	8,758	6,446	1.36
9	FWC 10	14,642	11,705	1.25
10	FWC 11	6,877	3,892	**1.77**
11	FWC 12	8,230	6,628	1.24
12	FWC 17	4,707	8,723	0.54
13	FWC 18	9,023	14,058	0.64
14	FWC 19	11,192	8,920	1.25
15	FWC 20	10,381	16,696	0.62
16	FWC 21	13,438	16,560	0.81
17	FWC 22	4,686	7,783	0.60
18	FWC 23	7,322	11,950	0.61
19	FWC 24	12,347	24,106	0.51
20	FWC 25	12,030	14,504	0.83
21	FWC 27	11,017	9,087	1.21
22	FWC 28	10,021	15,816	0.63
23	FWC 33	11,464	37,062	**0.31**
24	FWC 34	6,357	14,929	0.43
25	FWC 35	7,075	6,775	1.04
26	FWC 36	8,050	8,916	0.90
27	FWC 37	5,941	6,434	0.92
28	FWC 38	6,049	9,304	0.65
29	FWC 39	11,961	7,085	1.69
30	FWC 40	7,702	15,070	0.51
31	FWC 41	6,867	11,815	0.58
32	FWC 42	7,067	6,268	1.13
33	FWC 43	7,157	22,490	0.32
34	FWC 44	7,117	11,520	0.62
Total		310,556	405,375	0.77

Figure 18: Relationship between service volume and unit cost in the health centres



These results show a clear inverse relationship between the cost per visit and service volume among the 34 health centres. It seems that the optimum service volume, corresponding to the lowest average cost, has not yet been reached in the sample of centres. A significant factor appears to be the existence of substantial fixed costs associated with the delivery of general health services at these health centres. Although Table 41 above classifies staff as a recurrent item, in line with standard costing guidelines, in reality personnel costs are fixed in nature, at least in the short-term. Thus between 52% and 63% of the resources used to produce these services were fixed, which means that these resources change little, if at all, as the volume increases or decreases. Under these conditions, the results, as expected, show that up to a certain volume of service, a larger number of visits tends to reduce the average cost.

These results provide estimates on the supply-side cost of the ESP. Given that much of the variation in unit costs can be attributed to variation in the number of visits, these findings also suggest that differences in the factors affecting demand, such as distance living from the centre, information on services and perceived quality, are also at play, rather than differences in resource availability *per se*. Nevertheless, this analysis of the cost of union-level health centres suggests that further efficiency gains may be possible if the best performing centres are taken as the standard at different levels of outputs. Other centres could then be helped to achieve similar levels of productivity.

8.4 Summary

- Bangladesh began a sector-wide approach to extend health care to vulnerable populations in 1998, especially through an ESP emphasising maternal care, services for certain communicable diseases and child health;
- The current allocation of resources to union-level health centres bears little relation to either the size of the population or the number of patients treated. Therefore, the overall aim of this chapter is estimate the cost, from the provider perspective, of delivering the ESP and to describe variation in these costs;
- Data were collected from 36 health centres from three districts in Bangladesh (Brahmanbaria, Chandpur and Moulvi Bazar), although two health centres were excluded because of missing data. The aim in selecting the districts was to represent high, medium and low performing districts using a variety of indicators of disease, vaccination coverage, health service provision and access to health services;
- The health centres were costed by the 'ingredients' approach. Cost and output indicators were collected for the financial year July 2001 – June 2002;
- The average annual cost per health centre was $8,873 and the average annual number of visits was 11,582, resulting in a mean cost per visit of $0.77;
- The cost per visit ranged from $0.31 to $1.77 per visit (a 5.5-fold difference), among the sample of 34 health centres. The number of visits at the health centres largely

determines the unit cost of delivering care. A significant factor appears to be the existence of substantial fixed costs associated with the delivery of general health services at these health centres.

Ensor et al. (2003a) stated that the efficiency with which services in Bangladesh are delivered at the local level is "… an important subject for future investigation". The next chapter analyses these data using parametric and non-parametric efficiency measurement techniques in order to identify whether further efficiency gains might be possible.

Chapter 9

DATA ENVELOPMENT ANALYSIS AND STOCHASTIC FRONTIER ANALYSIS OF PRIMARY HEALTH CARE CENTRES IN RURAL BANGLADESH

This chapter presents the efficiency of health centres derived by SFA and DEA, using the data presented in Chapter 8. Following a brief introduction, there are five parts to this chapter. The first presents the data. The second describes the models used. In total, five models are included in this chapter; three for the DEAs and two for the SFAs. The third part provides an overview of the performance of health centres using 2001-02 data based on the five models, focussing on a summary of the efficiency scores, the number of efficient health centres and the number of times health centres are a peer. This section also includes a correlation analysis of the scores and ranks between the different specifications. The fourth section illustrates how DEA can be used to identify unit-specific and industry-level potential savings, and associated targets. The chapter concludes with a summary.

9.1 Introduction

As previously stated, the ESP consists of reproductive health care; child heath care; communicable disease control; and limited curative care. Unfortunately, the GoB faces significant resource constraints in funding the ESP. Previous reports have found that the potential for additional resource mobilisation is limited, and suggested that improvements in the internal efficiency of health care services must be a critical component of efforts to provide the ESP to the whole population (Rannan-Eliya and Somanathan 2003). The preceding chapter estimated a 5.5-fold difference in the cost per visit across a sample of health centres, which suggests there is scope to improve efficiency. Furthermore, because government-funded health facilities in Bangladesh are not profit-seeking entities, coupled with the fact that their input mix is largely determined by external rules and budgetary allocations, they cannot be assumed to operating efficiently.

Therefore, the overall aim of this chapter is to assess the efficiency of a sample of health centres in rural Bangladesh, in order to identify the scope to mobilise additional resources via improvements in their operating efficiency. The specific objectives are to:

- use DEA and SFA techniques to identify the level of technical and scale efficiency in the sample of health centres;
- compare and contrast these findings with those from the preceding chapter on the unit costs of the sample of health centres;
- investigate possible causes for differences in the efficiency scores;
- identify the potential savings and related targets for the sample of health centres.

9.2 Data

As stated previously in Chapter 8, the final data set consisted of 34 out of a possible 36 health centres in three Districts in Bangladesh (Brahmanbaria, Chandpur and Moulvi Bazar. Hence, 94.4% of all health centres in the sample were included. All data are for the period July 2001 – June 2002.

Data from Chapter 8 illustrated that on average 88.4% (range: 77.6 – 97.2%) of the total annual operating cost of the sample of health centres is accounted for by expenditure on staff and drugs. Therefore, on the input side, in addition to total cost, expenditure on staff and drugs were also included as model inputs. The number of patients treated aged below one year, between one and four years of age, and greater than four years of age were included as model outputs. In addition, the outputs were summed together to produce the total number of patient visits for all age groups. The descriptive statistics are given in Table 42.

Table 42: Descriptive statistics of inputs and inputs

Variable	Mean	Std. deviation	Min	Max
Inputs				
total cost ($)	8,812	2,656	4,686	14,642
staff expenditure ($)	4,138	1,340	1,376	6,441
drugs expenditure ($)	2,889	1547	269	7,775
Outputs				
total number of patients	11,310	6,737	3,478	37,062
total number of patients aged < 1 year	692	676	83	3,121
total number of patients aged 1-4 years	1,699	982	580	4,464
total number of patients aged 4+ years	8,920	5,738	2,742	32,010

A summary of the specifications, described in more detail below, is included in Table 43. The total number of input and output variables ranged from three to five. A common rule of thumb used with DEA suggests that the number of observations in the data set should be at least three times the sum of the number of input and output variables, i.e. 3 x (2 + 3) = 15 (Cooper et al. 2003). An alternative rule of thumb suggested by Dyson et al. (2001) states that the number of observations should be at least twice the product of the number of inputs and outputs, i.e. 2 x 2 x 3 = 12. Therefore, according to either of these rules of thumb, and the specifications chosen, the final sample size of 34 is acceptable for the proposed DEAs. However, the sample appears to be at the limit of acceptability for SFA, which will be explored in more detail in the results section.

Table 43: DEA and SFA specifications

Specification	DEA1	DEA2	DEA3	SFA1[72]
Inputs				
total cost ($)	✓	✓	✓	✓
staff expenditure ($)		✓	✓	✓
drugs expenditure ($)		✓	✓	✓
Outputs				
total number of patients		✓		✓
total number of patients aged < 1 year	✓		✓	
total number of patients aged 1-4 years	✓			
total number of patients aged 4+ years	✓		✓	
Number of variables	4	3	5	3

9.3 Methods

9.3.1 DEA models[73]

Given the Government's stated objective to mobilise additional resources via improvements in the operating efficiency of health facilities, an input-orientated specification has been adopted for each model. This considers what reduction in inputs is possible given existing levels of outputs. This specification is also consistent with the fact that treating fewer patients is clearly better, in the sense that it may reflect successful health promotion and prevention programmes. However, this needs to be balanced against the fact that utilisation rates of government health facilities are low, suggesting that there could be considerable unmet need. The assumption of CRS is only appropriate when all health centres are operating at an optimal scale. Yet, in reality

[72] Two models were run using the same inputs and outputs.
[73] The data used for models DEA1-3 are presented in Appendices 16-18.

there are reasons to suspect that health centres may not be operating at optimal scale, therefore a VRS specification is additionally adopted.

For the DEA models, the technology was initially constructed under CRS and strong disposability of costs (as costs increase, outputs must increase, *ceteris paribus*) TE_{CRS}. Allowances can be made in the restraints to allow for VRS TE_{VRS}. Further, the type of scale inefficiencies were determined by employing a third model TE_{NIRS}. In all these cases the definitions given by Färe et al. (1994) were followed. The DEAP programme by Coelli (1996a) has been used for the computations. The linear programming problems presented in Chapter 7 apply here too.

9.3.2 SFA models

Frontier Version 4.1 (Coelli 1996b) was used to estimate a Cobb-Douglas production frontier assuming a half-normal distribution (SFA1) and a translog production frontier assuming a truncated normal distribution (SFA2).

9.3.3 Analysis of environmental variables

The ANOVA and Kruskal-Wallis tests were conducted in order to test the null hypotheses that the mean technical, 'pure' technical and scale efficiencies of the delivery units are the same across the three districts against the alternative hypotheses that they differ from one another.

9.4 Results

9.4.1 DEA models[74]

Table 44 presents the efficiency results and shows that technical efficiency (TE CRS) was 0.53 in specification DEA1, 0.48 in specification DEA2 and 0.58 in specification DEA3. In other words, if health centres were technically efficient and operated at the correct scale, expenditure on staff and drugs could be reduced by 47%, 52% and 42% respectively without sacrificing the current level of outputs produced.

Table 44: Descriptive statistics of the DEA results (models DEA1-3)

Specification	Measure	Mean	SD	Min	Max
DEA1	Technical efficiency	0.53	0.25	0.18	1.00
	'Pure' technical efficiency	0.69	0.21	0.34	1.00
	Scale efficiency	0.75	0.20	0.32	1.00
DEA2	Technical efficiency	0.48	0.26	0.18	1.00
	'Pure' technical efficiency	0.68	0.12	0.31	1.00
	Scale efficiency	0.69	0.24	0.27	1.00
DEA3	Technical efficiency	0.58	0.28	0.21	1.00
	'Pure' technical efficiency	0.73	0.20	0.32	1.00
	Scale efficiency	0.77	0.22	0.30	1.00

By decomposing this CRS technical efficiency measure into 'pure' technical efficiency (TE VRS) and scale efficiency, it can be shown that slightly more of the technical inefficiency is due to health centres using too many inputs in treating the different age groups of patients rather than operating at the wrong size. However, both sources of this technical inefficiency must be addressed for these facilities to become less wasteful of scarce resources.

[74] The centre-specific results for models DEA1-3 are presented in Appendices 19-21.

Figure 19 illustrates the frequency distribution of efficiency scores of the three DEA specifications. It highlights that specification DEA3 produces slightly higher efficiency scores, while specifications DEA1 and DEA2 produce a spread of efficiency scores that are more average.

Figure 19: Distribution of efficiency scores for the three DEA specifications

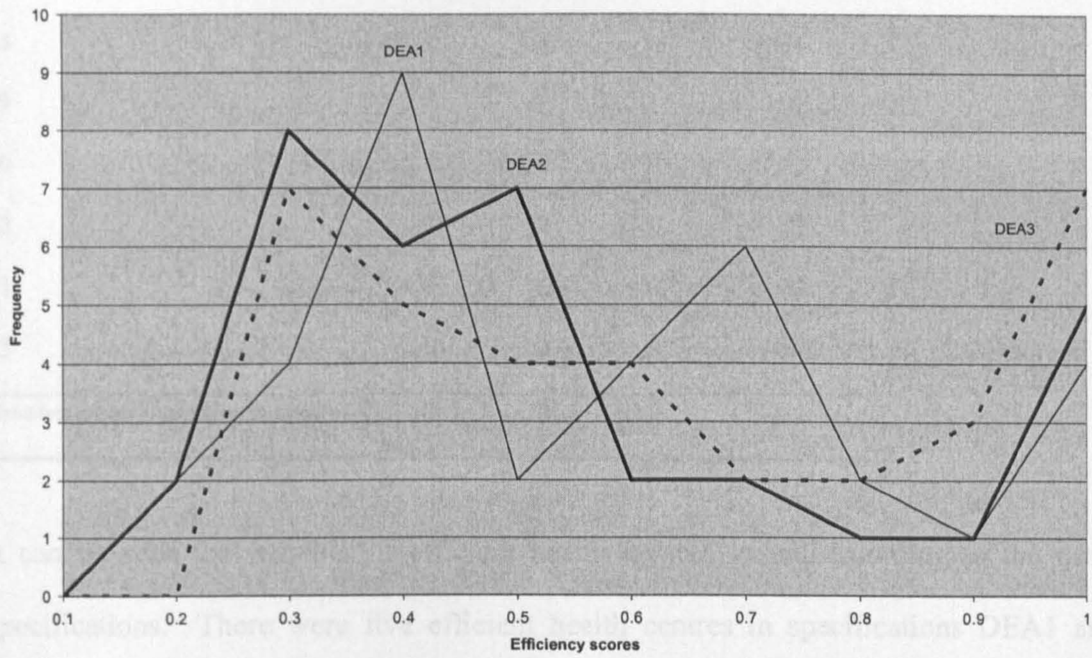


Table 45 shows that the majority of health centres, 30, 30 and 27 for specifications DEA1 and DEA2, and DEA3 respectively, exhibited IRS (implying that they are too small), four, four and six of the centres under specifications DEA1, DEA2 and DEA3 respectively exhibited DRS (implying that they are too large) and only one centre under specification DEA3 was the 'right' size.

Table 45: Returns to scale in the health centres (models DEA1-3)

Types of returns to scale	Number of health centres		
	DEA1	DEA2	DEA3
Increasing returns to scale	30 (88.2%)	30 (88.2%)	27 (79.4%)
Constant returns to scale	4 (11.8%)	4 (11.8%)	6 (17.6%)
Decreasing returns to scale	0 (0%)	0 (0%)	1 (3.0%)

Table 46 presents the efficient health centres by specification and number of times that each of these efficient health centres act as peers.

Table 46: Efficient health centres and the number of times they are a peer

Efficient health centre	Specification			Summation
	DEA1	DEA2	DEA3	
12	29	29	27	85
14	0	26	25	51
19	0	0[75]	1	1
20	3	0	10	13
22	11	0	0	11
23	9	13	10	32
33	10	3	9	22
Number of efficient health centres	5	5	6	

It can be seen that number of efficient health centres varied according to the three specifications. There were five efficient health centres in specifications DEA1 and DEA2 while there were six efficient centres in specification DEA3. The greater the number of input and output variables in a specification, the greater the number of efficient health centres. This finding is consistent with the dimensionality issue; increasing the number of dimensions used in the characterisation of production reduces the discriminatory power of the analysis, increasing measured efficiency and the number of health centres identified as fully efficient.

Among the health centres, seven of them were efficient and acted as peers between 1 and 85 times across all three specifications. However, only health centres 12, 23 and 33 were efficient across all three models; unit 12 is located in Moulvi Bazar whereas units

[75] Delivery unit #19 is efficient but does not act as a peer for any other delivery unit(s).

23 and 33 are located in Brahmanbaria. It is interesting to note that health centre 20, which was efficient in specification DEA2 and DEA3 but not DEA1, had a score of 0.425 in specification DEA1. Health centre 12 acted a peer the greatest number of times: 29, 29 and 25 for specifications DEA1, DEA2 and DEA3 respectively – a total of 85 times.

9.4.2 SFA models

The mean efficiency under models SFA1 and SFA2 was 0.998 and 0.929 respectively. The output file produced by Frontier 4.1 (Coelli 1996b) reported that "The likelihood value is less than that obtained using OLS! - try again using different starting values". As stated in Chapter 3, it is possible to specify the starting values in the instruction file of Frontier Version 4.1. Therefore, the candidate specified the starting values manually, but the output file reported the same message. The candidate contacted Tim Coelli, the author of Frontier via email for advice, who in response commented on the problems of small samples, and the fact that noise can be a particular problem in developing country data sets (Tim Coelli, personal communication 2005). He advised running an OLS regression, saving and plotting the residuals in order to identify any outliers[76].

Using SPSS, one case where the prediction was three standard deviations or more from the mean value of the dependent was identified (health centre number 23). This case was dropped from the analysis and the SFA models re-run. This made no difference to the findings, suggesting that the small sample size is the over-riding problem. It was

[76] The presence of outliers (that is, the presence of large residual variation) in the sample can cause stochastic frontier models to perceive that there is too much noise in the data and therefore may find little or no inefficiency in the sample, even in cases where there is some. As a result, all units may appear to be almost 100% efficient. In this way, the main potential advantage of SFA of decomposing the residual into noise and inefficiency has turned to be a great disadvantage as it fails to differentiate between units' efficiency.

reported in Chapter 3 that in empirical applications samples of size under 30 are usually considered to be small. It appears that in this instance the sample size of 34 health centres is too small. It is worth recalling that this sample was selected on the basis of available resources in light of the range of other costing activities taking place as part of the larger project described in Chapter 1[77].

9.4.3 Stability of efficiency assessment between the three DEA specifications

This section focuses on the stability of efficiency scores across the specifications. This will be explored in two ways: efficiency scores and efficiency ranking (Table 47). Because the efficiency scores are not normally distributed, it is important to consider the ranking of the scores as well. Moreover, ranking, unlike the score, does not depend on the number of variables included in the specifications.

There is a high degree of correlation between the scores and ranks of the three specifications (upwards of $r = 0.609$ and $r = 0.742$ respectively). Equally, there is a high degree of correlation between the unit costs presented in the preceding chapter and the scores and ranks of the three DEA specifications (upwards of $r = -0.688$ and $r = -0.792$ respectively). The relationship is negative illustrating that as efficiency increases, unit costs fall.

[77] Cost data were also collected from nine sub-district hospitals and three district hospitals. Furthermore, approximately 600 patient records were abstracted, 800 exit interviews were administered to caregivers taking children to receive vaccination services, 600 interviews were administered to caregivers of children with a vaccine-preventable disease treated on an out-patient basis and 75 interviews were administered to caregivers of children with vaccine-preventable diseases treated on an in-patient basis. Finally, approximately 30 interviews were conducted with physicians to ascertain how they usually treat children with vaccine-preventable diseases. However, it is also worth noting that a sample of 34 health centres is generous in comparison to most cost and cost-effectiveness analyses.

Table 47: Correlations of unit cost and DEA results[78]

	Unit costs	DEA1	DEA2	DEA3
Scores				
Unit costs	1.000			
DEA1	-0.841	1.000		
DEA2	-0.688	0.609	1.000	
DEA3	-0.750	0.810	0.884	1.000
Ranks				
Unit costs	1.000			
DEA1	-0.900	1.000		
DEA2	-0.811	0.742	1.000	
DEA3	-0.792	0.866	0.915	1.000

While correlations describe overall relationships, they are not a satisfactory way to examine the changes in efficiency scores across different methods and specifications, as they do not show what happens to individual vaccination sites' scores (Jacobs 2001). Therefore, it is worth considering the effect of alternative specifications on the efficiency estimates for individual delivery units (Street 2003).

Table 48 presents details of the maximum change in the efficiency score and rank across specifications. It was found that of 34 health centres, there were three whose efficiency did not vary between DEA1 and DEA2 (they were efficient under all three specifications). Among the remaining 31 health centres, the maximum difference in the efficiency score which an individual health centre obtained was 0.626, which equated to a change in rank of 20 places. However, there was an even greater change in ranking when comparing the results of the unit costs and specifications DEA2 and DEA3; health

[78] All correlations are significant at the 0.01 level (2-tailed).

centre 14 had a unit cost of \$1.25 per visit which ranked it 26th, while it was found to be fully efficient in specifications DEA2 and DEA3.

Table 48: Maximum changes in individual estimates of efficiency

	Unit costs	DEA1	DEA2	DEA3
Scores				
DEA1	NA	0		
DEA2	NA	0.626	0	
DEA3	NA	0.626	0.575	0
Ranks				
Unit costs	0			
DEA1	15	0		
DEA2	25	20	0	
DEA3	25	20	17	0

9.4.4 Analysis of environmental variables

Table 49 gives the technical, 'pure' technical and scale efficiency of the health centres by district.

Table 49: Technical, 'pure' technical and scale efficiency of health centres by location (models DEA1-3)

Location	Technical efficiency			'Pure' technical efficiency			Scale efficiency		
	DEA1	DEA2	DEA3	DEA1	DEA2	DEA3	DEA1	DEA2	DEA3
Brahmanbaria	0.65	0.65	0.81	0.74	0.79	0.88	0.87	0.81	0.91
Chandpur	0.30	0.28	0.34	0.55	0.54	0.57	0.55	0.59	0.63
Moulvi Bazar	0.63	0.50	0.60	0.77	0.72	0.75	0.80	0.67	0.78
Total	0.53	0.48	0.59	0.69	0.68	0.73	0.75	0.69	0.78

Table 50 shows that technical, 'pure' technical and scale efficiency varied systematically according to the location of the health centres. The districts were selected to reflect high (e.g. Chandpur), medium (e.g. Moulvi Bazar) and low (Brahmanbaria) performing districts using a variety of indicators of disease, vaccination coverage, health service provision and access to health services. The efficiency data presented in this thesis are not consistent with this classification. The results presented in Table 50 are for specification DEA1. However, the findings were also significant for specifications DEA2 and DEA3.

Table 50: Significance of selected environmental variables and technical, 'pure' technical and scale efficiency (models DEA1-3)

	Technical efficiency		'Pure' technical efficiency		Scale efficiency	
Environmental variable	F-test	Kruskal-Wallis	F-test	Kruskal-Wallis	F-test	Kruskal-Wallis
Location, i.e. districts (Brahmanbaria, Chandpur, MoulviBazar)	7.817 (0.002)	13.465 (0.001)	6.513 (0.004)	9.353 (0.009)	2.732 (0.081)	4.883 (0.087)

9.5 Policy implications

A potential strength of DEA is its diagnostic capability; DEA provides clues on how each inefficient health centre can improve efficiency in line with their peers. Accordingly, this section focuses on linking the results of DEA with two related policy implications: the level of potential savings and targets for performance.

9.5.1 Potential savings

The mean efficiency scores were 0.53, 0.48 and 0.59 respectively for specifications DEA1, DEA2 and DEA3. In other words, if health centres were technically efficient and operated at the correct scale, total costs, on average, could have been reduced by 47%, and specifically, expenditure on staff and drugs could have been reduced by between 41% - 52%, without sacrificing the current level of outputs produced. Given that the mean total cost of the 34 health centres was \$8,812, this equates to an average saving per health centre of \$4,142.

To achieve efficiency of all centres, the inputs, expenditure on staff and drugs, have to be reduced by a different percentage. In specifications DEA2 and DEA3, drug expenditure requires a greater percentage reduction than expenditure on staff. Therefore managing drug expenditure plays the most important role in efficiency improvement.[79]

9.5.2 Targets for efficiency improvement

DEA results can be used as a managerial tool to improve efficiency of health centres as it provides targets to achieve efficiency for each health centre. In light of the inevitable transaction costs of implementing efficiency improvement programmes, it is reasonable to target improvement to those centres which have most to gain, i.e. currently the least efficient.

Table 51 provides an example of identifying target health centres using specification DEA2. In this specification, there were 30 inefficient health centres of which nine (30%) accounted for almost 40% of the technical inefficiency of the whole sample. After identifying these target health centres, information from the DEA results can be used to set unique target levels for each type of input which inefficient health centres need to meet in order to become efficient.

---

[79] Staff should perhaps be excluded from this analysis given that wages are generally beyond the control of the health centres. This is an issue that will be explored in more detail in the next chapter.

Table 51: Nine most inefficient health centres using specification DEA2

Health centre	Share of inefficiency %	Cumulative inefficiency %
10	4.67%	4.67%
29	4.67%	9.34%
4	4.52%	13.86%
5	4.49%	18.35%
6	4.30%	22.65%
32	4.15%	26.80%
2	4.13%	30.93%
25	4.12%	35.05%
8	4.11%	39.16%

Table 52 presents actual and target resource use for these nine health centres. For example health centre 10 which spent $3,389 on staff and $3,311 on drugs was inefficient. In order to become an efficient health centre it needs to reduce consumption to $1,864 on staff and $1,821 on drugs.

Table 52: Actual and target resource use for the nine most inefficient health centres, using specification DEA2

	Actual resource use		Target resource use	
Health centre	Staff	Drugs	Staff	Drugs
10	3,389	3,311	1,864	1,821
29	6,441	2,780	2,969	1,281
4	3,854	3,900	1,821	1,843
5	2,558	2,314	1,961	1,774
6	5,871	6,058	1,824	1,882
32	3,701	1,755	2,837	1,345
2	2,711	3,660	1,486	2,006
25	3,869	2,191	2,590	1,467
8	3,588	4,010	1,701	1,901

9.6 Summary

- Unfortunately the GoB faces significant resource constraints in funding the ESP;
- Previous reports have found that the potential for additional resource mobilisation is limited and have suggested that improvements in the efficiency of health care services must be a critical component of efforts to provide the ESP to the whole population;
- Given the Government's stated objective to mobilise additional resources via improvements in the efficiency of health facilities, an input-orientated specification under VRS was adopted to assess the technical efficiency of a sample of 34 representative health centres;
- Technical efficiency was 0.53 in specification DEA1, 0.48 in specification DEA2 and 0.58 in specification DEA3. By decomposing this technical measure into 'pure' technical and scale efficiency, it was shown that slightly more of the inefficiency is due to health centres using too many inputs in treating the number of patients seen rather than operating at the wrong size;
- The majority of health centres exhibited increasing returns to scale;
- There is a high degree of correlation between the scores and ranks of the three specifications and the unit costs presented in the preceding chapter. However, among the inefficient health centres, the maximum difference in the efficiency score which an individual health centre obtained was 0.626, which equated to a change in rank of 20 places;
- Technical, 'pure' technical and scale efficiency varied systematically according to the location of the health centres.

The next chapter the results presented in Chapters 6 – 9 are discussed.

Chapter 10

DISCUSSION OF FINDINGS

This thesis set out to explore whether, and the extent to which, interventions to provide health care in Bangladesh are delivered efficiently. Previous chapters have presented data on the cost and efficiency of the provision of primary health care in urban and rural Bangladesh. In this chapter, the study results are discussed. It is divided into two sections. The first discusses methodological issues, in particular the limitations of the data, analysis and interpretation. The second section discusses the main findings of the thesis. In particular, it focuses on the: variation in unit cost data; variation in efficiency estimates; implications of inefficiency on CEA; constraints to efficiency improvement (particularly in relation to human resources); and cost-effectiveness of efficiency improvement programmes.

10.1 Methodological matters

Theoretically, in order to measure the technical and allocative efficiency of a firm, it is necessary to know the underlying production and cost functions for that firm. This requirement poses significant problems for 'real-world' applications in health care. First, the extreme heterogeneity and complexity of health care interventions effectively rules out the development of engineering-type production functions for all but the simplest interventions. If bottom-up engineering functions cannot be described, then some form of statistically derived estimation from observed data becomes necessary. In spite of this, it can only be assumed that a statistically estimated production or cost function reflects the underlying, 'true' function if it is assumed that firms are always technically and allocatively efficient in their operation. This thesis has shown that there are good reasons to conclude that health care facilities are unlikely to meet these conditions in reality. Consequently it is not possible to observe the isoquant or production possibilities frontier of an efficient facility, and any estimated production or cost function cannot be assumed to fully represent the production frontier or the underlying cost function.

Frontier estimation methods involve the estimation of an efficiency frontier (or envelopment surface) from an observed sample of data, based upon best performance within the sample. The efficiency of other facilities in the sample is defined relative to these best performers. Specifically, measurement of the deviation of individual firms from this frontier enables the calculation of relative efficiency scores and the computation of potential efficiency gains if all units could achieve best performance levels.

There are two major features that distinguish alternative empirical approaches for estimating the production frontier: whether they are parametric or not; and whether they are deterministic or stochastic. Parametric methods assume a specific functional form for the frontier, whereas non-parametric methods do not; and deterministic methods assume that the distance of a unit from its frontier is a result of inefficiency whereas stochastic methods assume that this is also partially due to random error. This thesis used parametric and non-parametric techniques to measure the efficiency of primary health care in Bangladesh. In doing so, it challenged the assumption of technical efficiency implicit in CEA.

The remainder of this section focuses on the issue of measurement error. In particular it considers the potential for measurement error in the data sets used in this thesis, discusses the potential implications of measurement error on DEA and SFA and reflects on a number of solutions that have been proposed in the literature to mitigate the potential effects of measurement error. In addition, it considers some of the issues when interpreting efficiency estimates and possible extensions to the models presented in chapters 7 and 9. Finally, it discusses a number of constraints related to conducting this research within the larger WHO- and DFID-funded projects.

10.1.1 Measurement error and other 'noise'

Measurement errors are errors in reading, calculating or recording a numerical value (Everitt 1995). It is the difference between observed values of a variable recorded under similar conditions and some fixed true value. Clearly errors in the original data cannot usually be rectified, but errors introduced at a later stage can be minimised if certain steps are taken in the process starting from the collection of the data to its analysis.

Appendices 5 and 15 present the data collection tools used at the vaccination delivery units and the health centres respectively. A coding sheet for data was prepared for both. Data were entered into Excel and SPSS. Errors in recorded data are common. For example, the recorded values may be wrong because of confusion over the correct units of measurement, digits may be transposed when data are transcribed, or data may be mistyped when being entered onto a computer. Data checking aims to identify and, if possible, rectify such errors in the data. To minimise these types of errors, the data were entered twice by two different people. Differences between the two files were checked and the 'correct' values obtained by consulting the original questionnaires. As noted in Chapters 6 and 8, before beginning the main cost and efficiency analyses, the data were screened, which involved producing histograms of variables, and pairs of variables were inspected by scatter diagrams. These plots gave a first idea of the average value, the variability, the shape of the distribution, and whether there were any outlying or missing values. Finally, analyses were conducted on Excel, SPSS, DEAP and FRONTIER.

As stated above, errors in the original data cannot usually be rectified. A report by Uddin et al. (2002) provides an indication of the quality of the data collected and analysed. As part of health sector reforms in Bangladesh (see Chapter 5), the Unified Management Information System Unit of the Directorate General of Health Services of the MOHFW introduced a new record-keeping and reporting system. The objective of the new system was to record and report on the ESP offered at the upazila level and below. From February 2000, service providers at the union level began to use the new record-keeping and reporting tools. Uddin et al. (2002) assessed the extent to which the new system was functioning at the union level.

Monitoring was conducted in 36 randomly selected health centres of Chittagong district and 15 randomly selected health centres of Jessore district during February 2000-March 2001. It was observed that the new record-keeping and reporting tools were being used, and fulfilled the record-keeping and reporting requirements at the union level of both the districts. The service providers committed less than 10% of errors, such as omitting data and entering incorrect data, when they were observed during service delivery, and the rate of errors increased to as much as 34% when they were not. The workload during peak hours, inadequate training and inadequate supervision contributed to such errors. Unclear instructions from the national level to discontinue the use of some record-keeping and reporting tools of the previous system also contributed to errors. With systematic monitoring and supervisory support, the authors believed that the extent of errors could be reduced gradually.

It is therefore apparent that the potential for measurement errors existed. This potential has been exacerbated among the health centres included in this thesis by the small sample size. For these reasons, it is pertinent to investigate the possible influence of measurement error and other 'noise' on the study findings.

10.1.2 Influence of measurement error and other 'noise'

One of the primary criticisms of deterministic frontier models, such as DEA, is that no account is taken of the possible influence of measurement errors and other 'noise' upon the frontier (Coelli et al. 1998). All deviations from the frontier are assumed to be the result of technical inefficiency. Yet, as with regression analysis, deviations from the frontier may be due to a number of factors other than inefficiency such as omitted variables and measurement errors.

These factors are not testable. As a result, interpreting DEA scores as measures of efficiency requires a high degree of 'blind' faith in the model. This is because, when there are outliers the method envelops the outermost observations without asking whether these observations are genuine or the result of an error. Even a single outlier can result in finding huge inefficiencies for most comparators without this being necessarily true. This is particularly the case where an observation contains inputs which are significantly smaller, or outputs which are significantly larger, than other observations employing a similar input mix or producing a similar output level (Coelli et al. 1998; Cooper et al. 2003).

SFA recognises the presence of errors and aims, in principle, to separate these error components from the measures of inefficiency. In practice, this effort is not always successful as, often the estimated inefficiency component represents a small fraction of the overall residual variation (Kumbhakar and Lovell 2000). This practical nuance may cause many problems in the analysis. For example, it can make SFA vulnerable to outliers.

The presence of outliers in the sample can cause the stochastic frontier model to indicate that there is too much noise in the data and therefore may find little or no inefficiency in the sample, even in cases where there is some. As a result, all units may appear to be almost 100% efficient, which at first sight appeared to be the problem faced by the stochastic frontier models used in this thesis in Chapter 9 (although this appears to be due primarily to the small sample size). In this way, the main potential advantage of SFA of decomposing the residual into noise and inefficiency becomes a disadvantage as it fails to differentiate between units' efficiency. There are other instances in which the stochastic frontier model ceases to have the role it is intended to have. Sometimes, SFA can suggest that the noise residual has been drawn from a distribution with a very small variance. Consequently, deviations from the frontier are almost entirely due to the residual supposed to measure inefficiency. In these cases SFA collapses to a deterministic form, with the result that the frontier 'envelopes' the observations from below, resulting in at least one unit estimated to be 100% efficient.

Therefore, outliers can cause problems in both SFA and DEA but for completely different reasons: while SFA can fail to find any inefficiency at all, DEA is likely to find too much inefficiency in the sample.

10.1.3 Approaches to mitigate the potential effects of measurement error

Outliers could be removed from the analysis to remedy this problem and find 'sensible' scores of inefficiency, although any such choice would be inherently arbitrary and difficult to be make. Timmer (1971) adopted the suggestion of Aigner and Chu (1968) of deleting a percentage of the sample firms closest to the estimated frontier, and re-estimated the frontier using the reduced sample. However, the arbitrary nature of the selection of a percentage of observations to delete has meant that this so-called 'probabilistic' frontier approach has not been widely followed (Coelli, Rao and Battese 1998).

An analysis of the data sets for outliers, which identifies observations with inputs or outputs lying more than three standard deviations on either side of the sample mean, indicated that the vaccination delivery units and health centres used inputs and produce outputs commensurate with size. Therefore, no significant outliers were discovered. However, future research could include the application of more sophisticated methods to identify influential outliers in DEA using modifications suggested by Wilson (1995) and Lovell et al. (1993).

This thesis also sought to account for the deviations from the identified frontiers by performing a two-stage approach whereby the resulting efficiency scores were analysed against an array of independent factors that may affect efficiency but are out of the managers' or policy-makers' direct control. While the two-stage approach recommended by Coelli, Rao and Battese (1998) was used in this thesis, alternative techniques exist, which were reviewed in Chapter 3.

Valdmanis (1992) (based on Nunamaker 1985) suggests, as a possible answer to these problems, that researchers run a number of different models for each data set and evaluate the sensitivity of the results to changes in model specification. These changes may take the form of alternative input and output definitions. While this approach does not address the issue of measurement error *per se*, if different methods suggest similar directions for results then the validity of such findings is enhanced. The purpose of the sensitivity analysis would be to assess whether the ranking and efficiency of an individual firm is variable-specific (or model-specific) or whether the results are robust to changes in data set specification. Valdmanis (1992) cautions that '... for a model to be considered robust, it must be shown that minor changes in the list of variables cannot alter fundamentally the conclusions of the DEA model'.

Another method of evaluation is to compare the results of DEA and SFA applied to the same data sets. As Hollingsworth (2003) notes that "... given the limitations of frontier techniques at present it may be that they are best employed in tandem". As a result, the best approach is the use of different techniques in tandem. Thus both methods serve as signalling devices (Jacobs 2001). To the extent that there is no *a priori* reason to prefer one methodology over another, it seems prudent to analyse efficiency using a broad variety of methods to 'cross-check'.

This thesis has examined the consistency and robustness of efficiency scores across DEA and SFA techniques when applied to the same data sets. Sensitivity analysis was carried out within the DEA and SFA models by changing the model specifications (omitting and including different variables) and testing for the robustness of the results (Jacobs 2001; Valdmanis 1992; Nunamaker 1985). While models proved to be relatively robust in this respect, there was some inconsistency across the different methodologies. Caution is therefore warranted against literal interpretations of units' efficiency scores and rankings. Reasonable correlations suggested convergent validity. However, while on average, scores and rankings were fairly stable across specifications, some units experienced dramatic movement in where they were ranked. This implies that it would be inadvisable to rely on a single specification if the objective was to set unit specific efficiency targets, such as those presented in Chapter 7 and 9.

The different efficiency scores should not therefore be interpreted as accurate point estimates of efficiency, but might more usefully be interpreted as indicating general trends in inefficiency for certain units (Jacobs 2001). The point estimates of inefficiency in either method are indeed sensitive to specification, measurement and data errors. However, when several specifications were used, general trends could be discerned as to which units usually came out as being more efficient and which ones generally emerged as inefficient. It is therefore imperative that several specifications be employed to gauge an overall picture of efficiency. It might also be useful to explore a number of DEA re-sampling techniques (including jack-knifing and bootstrapping) which have been developed to obtain more statistically robust measures of estimated frontiers (e.g. Ferrier and Hirschberg 1997; Atkinson and Wilson 1995).

Ultimately, data accuracy is paramount to parametric and non-parametric analyses as inaccurate data in, for instance the DEA methodology, will affect not only that unit's efficiency rating but also potentially the efficiency ratings of other units as well. The level of random 'noise' is a reflection of the quality of the data and will affect the ability to measure efficiency. Nevertheless, in spite of these problems, it should be recognised that the approaches used in this thesis are only relative and that further efficiency gains could still be possible beyond the identified frontiers.

From the GoB's point of view, improving on data deficiencies would probably contribute more to better efficiency estimates than further experimentation with alternate specifications and estimation techniques. In particular, data on outputs and quality of care and outcome indicators would be important. There may not be a strong self-interest in the accurate reporting of data and as such incentives might be needed to ensure this.

10.1.4 Interpreting efficiency scores

As noted in Chapters 7 and 9, a potential strength of DEA is that it can identify potential efficiency gains and the targets that need to be met in order to realise such gains. However, while these methods prove useful diagnostic tools it would be inappropriate to base funding and resource decisions or indeed efficiency targets entirely on the efficiency estimates arrived at (Skinner 1994; Newhouse 1994; Hadley and Zuckerman 1994). Relative efficiency assessment and target setting based on only one method may provide inappropriate incentives to managers. Therefore, given the limitations of frontier techniques at present it may be that they are best employed in tandem when possible; if different methods suggest similar directions for results then the validity of such findings is enhanced.

Particularly under DEA, the *Pareto* efficiency criterion has the advantage of regarding each separate input and output as being equal in value thus allowing units to be rated along their best dimensions. However, this same advantage could also create the perverse incentive for managers to act in a dysfunctional manner trying to improve their efficiency rating by engaging in creative accounting, and alteration of the input / output mix (Nunamaker 1985) if DEA performance measures were incorporated into an incentive scheme.[80]

Poorly constructed output measures in any method could also lead to units devoting more resources to achieving low priority outputs simply to improve their perceived efficiency. For example, the different weighting schemes used in Chapter 7 illustrate that the methods used to aggregate different outputs into a single measure of output, which is required for stochastic frontier methods, can influence findings. Similarly, failing to list ESP-specific outputs, rather than the number of age-specific visits, weakens the ability of the techniques to set targets to improve the efficiency of the ESP. Equally, the examination of vaccination services in isolation is also of concern, when many of the sites provided a range of other health services. It is plausible that, allowing a certain level of inefficiency in vaccination services may allow sufficient flexibility for a health centre deliver to other services more efficiently.

[80] An incentive scheme has been introduced by the Global Alliance for Vaccines and Immunization as a means to increase vaccination coverage rates. The scheme is supported by a data quality audit exercise, which is used to verify reported performance (GAVI 2005).

Consideration needs to be given to whether units should be allowed some time to adjust their activities in such a way that they are more directed at agreed priority outputs before their relative efficiency is assessed. However, this should not detract from the usefulness of a baseline assessment of efficiency which can help inform the process. If these or similar results were to form the basis of a performance target-setting regime, careful consideration would have to be given to the potential incentives provided by the implicit weights provided by the model selections.

A note of caution with regards interpreting targets based on efficient 'peers'. Efficient peers give a measure of efficiency that is empirically obtainable in a given scenario (e.g. given available resources and institutional set-up), as firms are directly compared against a peer or combination of peers. Hence one can compare the efficiency of individual facilities or administrative areas against realistic benchmarks. For example, unions, upazilas or zones which are classified as efficient health service delivery units, or 'peers', could become 'model' areas where new policies and procedures for improving efficiency, quality, promoting community involvement, and fostering sustainability are implemented and closely monitored. Such administrative areas would have the potential for becoming training sites where field staff could be trained by persons working with programmes currently engaged in the effective provision of the ESP. These same programme sites could offer strong potential for carrying out local operations research activities to strengthen the efficiency of service delivery. However, in some instances, the hypothetical target unit on the frontier will consist of a combination of the largest and smallest efficient 'peer' units. Thus, there is a pedagogical problem because the manager of the in-between-sized unit may not find it interesting to compare himself neither to the largest units, nor to the smallest units. The results may therefore, in some instances, give little help to practical policy questions.

10.1.5 Extensions to the models presented

The models considered in this thesis are cross-sectional in the sense that each vaccination delivery unit and health centre was observed at a single point in time. In a panel data set, each unit is observed not only once but over a period of time and thereby the ability to make statistical inferences increases. Panel data models tend to be less susceptible to multicollinearity and degrees of freedom problems (Coelli et al. 1998; Cooper 2003). Furthermore, if assumptions about the functional form of the distribution of the inefficiency effects are difficult to justify, and the functional form of the relationship between cost and outputs requires a lot of data for the estimation to proceed, it is desirable to use panel data analysis. In particular, panel data analysis avoids making strong distributional assumptions about the inefficiency effects. By contrast, these effects are usually assumed to be either fixed or random – a number of these types of studies were reviewed in Chapter 4. This means that in the fixed effects models the firm-specific inefficiency effects are treated as fixed (Skinner 1994), while in random effects models the firm-specific inefficiency effects are treated as realisations of some random process (common for all facilities). The standard error of each effect can then be used to make assessments of how far each unit differs from the 'best practice' units.

A specific issue that arises in panel data is that of modelling the time aspect of inefficiency. In traditional panel data models efficiency is assumed to remain unchanged over time. This assumption may be difficult to justify particularly when observing the same units over a long period of time. For example, it is natural to think that efficiency will improve over time as new working practices are developed, and differences in efficiency may narrow if units can learn from 'doing' and / or each other's practices. Chapter 7 suggested that the efficiency of the vaccination delivery units was related to such learning effects. With panel data analysis it is possible to not only the check whether a unit's efficiency is improving over time relative to the frontier, but also whether the frontier itself is shifting. DEA Malmquist indices can also be used to examine productivity change over time (Hollingsworth, Dawson and Maniadakis 1999).

Scope exists to develop a panel data set of vaccination services, as data could be collected alongside routine national coverage evaluation surveys. Not only would this enable extension to the analyses presented in this thesis, but it would likely improve knowledge about variation for different scales of production and settings. These data would inform whether and how costs vary with the level of production of vaccination services, which would guide decisions about whether and how much to expand existing vaccination programmes. It should also allow a better-defined relationship between costs and effects of provision of existing services and new services to be questioned.

10.1.6 Input- versus output-orientated models

Given the GoB's stated objective to mobilise additional resources via improvements in the efficiency of health facilities, input-orientated DEA specifications were adopted in Chapters 7 and 9. However, as stated in Chapter 7, an input-orientation runs contrary to the Government's stated objective of full vaccination coverage. And while an input-orientation is consistent with the fact that treating fewer is clearly better, in the sense that it may reflect successful health promotion and prevention programmes, it needs to be balanced against the fact that utilisation rates of government health facilities are low, suggesting there could be considerable unmet need (see Chapter 5).

Input-orientated measures help identify by how much input quantities can be proportionally reduced without changing the output quantities produced. In essence this reveals the quantity of variable resources that could be reduced, as fixed resources cannot be reduced in the short-term. The alternative output-orientated specification would reveal by how much output quantities can be proportionally expanded without altering the input quantities used. This is perhaps more palatable from a policy-makers perspective. However, as noted in Chapter 5, under-utilisation of union-level facilities is well documented, so the recommendations would have to be demand-led, although they would need to ensure that supply-side considerations have been met.

The indivisibility of inputs may explain why reduction or substitution predicted by the models would not take place in practice. The context of the health care setting concerned, in particular the way in which the supply and demand for factor inputs is regulated, may also have important implications for the divisibility of inputs. For example, although it might improve technical efficiency for staff to be made redundant, contracts and trade unions obviously ensure that this cannot be enacted in the short-term. If an output-orientation were adopted, it would hold constant all inputs, which would 'solve' this particular problem. But equally there are challenges in interpreting output-based models, particularly within the health sector as demand is stochastic.

10.1.7 Constraints of conducting this research within a larger project

As stated in Chapter 1, these data were not collected for the purpose of parametric and non-parametric efficiency measurement analyses. Rather, they were collected for the purpose of CEAs. As illustrated in Chapter 2, such studies do not examine the relative efficiency of production units as one of their objectives because efficiency is already assumed. The focus of CEAs tends to be on providing a point estimate of the cost of a given service or intervention, where wider consideration of efficiency requires comparison of a sample of several production units. Therefore, it is important to ask what the immediate implications were for this thesis.

The main implication is perhaps the sample size of each case study. The case studies examined data from 110 vaccination delivery sites and 34 health centres providing basic primary care. Given that the sample sizes for both case studies were not selected for the purposes of efficiency analyses it was thus fortunate that the sample of 110 vaccination delivery units was adequate for both parametric and non-parametric analyses. The same fortune did not extend to the 34 health centres; while this thesis has illustrated that DEA can be implemented on a relatively small data set, Chapter 9 highlighted the problems of conducting parametric analyses on a small sample.

Although the power to differentiate firms diminishes as the sample size falls, DEA still gave meaningful results with the limited sample. However, it should be noted that in DEA, the efficiency scores tend to be sensitive to the choice of input and output variables and, in some circumstances, relatively small samples may lead to relatively inefficient firms defining the frontier. This is because there is likely to be at least one factor (use of input or production of an output) for which a firm is distinct. Even if this is not in fact an important variable, its use in a DEA could put that firm on the frontier. On the other hand, SFA requires a larger minimum sample size in order to stand up to statistical testing. Indeed, Chapter 9 illustrated that the sample of 34 health centres was too small. Were this research undertaken again, a larger sample of health centres should be sought.

A brief note regarding missing values is also merited at this point. In general terms, there are three possible approaches for analysing datasets with incomplete or missing observations. The simplest solution is to ignore the problem and work only with the subset of observations with complete data. Alternatively, the analyst can use a range of alternative statistical techniques and perform a single imputation for each missing value, whereby the incomplete observations are replaced with a single imputed value. The final approach is to use methods such as multiple imputation using Markov chain Monte Carlo techniques (Briggs et al. 2003; Manca and Palmer 2005).

In both case studies, a *complete case analysis* approach was adopted in which all cases with missing data were excluded from the analysis. Given that the incomplete observations in both case studies were missing completely at random, there is no reason to believe the results are biased in any way. As evidenced by the analyses presented in Chapter 7, the exclusion of 22 vaccination delivery sites because of missing observations did not affect the ability to perform parametric and non-parametric efficiency measurement analyses. Chapter 9 highlighted that a sample size of 34 health centres was too small to perform parametric analyses. The use of methods to impute values for missing observations in provide a full sample of 36 health centres with which to work with would not have resolved this problem; a sample of 36 health centres is equally as small, and restricts analysis to the degree.

The second implication relates to the selection of the inputs and outputs, which were therefore essentially *post hoc*. As noted above, improvements in the output data should be seen as a priority of future similar studies, but which have efficiency measurement as their main focus (or at least an objective). Similarly, the choice of environmental variables was also essentially *post hoc*. It would have been interesting to complement the quantitative analyses presented here with qualitative research undertaken to better understand the unit-level managerial characteristics of good practice. Unfortunately this was not possible given the requirements of the larger project.

Related to the sample size issue, it might have been preferable to have collected data from urban and rural settings in Bangladesh on the same service. However, in retrospect, the fact that two different services have been investigated and both found to exhibit large degrees of inefficiency strengthens the study. Both of the services are delivered at the primary level. Future research of efficiency using methods similar to those followed in this thesis might be extended to also include secondary and tertiary health care facilities to give a more comprehensive indication of the efficiency of health services in Bangladesh. Although the 'Bangladesh health facility efficiency survey' conducted by Rannan-Eliya and Somanathan (2003) collected data from these higher levels of care, they presented simple ratio measures as their indicators of efficiency.

10.2 Discussion of findings

The discussion of findings covers the following aspects: variation in unit costs; variation in technical and scale efficiency; the implications of inefficiency on CEA; the constraints to improving efficiency (particularly in relation to human resources); and the cost-effectiveness of efficiency improvement programmes.

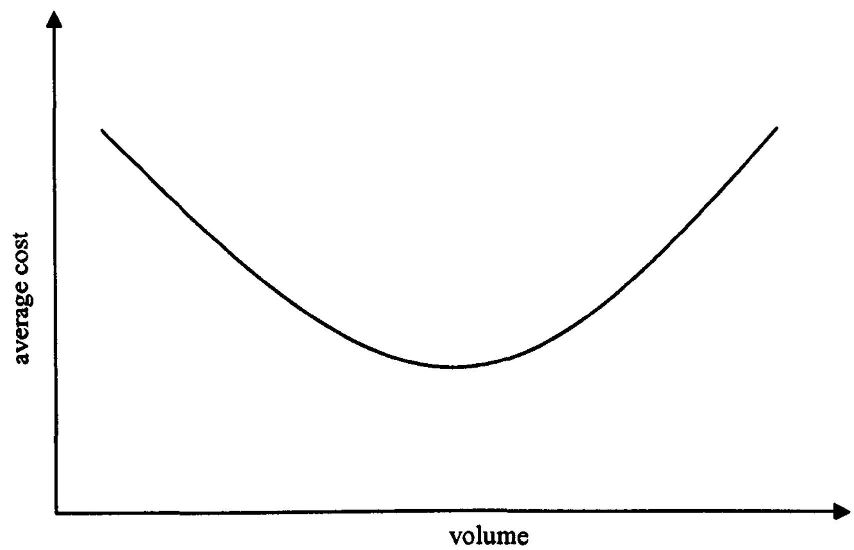
A relatively large body of literature exists presenting cost studies for different levels and types of health care provider in many developing countries (e.g. Barnum and Kutzin 1993; Adam et al. 2003). The most commonly used technique for measuring costs of public health interventions in developing countries is the accounting approach, which was used in Chapters 6 and 8.

10.2.1 Variation in unit costs

Chapters 6 and 8 focussed on the 'cost' in cost-effectiveness and in particular the need to describe variation in the cost of providing health care services. Chapter 6 reported a 40-fold difference in the cost per dose administered among the 110 vaccination delivery units. Chapter 8 reported a 5.5-fold difference in the cost per visit among the 34 health centres. Importantly, the same methodology was used to cost each of the vaccination delivery units and health centres, which means that methodological inconsistencies can be excluded as a potential source of the variation in the unit costs observed.

These results illustrate that service volume at the vaccination delivery units and health centres appears to be one of the most important factors influencing the unit costs. The findings presented in Chapter 6 and 8 show a clear inverse relationship between the unit costs and service volume. Good managers of health facilities try to choose combinations of personnel (of various types), supplies and other inputs that will minimise the costs for a given volume, at the prevailing rates of pay and prices. When the total costs (of the cheapest combination) are divided by the service volume, an average cost function is derived. When presented in a graph, the curve shows that the average cost first falls and then rises as the service volume increases (Figure 20). It seems that the optimum service volume, corresponding to the lowest average cost, has not yet been reached in either of the samples because neither of the average cost curves showed a trough throughout the range of service volumes studied. Both case studies therefore, suggest that increasing the volume of service (doses or visits), in those units that saw relatively few visits, or administered relatively few doses, would lower the average cost.

Figure 20: Typical curve showing relationship between service volume and average cost



A significant factor contributing to this relationship appears to be the existence of substantial fixed costs associated with the delivery of both services. Although both cost analyses classified staff as a recurrent item, in line with standard costing guidelines, in reality personnel costs are fixed in nature, at least in the short-term, which means that a significant portion of the resources required for the delivery of both services change little, if at all, as the volume increases or decreases. Under these conditions, therefore, the results, as expected, show that up to a certain level of utilisation of resources, a larger volume of services tends to reduce the unit cost. However, there must be some point beyond which a higher volume of service is accompanied by inefficient utilisation of resources, which raises the average cost. This point does not appear to have been reached in any of the units analysed in Chapters 6 and 8.

With respect to Chapter 6, vaccine wastage was noted as another cause of variation in the unit costs. Vaccine wastage is important as it can indicate programme errors. For example, it can highlight that too many drops of OPV or the wrong dosage for other vaccines is used; cold-chain failures or poor logistics; and false reporting of more vaccinations administered than vaccine received. There are also economic implications associated with wastage. If wastage can be reduced without affecting coverage, it can result in significant resource savings for programmes. This is especially true for very poor countries, such as Bangladesh, which do not typically have budgetary flexibility to expand programme financing (Dervaux et al. 2003).

The literature appears to suggest that distance could be important in determining vaccine wastage rates. Unfortunately, no data were available on the distance of the fixed delivery units from the outreach sites with which to test this theory. Nor were data available regarding the distribution of households around these delivery units. However, the outreach sites clearly experienced higher rates of wastage, suggesting that a relationship exists between distance and wastage rates which is worth exploring further.

A study from Benin and Guinea suggests that these problems, if identified, can be addressed for little extra cost, and can result in a rapid increase in vaccination coverage, as well as the more efficient use of other primary health care facilities (Soucat et al. 1997). To investigate differences in the cost of delivering vaccines between centres the authors measured various parameters associated with access to, and availability of, services. In Benin outreach activities increased accessibility (from 77% to 95%) in the worst group, but utilisation remained low when compared to the best group, demonstrating that improved accessibility does not necessarily increase utilisation. Those centres performing best in terms of coverage had effective social mobilisation and channelling strategies, and those performing worst had a severe problem with drop-out rates. In Guinea, the worst performing centres were those for which access was difficult – where the service is available, but not readily accessible. The better accessibility in Guinea was explained by improved outreach activities in the best-performing group. Further studies of this type are required to help inform decision-making regarding the optimal use of additional resources for vaccination services.

The results discussed here provide estimates on the supply-side cost of the ESP. The results are based on the average level of efficiency in the sample zones and unions. Systematic and significant variation in unit costs between production units, can present a powerful basis for benchmarking and for identifying high cost and thus relatively inefficient units. Thus, further efficiency gains might be achieved if, instead, the best performing facilities were taken as the standard at different levels of output. Other facilities would then be helped to achieve similar levels of productivity. Chapter 7 and 9 analysed these data in order to discern whether and to what degree the health facilities were being operated efficiently.

10.2.2 Variation in efficiency

While there has been a recent expansion in the number of efficiency evaluations (Hollingsworth 2003), there remains a dearth of literature on the measurement of efficiency from low- and middle-income countries (see Chapter 4). This is disappointing given the developing world's greater scarcity of resources, which results in the inefficient use of resource exacting a much greater penalty in terms of foregone benefits.

Chapters 7 and 9 presented the efficiency of providing vaccination services in Dhaka and the efficiency of providing primary health care at union-level health centres in rural Bangladesh. The findings illustrated that these services were being provided inefficiently. Not only was there a large degree of technical inefficiency present, but the majority of the units were operating at IRS, which questions the applicability of cost-effectiveness analyses that assume CRS.

With respect to technical efficiency, Chapter 7 illustrated that differences in ownership and type of vaccination delivery units made no difference. However this was not true for technical and scale efficiency, where it was found that NGO-outreach delivery units were the most technically efficient and GoB-fixed delivery units were the most scale efficient. The technical, 'pure' technical and scale efficiency of the delivery units varied systematically by location.

It is difficult to interpret these findings. It appears that delivery units trade-off scale and technical efficiencies. The findings suggest that government-owned units, perhaps due to more centralised control, were better at long-term planning. It was also found that units that had been practicing longer were relatively more scale efficient, which is perhaps attributable to learning effects. This suggests that merging the smaller sites would reduce excessive costs attributed to scale diseconomies. However, mergers should not be pursued too hastily, especially if access to vaccination services would be compromised.

Chapter 9 explored whether location affected the efficiency of health centres. Technical, 'pure' technical and scale efficiency varied systematically according to the location of the health centres. The districts were selected to reflect high (Chandpur), medium (Moulvi Bazar) and low (Brahmanbaria) performing districts using a variety of indicators of disease, vaccination coverage, health service provision and access to health services. The efficiency data presented in this thesis are not consistent with this classification.

It would be interesting to better-understand why location influences efficiency. It may boil down to the management skills of the zonal programme managers. Future analyses should include a qualitative component that attempts to tease out the characteristics of good practice.

10.2.3 Technical efficiency and cost-effectiveness analysis

Chapters 7 and 9 both illustrated that the case studies were delivered at less than full technical efficiency. Failing to account for differing levels of technical, and therefore by definition allocative, efficiency among providers could have significant implications for the validity of the results of economic evaluations. If technical inefficiencies exist it means that a cost-effectiveness ratio does not reflect the minimum efficient point of production at a given level. Not knowing whether, or the degree to which, a cost-effectiveness ratio incorporates technical inefficiency could have a significant impact on decisions. Consider, for example, a facility operating inefficiently – excess costs are incurred given the outputs produced. If these excess costs could be reallocated elsewhere, then there exists the possibility of potential *Pareto* efficiency gains.

This thesis has illustrated that inefficiencies occur in the provision of preventive and basic curative services[81]. In the case of successful preventive programmes, the need for the associated curative services will be averted. A question which therefore needs to be addressed is whether accounting for inefficiencies in both of these elements of the 'cost' of the cost-effectiveness ratio would cancel each other out. The extent to which this might happen is related to health care seeking behaviour.

10.2.4 Scale efficiency and cost-effectiveness analysis

Using DEA, Chapters 7 and 9 illustrated that the case studies exhibited VRS, thus violating the frequently stated assumption of CRS in the provision of health services (Elbasha and Messonnier 2003; Jacobs and Baladi 1996). Whether economies of scale are likely to be exhibited to the same degree in other health services or other health settings is an empirical question. However, assuming CRS when costs and cost-effectiveness ratios in reality change with production, will produce biased estimates of any change in production and the bigger the expected change the larger the bias is likely to be. Not investigating or accounting for these economic forces could produce biased results that might mislead policy.

The presence of VRS has other implications. First it means that interventions cannot be treated as divisible for a population and retain the same average level of cost-effectiveness. In such a case, Johanesson (1996) suggests that the decision-maker's willingness to pay approach for choosing the allocation of health interventions would be more appropriate than maximising outcomes subject to a fixed budget. This is simply

[81] It is likely that an examination of the efficiency of diagnostic services would highlight inefficiencies also. No empirical studies of these services were identified in Chapter 4.

because the latter either provides a programme or not, and therefore there is no division of programmes.

Johanesson (1996) also suggests that interventions with IRS are presented as 'mutually exclusive' options for each level of production within a league table. However, in the event that IRS exist over all levels of production, the greatest level of production will demonstrate extended dominance over all other options, and therefore all production prior to the minimum efficient point would be excluded. Thus interventions with IRS are less likely to be provided, but conversely, assuming CRS when IRS exist means that interventions are likely to be over provided.

The recent review on learning effects with health technology (Ramsey et al. 2000) and its potential application to understanding economies of scale and CEA is also of interest. However, researching this issue will require larger sample sizes for the resources and costs of providing services than usually underpin CEAs in practice. Therefore it is important that randomised clinical trials provide costs from each trial centre (e.g. Coyle and Drummond 2001; Raikou et al. 2000; Wilke et al. 1998) and that the analysis of variation include analyses of technical and scale efficiency.

10.2.5 Efficiency and the generalisability of cost-effectiveness data

Ignoring the possible existence of technical inefficiencies and VRS would make generalisability of cost-effectiveness ratios suspect and could lead to a misallocation of resources. Consider, for example, a more efficient health system incorporating cost-effectiveness ratios of health interventions from a less efficient health system. The cost-effectiveness ratios will be higher than could be expected if the services were provided within their own system. If the transported ratio is used, the intervention would be less likely to be adopted and hence inefficiencies in one system are imported into another.

Failure to adequately incorporate some assessment of the relative efficiency of providers may therefore also bias the outcomes of CEA within health systems. For example, consider a new intervention provided by a highly motivated and efficient provider compared to standard care at a low-efficiency provider. As a minimum, therefore, good practice in economic evaluation should seek to compare interventions between providers with similar levels of technical efficiency, while sensitivity analysis should attempt to consider the impact of different levels of technical efficiency on results.

10.2.6 Constraints to efficiency improvement

Human resources policy has the potential to be an important support to, or major brake upon, efforts to improve efficiency, therefore the potential importance of attitudes to job losses cannot be over-stated. Where major inefficiencies have been identified, it is highly unlikely that equivalent savings can be realised without job losses. Chapters 6 and 8 illustrated that staff accounted for a considerable proportion of the costs of providing services in the case studies. Chapters 7 and 9 identified scope for substantial reductions in total costs, and by implication staff costs. However, as WHO (2000) notes, 'tensions' may arise between managers and politicians when the right to shed workers is withheld due to political pressure. It is particularly important that politicians understand that they will not be able to have both savings and no job losses, and that squeezing non-personnel funds is likely simply to exacerbate existing inefficiencies.

Therefore, employment contracts should make some provision for reassignment of duties or redeployment (functionally or geographically), even if fixed-term contracts are not felt to be feasible (the latter clearly offer the opportunity of non-renewal, greatly facilitating skill substitution). Careful assessment of skill requirements and skill-mix should be undertaken regularly, so that opportunities presented by routine departures of staff (promotion, job moves, retirement etc.) can be exploited to allow skill substitution. Institutional and professional inflexibilities can easily jeopardise attempts at skill substitution.

Remuneration policies and practices also seem likely to have a significant impact on the efficiency of service delivery in developing countries (Hensher 2001). Ensuring that the remuneration of skilled health workers is adequate seems frequently to be overlooked in the attempt to control costs and expenditure. However, there are several persuasive arguments as to why inadequate remuneration of skilled health workers will undermine efficiency. For example, the generic theory of efficiency wages (e.g. Stiglitz, 1987; Yellen, 1984), argues that productivity is directly affected by wage levels through attracting and retaining higher quality workers, and through motivating higher levels of effort and morale. However, this argument may not hold when, in effect, the government sets the market-clearing rate for health workers in Bangladesh.

Failure to attract and retain adequate quantities and quality of staff will lead to technical inefficiency because of skill shortages. For example, unfilled vacancies in key posts mean critical activities do not take place, and efficient operation becomes significantly degraded. Finally, persistently low pay almost always opens the door to unofficial 'private practice' using public facilities and time, if not to full-blown theft and corruption. The wastage of resources and low productivity that result may well outweigh the 'saving' in salaries achieved by a low wage policy.

The case studies in this thesis identified significant operating inefficiencies in the delivery of vaccination services and primary healthcare. In both instances, low utilisation and high fixed staffing costs appear to be contributing significantly to the levels of inefficiency. As a result, the question is whether policy-makers want to maintain current output, and thus release inputs for other uses, or whether they want to increase outputs until current inputs are efficiently employed. In other words, do they want a higher utilisation of primary care in the community served by the health centres, or would it be more cost-effective to take the efficiency savings and invest them into another programme, or another community? This thesis has not sought to answer these questions. However, if, after careful consideration, policy-makers in Bangladesh decided that it was indeed cost-effective and desirable to increase output of the existing providers and their services, then they would essentially be faced with a series of tasks related to improving the productivity of the current units. If, however, they required only current output levels, and what they really wanted were the efficiency savings, then they face two sets of tasks: how to improve productivity (of those who are going to keep their jobs) and how to identify the surplus inputs and extract them / convert them into savings.

If it is assumed that policy-makers have decided to put efficiency savings to some alternative use, they must then consider how far and how rapidly the current inputs can be converted to a new application, either physically (if they are suitable for redeployment), or via realisation into cash savings. This is clearly very much a problem of short-run versus long-run, and will largely depend on the extent to which institutional factors constrain the adjustment process. Returning to the case studies in this thesis, assume it has been identified that the highest priority use for the efficiency savings is to establish additional health centres in currently under-served regions. Therefore, at the health centres identified as inefficient, the policy-makers might have good reason to believe that some or all of the resources released by efficiency improvements could be directly transferred. For example, the staff probably already have the appropriate training and spare equipment could be moved. However, even in this relatively straightforward case, constraints may still be faced, e.g. will the redeployed staff be willing to work for a new employer (perhaps moving from the local government to the MOHFW), and how will their contracts be transferred? Do their employment contracts allow policy-makers to transfer staff involuntarily? What period of consultation with trade unions may be required before definite decisions could be taken? Do procedures exist to allow policy-makers to initiate a process of redundancies? Do they have an effective human resources policy to allow them to select those who will stay and those who will go? If not, can they plausibly retrain the staff – if not to go to the highest priority programme, then at least to do something deemed more valuable than their current role? This option reduces the scale of the efficiency saving that would ultimately be realised, but at least provides a solution that is less allocatively inefficient than the present situation. How long will all these processes take to work through?

10.2.7 The cost-effectiveness of efficiency improvement programmes

In light of the above discussion, when is it cost-effective to introduce an efficiency improvement programme? Efficiency improvement and implementation methods seek to change the behaviour of individuals or organisations in response to inefficiencies.

Behavioural change comes at a certain cost and achieves a certain level of change; it is never costless. The economics of efficiency improvement could provide a way of thinking about inefficiency and identify, for policy-makers and practitioners, the best use of scarce resources to achieve efficiency improvement goals. A model for working through the economic issues of efficiency would combine the costs and effects of corrected inefficiency with the costs and degree of behavioural, institutional or system change achieved by an efficiency improvement method in the policy maker's locality.

Because it is hardly ever possible to have one empirical study that gathers all the data needed to study the cost-effectiveness of an efficiency improvement strategy (and this thesis unfortunately has failed to do so), the models by Mason et al. (2001), Lobo et al. (2003), Severens (2003) and Verstappen et al. (2004) developed to examine when it is cost-effective to introduce a quality improvement programme are discussed below as they can provide guidance on the likely design of such a model.

In the economic evaluation of quality improvement interventions, costs have been subdivided into different phases of the quality improvement process (Verstappen et al. 2004). First, there are costs related to the task of collecting evidence to identify best practices, and conversely, poor or inefficient practice. Therefore, the costs of the research presented in Chapters 6 and 8 reflect these costs, which can be classified as developmental costs (fixed costs). Second, there are costs associated with organising an efficiency improvement programme, e.g. replace jobs-till-old-age-retirement with shorter term renewable contracts. Such costs are basically one-time costs and can therefore be considered fixed costs, unless the intervention used after the experience that is gained is subject to change. In that case, the efforts associated with a revision of the strategy must be considered execution costs. If the efficiency programme targets behaviour rather than legislation, the magnitude of behavioural change is unlikely to remain constant over time (Durieux et al. 2000), and a decision would need to be taken as to whether efficiency improvement is a 'one off' or whether periodic reimplementation should be costed. On the other hand, the costs of the actual execution of the efficiency improvement strategy are not relevant until the moment the strategy is executed (Lobo et al. 2003). Such costs can be considered fixed or variable, depending on the amount of detail included in the cost study.

Costs might sometimes be associated with a change in health care provision as a result of the application of an efficiency improvement strategy. Of course, this would depend on whether the analysis was measuring output-orientated efficiency, which addresses the question, "By how much can output quantities be expanded without changing the input quantities used?". The analyses presented in this thesis have measured input-orientated technical efficiency, which would result in a supply-side recommendation such as replacing jobs-till-old-age-retirement with fixed shorter term renewable contracts; no change in health care provision would be examined in this situation.

However, an output-orientated analysis would need to develop a demand-side intervention, which is more challenging (Ensor and Cooper 2004). And in this situation, as a result of the application of a social mobilisation intervention, for instance, vaccinators may see more children. Non-medical costs, such as parents' cost for time and travel, and possibly costs resulting from absence from work, can also be analysed on this level. These changes in health care provision costs are always considered variable costs.

Mason et al. (2001) distinguish between treatment cost-effectiveness (the incremental costs and benefits of a treatment) and policy cost-effectiveness (combining treatment cost-effectiveness with the cost and magnitude of change achieved by an improvement programme). Policy cost-effectiveness is most likely to remain attractive in those treatments that are highly cost-effective, e.g. vaccination services and primary health care more generally (World Bank 1993; Doherty and Govender 2004), and most likely to become unattractive when the cost-effectiveness of treatment is borderline.

As a general rule, the larger the efficiency deficit, the lower the marginal implementation cost of an efficiency programme. Therefore, an efficiency deficit must reach a minimum threshold before an efficiency improvement programme becomes economically attractive, that is, saves costs or shows an acceptable marginal cost-effectiveness ratio. Similarly, it may not be economically attractive to further improve the efficiency of care once an efficiency deficit is reduced to a certain size.

This chapter has discussed the findings of this thesis. The next, and final chapter, concludes the thesis. It considers the generalisability of the findings, and provides some policy recommendations for programme managers and decision-makers. It will also consider some research priorities for the future.

Chapter 11

CONCLUSIONS

This chapter is divided into five sections. The first section summarises the objectives, methods and findings of the thesis. The second section presents the thesis' conclusions. The third section considers the generalisability of the findings. The fourth section provides some policy recommendations for programme managers and decision-makers. And finally, the fifth section considers some future research priorities.

11.1 Summary of thesis

This thesis has contributed to the methodological development and application of cost, and more broadly cost-effectiveness, analysis of health care programmes by exploring whether, and to what degree, health care is delivered efficiently in one developing country, Bangladesh. It compared and contrasted two different efficiency measurement techniques, and applied them to the delivery of primary health care in urban and rural Bangladesh.

The specific objectives of the thesis were to:

1. Describe the empirical evidence on the efficiency of health care programmes in low- and middle-income countries and regions;
2. Estimate the cost of delivering vaccination services among a sample of vaccination delivery units in Dhaka City;
3. Estimate the cost of delivering primary health care among a sample of health centres in rural Bangladesh;
4. Estimate the efficiency of delivering these services using DEA and SFA;

5. Describe the variation in efficiency among the units and explore some of the causes of this variation;
6. On the basis of these findings, describe the potential implications of inefficiency in the delivery of health care programmes;
7. On the basis of these findings, make recommendations on how policy-makers in Bangladesh and elsewhere could improve efficiency, and make recommendations on further research relevant to health care efficiency issues.

This thesis has addressed the study objectives in the following ways. While there has been a recent expansion in the number of efficiency evaluations (Hollingsworth 2003), and despite a large and growing body of literature on the measurement of health facility costs in developing countries (Barnum and Kutzin 1993; Adam et al. 2003), a review of the literature revealed that there is a paucity of data on the efficiency of health care in the developing world (objective 1).

Standard costing methods were employed to estimate the cost of delivering vaccination services and primary health care in urban and rural Bangladesh respectively (objectives 2 and 3). In essence, standard costing methods assume full technical efficiency, but as cost data are generally summarised into a single estimate, they reflect the average level of efficiency exhibited among the sample of facilities costed. These analyses identified a large degree of variation in unit costs which could be indicative of varying degrees of technical efficiency. Therefore, parametric and non-parametric efficiency measurement techniques were applied to the same data (objectives 4 and 5). Using DEA and SFA a large degree of inefficiency among both the vaccination delivery units and primary health care centres was identified.

Objective 6 was addressed in the preceding chapter, and objective 7 is addressed in the rest of this concluding chapter. More specifically, the remainder of this chapter reflects on what has been presented in the preceding ten chapters and draws lessons from the theoretical and empirical information. It discusses the generalisability of the findings within and beyond Bangladesh. It makes recommendations on how policy-makers in Bangladesh and elsewhere could best approach the issue of inefficiency within the health sector. Areas for future research are then outlined.

11.2 Thesis conclusions

From this research the following can be concluded:

1. Based admittedly on limited evidence, health care systems in both developing and developed countries, display significant intra-system variations in technical efficiency.

2. There is scope for significant savings from reductions in relative inefficiency achieved by pulling poor performers up to benchmark performance levels (notwithstanding any scope to further improve the efficiency of 'frontier' production units).

3. Technical and scale inefficiency is present, to a large degree, in the delivery of health care in both urban and rural Bangladesh.

4. When technical inefficiency exists, as illustrated in the case studies, it means that a cost-effectiveness ratio does not reflect the minimum efficient point of production at a given level. A facility operating inefficiently incurs excess costs given the outputs produced. If these excess costs could be reallocated elsewhere, than there exists the

possibility of potential *Pareto* efficiency gains, which would by definition make programmes more cost-effective.

5. Health programmes are administered in settings that often violate the frequently stated assumption of constant returns to scale. Assuming constant returns to scale when average costs and cost-effectiveness ratios in reality change with production, will produce biased estimates of any change in production and the bigger the expected change the larger the bias.

6. Ignoring the possible existence of technical inefficiencies and variable returns to scale will make the generalisability of cost-effectiveness ratios suspect and could lead to a misallocation of resources.

11.3 Generalisability of findings

The importance of the findings depends on the extent to which they can be generalised.

1. The results of parametric and non-parametric efficiency measurement studies are sample specific[82]. The scores only reflect the dispersion of efficiencies within each sample and they say little about the efficiency of one sample relative to another. When efficiency scores for two different samples of health facilities are compared, as each sample is compared it is not possible to make conclusions on their relative efficiency as each sample is compared to the most efficient production unit in its own sample. A meaningful comparison would require samples to be combined, which may not be

[82] CEA results are equally site-specific although much current research is seeking ways in which to increase the generalisability or transferability of findings from one setting to another (e.g. Sculpher et al. 2004). Nevertheless, it is commonly implicitly and sometimes explicitly assumed (incorrectly) that results can be readily generalised among different settings.

possible among (or even within) countries where outputs and inputs are defined differently and costs would have to be converted to a common currency which again reduces comparability. As a minimum, therefore, good practice in economic evaluation should seek to compare interventions between providers with similar operational efficiency levels, while sensitivity analysis should attempt to consider the impact of different levels of technical and economic efficiency on results.

2. The relative nature of measuring efficiency requires that each country develop a strategy of its own and that, in turn, its own efficiency improvement programme. There is scope for sharing of experience and expertise, both in measurement and in implementing efficiency improvement measures, but it is essential to identify specific problems related to inefficiency from the top to the very lowest level of the system, and to develop solutions which will fit local realities and overcome particular local obstacles and constraints. Hensher (2001) proposed that a successful national-level approach to developing an efficiency improvement programme would contain the following components (which are not all sequential steps):

- identification and quantification of major areas of technical inefficiency and potential gains from efficiency improvement;
- assessment of priority employment of funds / resources released through efficiency improvements;
- identification of key causes of identified inefficiencies;
- assessment of possible interventions to improve efficiency;
- assessment of likely constraints acting upon efficiency improvement options, and estimation of likely scale of savings realisable;
- implement structural changes required to facilitate major or one-off improvements;

- implement organisational and cultural shifts to continuous productivity improvement, including appropriate performance management and incentive systems.

Each of these steps are briefly discussed in turn.

Identifying and quantifying major inefficiencies

The review of methods (Chapter 3) and studies (Chapter 4), in addition to the applications presented in this thesis (Chapters 6 – 9) provide a clear sense of the range of techniques available for deployment in the search for inefficiencies in health systems. Most critical, however, is the development, full implementation and subsequent maintenance of a basic data reporting system, which provides useful, meaningful information on activity, expenditure, productivity and efficiency. A basic level of confidence in their quality and comparability is required before they can be used to inform efficiency improvements programmes. As noted in Chapter 10, as part of health sector reforms in Bangladesh, the Unified Management Information System Unit of the Directorate General of Health Services of the MOHFW introduced a new record-keeping and reporting system. This system should provide the necessary data to apply the techniques for efficiency measurement presented in Chapter 3 (although facility-specific expenditure data may be lacking, and as discussed in Chapter 10, additional data on case-mix and the quality of care would be desirable). As Uddin et al. (2002) illustrated, there is scope to improve the quality of these routinely reported data. Nevertheless, without regularly available routine data, policy-makers are forced to rely on one-off sample data and special studies, such as those described in this thesis. These stop-gap approaches to data militate strongly against measurement of progress and improvement over time, and generally fail to cover all providers, which are both serious impediments to the process of efficiency improvement. It is probably preferable, as is the case now in Bangladesh, to obtain maximum coverage of even a very crude data system than it is to focus on obtaining more sophisticated data at pilot sites – because without the former, no analysis will be possible at any sites other than these pilots. However the parametric and non-parametric approaches used in this thesis are likely to require specialised technical and academic expertise in order to employ them, which may be lacking in many low- and middle-income countries.

Assessment of priorities for additional resources

Where sophisticated sectoral resource allocation processes are being developed (e.g. application of sectoral cost-effectiveness analysis) the assessment of priorities for additional resources is likely to be relatively straightforward, in the sense that analyses already undertaken can be used directly. In the absence of such work, some discussion will need to take place regarding the stated health priorities of the country, and their likely fit with the level and mix of resources which are likely to become available given the nature of the inefficiencies that have been identified. The core question here is to ask whether more of the same is desired (i.e. increased output for current inputs), or whether the desire is to release resources for other uses (current output for reduced inputs), in order to plan efficiency improvement measures accordingly.

Identification of causes of major inefficiencies

It is essential to understand why particular inefficiencies are arising if there is to be any realistic chance of reducing them. While an analysis of environmental variables can help shed light on the causes, it is likely that this will be a qualitative exercise. The people responsible for the inefficient services are likely to be the best source of insight into causes of inefficiency. Whether more formal qualitative research methods are used to elicit their views, or whether managers simply spend time to ask questions and listen to opinions, those who are caught up in the heart of inefficient practices must be questioned in detail about why things happen as they do and, critically, how things might be improved. A significant portion of technical inefficiency probably relates to extremely micro-level custom and practice, which general managers or researchers may not necessarily be able to identify as inefficient. Overdyk et al. (1998) provide a fascinating discussion of the extremely micro-level changes in scheduling, organisation and day-to-day operation which they undertook to achieve significant improvements in the efficiency of their operating rooms, involving a level of intervention that no centralised policy could effectively capture.

Assessment of possible interventions and likely savings

Identifying potential remedies to inefficiencies requires a two-track process. At one level is the grass-roots approach of involving workers and stimulating process improvements and initiatives by all those involved in the process of health care delivery. However, there is, of course, an extensive stock of experience and knowledge already available internationally, which can be drawn upon to provide rather more fundamental changes and innovations. The Effective Practice and Organisation of Care topic group in the Cochrane Collaboration undertakes systematic reviews of interventions to improve health professional practice and the organisation of health services. However the great majority of reviews are based largely on studies in high-income countries and there are few intervention studies in low- and middle-income countries of strategies to improve the coverage of effective interventions (Haines et al 2004).

It will also be important to appreciate that many savings will take a considerable time to realise, and may well require up-front investment, i.e. there needs to be an acceptance that significant savings are unlikely to be realised without some up-front investment ('spend to save'). Thus, for example, the shedding of excess staff will require funding for redundancy packages, retraining measures, etc.

Structural change and 'Big Push' efficiency improvements

Eliminating very pronounced inefficiencies may well require concerted, deliberately planned structural change. Substantial analytical and planning effort will be required, while significant additional funding will be required for implementation. Key areas requiring funding include redundancy payments for excess staff; capital costs of site closure and disposal (which can be significant); increased expenditure on professional management; improvement works to upgrade facilities which are remaining open; and, quite possibly, new infrastructure. The provision of such capital transformation funding would seem to be an ideal use of donor funding; a discrete, non-recurrent programme whose explicit aim is to unlock efficiency savings.

Shifting to continuous efficiency improvement

In general terms, developed countries have consistently improved productivity in health care over a long time span (Hensher et al. 1999). Yet on the basis of Zere et al. (2001) (reviewed in Chapter 4), it seems likely that many developing countries have faced either static or negative productivity and efficiency change over recent years. A number of factors have probably contributed to this lack of demonstrated efficiency gain. Foremost has been a general insufficiency of funds, leading to bottlenecks and inefficient input mixes. But Hensher (2001) argues that another key contributor has been the continuing failure to develop a strong cadre of non-medical, professional health service managers in most developing countries. He argues that:

> "The continued dominance of medically qualified administrators, often with very little or no management training, loath to take on their colleagues, and often still practicing clinical medicine for much of their working day, represents a lost opportunity to spark (or, if necessary, to bludgeon) change" (Hensher 2001).

Professional managers, armed with data with which to benchmark and compare performance, given basic authority to adjust resource use and production processes, themselves judged significantly upon their ability to improve efficiency, are required. This would represent a fundamental change in the commitment of health systems in developing countries to improving both technical and allocative efficiency (Hensher 2001).

3. A model to examine when it is cost-effective to introduce an efficiency improvement programme should be developed which would be generalisable in structure. In principle, such a model would enable policy-makers to work through the steps listed above.

11.4 Policy recommendations

This section addresses the penultimate objective of the study by making recommendations on how policy-makers in Bangladesh and elsewhere could improve efficiency.

There is widespread agreement that MOHFW service providers and upazila-level management staff have learned how to serve the 'system' better than they have learned how to serve their clients. In preparation for the Government's Health and Population Sector Program, 1998-2003, the Government's Task Force on Community and Stakeholder Participation carried out an assessment of the local perception about Government health and FP services in five villages in different regions of Bangladesh using a participatory rural appraisal methodology (Task Force 8, 1997a, b and c). The assessment showed that, according to the villagers, even though Government health services are officially free, poor people are commonly charged fees by the staff. Village practitioners, in contrast, charge fees which are transparent, well-known in the community, and affordable. Furthermore, the villagers who participated in these discussions with the Task Force maintained that Government service providers treat them with disrespect, and the providers give priority to the better-off clients. The villagers stated that the Government health care facilities are dirty and lack waiting rooms, toilet facilities and privacy. Finally, they complained that the providers (and particularly the doctors) were rarely there, the facilities were often closed, and that the facilities, more often than not, lacked drugs. The facilities were also frequently inconveniently located, often at some distance from the markets where they are accustomed to consulting private local practitioners (Task Force 8, 1997a, b and c). Not surprisingly, villagers often view Government health services only as a provider of last resort, when local village practitioners have failed in their attempt to resolve the problem and the family is becoming desperate.

Thorough, systemic changes will be required in the MOHFW which promote accountability to the community, improve productivity and performance of health staff, encourage decentralisation, improve quality of care, increase the responsiveness of the providers to the needs of clients, promote community and NGO involvement, and provide local monitoring based on accurate information. Although the need for these changes is recognised, the capacity of the Government system to reform itself is a major issue (Perry 1999).

Below are a series of comments on how health care needs to be improved in the future years, if resources are going to be used more efficiently.

1. Systemic changes in the MOHFW should be seen as equal in importance to technical and financial support for improving service delivery at the local level. High-level political support along with strong managerial and technical support will be needed to carry out these proposed changes. Fostering competition between the Government health service system and the private sector might promote change within the Government system, as might the increasing practice of 'contracting out' basic Government services to private organisations, including NGOs. Such an approach is currently being implemented in the metropolitan areas of Bangladesh by the Ministry of Local Government, Rural Development and Cooperatives through a project for urban primary health care funded by the Asian Development Bank (Loevinsohn and Harding 2005). The extent to which this increases health service efficiency needs to be evaluated before such schemes are widely replicated.

2. Strengthening independent monitoring of health status and utilisation of services at the upazila level would make it possible for the MOHFW to more rationally direct its limited resources to those areas with the greatest need (Ensor et al. 2003a). Funds could be directed to those upazilas and urban zone with, for instance, the highest rates of morbidity and mortality.

3. Unions and upazilas should be identified which have efficient (and high-quality) Government health service delivery units which can become 'model' upazilas where new policies and procedures for improving efficiency, quality, promoting community involvement, and fostering sustainability can be implemented and closely monitored. Such upazilas would have the potential for becoming training sites where field staff could be trained by persons working with programmes currently engaged in the effective provision of the ESP. These same programme sites could offer strong potential for carrying out local operations research activities to strengthen the efficiency of service delivery.

4. The need for documentation and evaluation of local service delivery activities will continue, and future progress in reaching the Millennium Developed Goals by 2015 will depend in part on scaling-up activities that have been proven to be successful on a smaller scale and which are carefully monitored and adjusted during the scaling-up process. Thus, there will need to be strong financial support for these operations research activities.

5. Compared to many developing countries, Bangladesh has a dynamic and innovative health sector, and the country's experience with operations research concerning health services is one of the most extensive in the world (Perry 1999). There has been little effort so far, however, to review and synthesise the lessons learned from these experiences or to assess their implications for the further development of primary health care services at the local level. Of particular concern in a country like Bangladesh is ensuring that efficient quality primary health care services reach those most in need.

There is a need for operations research which is population-based and focussed on health and demographic outcomes as well as on the process of service delivery.

11.5 Agenda for future research

In addition to the operations research suggested above, the following are topics for future research.

1. Future research should consider ways to improve the models presented in this thesis through the possible inclusion of some alternative variables, such as data on other services provided by the vaccination delivery units or ESP-specific outputs, rather than the age-specific number of visits.

2. Improved and more comprehensive quality measures would be extremely useful as staff may very well argue that they are less efficient because they are providing better quality patient care. Quality variables relating to patient outcomes would be very useful to include in such analyses. Vaccination services may be a 'special case', given that quality does not vary much, so there is little in the way of trade-off between cost and quality. Unlike the provision of a service such as vaccination, which can be quite easily defined, and its production relatively well-planned and managed, treatment of patients as a product, for example, is more challenging. However, the cost implications of meeting minimum quality standards are unknown since the link between quality and outcomes is unclear. Therefore, efficiency data should be linked to the data on the management of patients, which in turn should be linked to available guidelines, such as those developed for the Integrated Management of Childhood Illness. This will help 'tease out' whether inefficiency is due to using too many resources to manage patients

with particular conditions, or whether facilities / providers identified as being among the best performers are in fact inappropriately treating patients.

3. Longitudinal data would be useful to highlight changes in efficiency and the productivity of units relative to peers and relative to their own performance and may help produce more robust efficiency estimates. It may be the case that over time, a health facility's activity rises, and hence its capacity utilisation and measured efficiency changes. A longer term examination of changes in capacity utilisation and efficiency could assess how progress is being made towards achieving potential efficiency improvement targets. Longitudinal data would help clear up several questions such as whether some outliers are merely one-off data anomalies, whether inefficient units are truly that, or have made improvements on prior performance, and more importantly whether efficiency scores change from year to year and display inconsistency.

4. More research is necessary in order to better understand the determinants of efficiency. Although advances have been made in productivity analysis in recent years, the effective use of productivity measures is dependent on the consideration of a host of factors that may influence organisational performance. There is a need to determine the relative impact of different strategies and policies on productivity and efficiency. In particular, the role of institutions and culture, as well as financial and organisational factors, in the incentive structure governing manager and provider behaviour, needs to be better understood if inter-country comparisons are to be interpreted correctly and if best practice is to be applied successfully across countries. Detailed investigative studies in a sample of relatively efficient health facilities to document key attributes of best practice should be performed.

5. A related question is when is it cost-effective to implement an efficiency improvement programme? As a starting point, an evidence-base on the costs and effects of strategies to improve efficiency needs to be collated. It will be important to recognise the context-specific nature of many strategies, but consideration should be given to whether and how a matrix can be developed to summarise certain scenarios.

In conclusion, if something is deemed worth doing then it should be carried out in a way which ensures the optimum use of scarce resources. An exclusive focus on switching resources from less cost-effective to more cost-effective activities will not realise the full benefits in terms of improved allocative efficiency if providers on the ground are not producing services at lowest cost. Furthermore, while more money is certainly needed to tackle poor countries' health problems such as Bangladesh, how it is spent is more important than how much is spent.

REFERENCES

Abbassi K (1999) The World Bank and world health; focus on South Asia, 1: Bangladesh. *British Medical Journal* 318: 1066-1069.

Adam T, Evans DB, Murray CJ (2003) Econometric estimation of country-specific hospital costs. *Cost-Effectiveness and Resource Allocation* 26; 1(1).

Afriat SN (1972) Efficiency estimation of production functions. *International Economic Review* 13: 568-598.

Aigner DJ, Chu SF (1968) On estimating the industry production function. *American Economic Review* 58: 826-839.

Aigner DJ, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6: 21-37.

Al-Shammari M (1999) A multi-criteria data envelopment analysis model for measuring the productive efficiency of hospitals. *International Journal of Operations and Production Management* 19(9): 879-890.

Altman DG (1991) *Practical Statistics for Medical Research*. London: Chapman and Hall.

Arifeen SE, Bryce J, Gouws E, Baqui AH, Black RE, Hoque DME, Chowdhury EK, Yunus M, Begum N, Akter T, Siddique A (2005) Quality of care for under-fives in first-level health facilities in one district of Bangladesh. *Bulletin of the World Health Organization* 83(4): 260-267.

Atkison SE, Wilson PW (1995) Comparing mean efficiency and productivity scores from a small sample: A bootstrap methodology. *Journal of Productivity Analysis* 6: 132-152.

Avkiran NK (2002) *Productivity Analysis in the Service Sector with Data Envelopment Analysis.* (2nd edition). Queensland, Camira.

Baltussen RM, Adma T, Tan Torres T, Hutubessy RC, Acharya A, Evans DB, Murray CJL (2002) *Generalized Cost-Effectiveness Analysis: A Guide.* Geneva: World Health Organization.

Bangladesh Bureau of Statistics (1997) *Statistical Pocketbook of Bangladesh, 1996.* Dhaka: Bangladesh Bureau of Statistics.

Bangladesh Bureau of Statistics (2003) *Statistical Pocketbook of Bangladesh, 2000.* Dhaka: Bangladesh Bureau of Statistics.

Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science* 30(9): 1078-1092.

Banker RD, Morey RC (1986a) Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research* 34: 513-521.

Banker RD, Morey RC (1986b) The use of categorical variables in data envelopment analysis. *Management Science* 32: 1613-1627.

Barkat A, Howlader SR, Chakma PB (1999) *Activity and Cost Analysis of Essential Health and Family Planning Services in Bangladesh.* Dhaka: Health Economics Project, MOHFW.

Barnum H, Kutzin J. (1993) *Public Hospitals in Developing Countries.* Baltimore, John Hopkins Press.

Battese GE, Corra GS (1977) Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia. *Australian Journal of Agricultural Economics* 21: 169-179.

Bhat R, Verma BB, Reuben E (2001) Hospital efficiency: an empirical analysis of district hospitals and grant-in-aid hospitals in Gujarat. *Journal of Health Management* 3(2): 167-198.

Birch S, Gafni A (2002) On being NICE in the UK: guidelines for technology appraisal for the NHS in England and Wales *Health Economics* 11:185-191

Boussofiane A, Dyson RG, Thanassoulis E (1991) Applied data envelopment analysis *European Journal of Operational Research* 52: 1-15.

Briggs A, Clark T, Wolstenholme J, Clarke P (2003) Missing ... presumed at random: cost-analysis of incomplete data. *Health Economics* 12: 377-392

Brenzel L, Claquin P (1994) Immunization programs and their costs. *Social Science and Medicine* 39(4): 527-536.

Buse K, Gwin C (1998) The World Bank and global cooperation in health: the case of Bangladesh. *Lancet* 351: 665-669.

Cassells A, Janovsky K (1998) Better health in developing countries: are sector-wide approaches the way of the future? *Lancet* 352: 1777-1779.

Chakrabarti A, Rao DN (in press). Efficiency in production of health: a stochastic frontier analysis for Indian states. In Colombus F (Ed) *Focus in Global Economics.* New York: Nova Science Publications

Charnes A, Cooper WW, Lewin AY, Seiford LM (1995) *Data Envelopment Analysis: Theory, Methodology and Application.* London: Kluwer Academic Publishers.

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research* 2: 429-444.

Chaudhury N, Hammer JS (2004) Ghost doctors: absenteeism in rural Bangladeshi health facilities. *World Bank Economic Review* 18(3): 423-441.

Chen L, Huq E, D'Souza S (1981) Sex bias in the family allocation of food and health care in rural Bangladesh. *Population and Development Review* 7: 55-70.

Coelli TJ (1995) Estimators and hypothesis tests for a stochastic frontier function: a Monte Carlo analysis. *Journal of Productivity Analysis* 6: 247-268.

Coelli TJ (1996a) *A Guide to DEAP Version 2.1: A Data Envelopment Analysis Computer Program*, CEPA Working Paper No. 8/96, Department of Econometrics, University of New England.

Coelli TJ (1996b) *A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation.* CEPA Working Paper No. 7/96, Department of Econometrics, University of New England.

Coelli T, Rao DSP, Battese GE (1998) *An Introduction to Efficiency and Productivity Analysis*. London: Kluwer Academic Publishers.

Conrad J (1902) *Heart of Darkness*. London: Blackwood's.

Cooper WW, Seiford LM, Tone K (2003) *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software.* Boston: Kluwer Academic Publishers.

Coyle D, Drummond M (2001) Analyzing differences in the costs of treatment across centres within economic evaluations. *International Journal of Technology Assessment in Health Care* 17(2): 155-163

Creese A, Parker D (1994) *Cost Analysis in Primary Health Care: A Training Manual for Programme Managers*. Geneva: World Health Organisation.

Culyer A (1992) The morality of efficiency in health care: some uncomfortable implications. *Health Economics* 1(1): 7-18.

Debreu G (1951) The coefficient of resource utilisation. *Econometrica* 19: 273-292

Dervaux B, Leleu H, Valdmanis V, Walker D (2003) Parameters of control when facing stochastic demand: a DEA approach applied to Bangladeshi vaccination sites. *International Journal of Health Care Finance and Economics* 3(4): 287-299.

Doherty J, Govender R (2004) The cost-effectiveness of primary care services in developing countries: a review of the international literature. Disease Control Priorities Project, Working Paper No. 37.

Donaldson C (1990) The state of the art of costing health care for economic evaluation. *Community Health Studies* 9: 341-356.

Donaldson C, Currie G, Mitton C (2002) Cost effectiveness analysis in health care: contraindications. *British Medical Journal* 325: 891-4

Drake L, Howcroft B (1994) Relative efficiency in the branch network of a UK bank: an empirical study. *OMEGA* 22(1): 83-90.

Drummond MF, O'Brien B, Stoddart GL, Torrance GW (1997) *Methods for the Economic Evaluation of Health Care Programmes* (2nd edition) Oxford: Oxford University Press.

Durieux P, Nizard R, Ravaud P, Mounier N, Lepage E (2000) A clinical decision support system for prevention of venous thromboembolism: effect on physician behavior. *Journal of the American Medical Association* 283(21): 2816-2821.

Dyson RG, Allen R, Camanho AS, Podinovski VV, Sarrico CS, Shale EA (2001) Pitfalls and protocols in DEA. *European Journal of Operational Research* 132: 245-259.

Elbasha EH, Messonnier ML (2004) Cost-effectiveness analysis and health care resource allocation: decision rules under variable returns to scale. *Health Economics* 13(1): 21-35.

Elixhauser A, Halpern M, Schmier J, Luce BR (1998) Health care CBA and CEA from 1991 to 1996: an updated bibliography. *Medical Care* 36(5 Suppl): Ms1-9, Ms18-147.

Elixhauser A, Luce BR, Taylor WR, Reblando J (1993) Health care CBA/CEA: an update on the growth and composition of the literature. *Medical Care* 31(7 Suppl): Js1-11, Js18-149.

Ensor T, Cooper S (2004) Overcoming barriers to health service access: influencing the demand side. *Health Policy and Planning* 19(2): 69-79.

Ensor T, Dave-Sen P, Ali A, Hossain A, Begum SA, Moral H (2002) Do essential service packages benefit the poor? Preliminary evidence from Bangladesh. *Health Policy and Planning* 17(3): 247-256

Ensor T, Hossain A, Dave-Sen P, Ali L, Begum SA, Moral H (2003a) Geographic resource allocation in Bangladesh. In Yazbeck AS, Peters DH (Eds) *Health Policy Research in South Asia: Building Capacity for Reform* Washington DC: World Bank.

Ensor T, Hossain A, Dave-Sen P, Ali L, Begum SA, Moral H (2003b) Public expenditure review of the health and population sector program in Bangladesh. In Yazbeck AS, Peters DH (Eds) *Health Policy Research in South Asia: Building Capacity for Reform* Washington DC: World Bank.

Ersoy K, Kavuncubasi S, Ozcan YA, Harris JM (1997) Technical efficiencies of Turkish hospitals: DEA approach. *Journal of Medical Systems* 21(2): 67-74.

Evans DB, Tandon A, Murray CJ, Lauer JA (2001) Comparative efficiency of national health systems: cross national econometric analysis. *British Medical Journal* 323(7308): 307-310.

Everitt BS (1995) *The Cambridge Dictionary of Statistics in the Medical Sciences*. Cambridge: Cambridge University Press.

Expanded Programme on Immunization (1998) *EPI Programme Review, 1998*. Dhaka: Directorate General of Health Services, Ministry of Health and Family Welfare, People's Republic of Bangladesh.

Expanded Programme on Immunization (2000) *Coverage Evaluation Survey, 1999.* Dhaka: Directorate General of Health Services, Ministry of Health and Family Welfare, People's Republic of Bangladesh.

Expanded Programme on Immunization (2001) *Coverage Evaluation Survey, 2000.* Dhaka: Directorate General of Health Services, Ministry of Health and Family Welfare, People's Republic of Bangladesh.

Färe R, Grosskopf S, Lovell CAK (1994) *Production Frontiers.* Cambridge: Cambridge University Press.

Farrell MJ (1957) The measurement of productive efficiency. *Journal of the Royal Statistical Society* Series A, CXX, Part 3: 253-290.

Ferrier GD, Hirschberg JG (1997) Bootstrapping confidence intervals for linear programming efficiency scores. *Journal of Productivity Analysis* 8: 19-33.

Fried HO, Lovell AK, Schmidt SS (1993) *The Measurement of Productive Efficiency: Techniques and Applications.* New York: Oxford University Press.

Fried HO, Lovell CAK, Schmidt SS, Yaisawarng S (2002) Accounting for environmental effects and statistical noise in data envelopment analysis. *Journal of Productivity Analysis*, 17(1/2): 157-174.

Global Alliance for Vaccines and Immunization (2005) The Data Quality Audit. Available at: www.vaccinealliance.org/Support_to_Country/Monitoring_Evaluatio/dqa.php/ Accessed on December 2005

Gravelle H, Jacobs R, Jones AM, Street A (2003) Comparing the efficiency of national health systems: a sensitivity analysis of the WHO approach. *Applied Health Economics and Health Policy* 2(3): 141-7.

Hadley J, Zuckerman S (1994) The role of efficiency measurement in hospital rate setting *Journal of Health Economics* 13: 335-340.

Haines A, Kuruvilla S, Borchert M (2004) Bridging the implementation gap between knowledge and action for health. *Bulletin of the World Health Organization* 82: 724-733.

Hensher M (2001) *Financing Health Systems Through Efficiency Gains*. Commission for Macroeconomics and Health, Working Group 3, Working Paper Series Paper No.2.

Hensher M, Edwards N, Stokes R (1999) International trends in the provision and utilisation of hospital care. *British Medical Journal* 319: 845-848.

Hollingsworth B (2003) Non-parametric and parametric applications measuring efficiency in health care. *Health Care Management Science* 6(4): 203-218.

Hollingsworth B, Dawson P, Maniadakis P (1999) Efficiency measurement of health care: a review of non-parametric methods and applications. *Health Care Management Science* 2(3): 161-172.

Hollingsworth B, Wildman J (2003) The efficiency of health production: re-estimating the WHO panel data using parametric and nonparametric approaches to provide additional information. *Health Economics* 12(6): 493-504.

Hossain B, Begum K (1998) Survey of the existing health workforce of Ministry of Health, Bangladesh. *Human Resources for Health Development Journal* 29(2): 109-116.

Hossain MM, Glass R (1988) Parental son preference in seeking medical care for children less than five years of age in a rural community in Bangladesh. *American Journal of Public Health* 78: 1349-1350.

Huq M (1991) *Near Miracle in Bangladesh*. Dhaka: University Press Limited.

Jacobs P, Baladi JF (1996) Biases in cost measurement for economic evaluation studies in health care. *Health Economics* 5(6): 525-529.

Jacobs R (2001) Alternative methods to examine hospital efficiency: data envelopment analysis and stochastic frontier analysis. *Health Care Management Science* 4: 103-115.

Jamison DT, Alleyne G, Breman J, Claeson M, Evans DB, Jha P, Measham AR, Mills A, Musgrove PR (2006) *Disease Control Priorities in Developing Countries* (2nd edition) Oxford: Oxford University Press.

Jamison DT, Mosely WH, Measham AR, Bobadilla JL (1993) *Disease Control Priorities in Developing Countries*. Oxford: Oxford University Press.

Jacques, NK, Koch S (2005) Improving policy implementation by the use of efficiency models: an application of DEA on public hospitals. Paper presented at the African Economic Society's 10th Annual Conference on Econometric Modelling in Africa.

Johanesson M (1996) *Theory and Methods of Economic Evaluation of Health Care*. Dordrecht: Kluwer Academic Publishers.

Kathuria V, Sankar D (2005) Inter-state disparities in health outcomes in rural India: an analysis using stochastic production frontier approach. *Development Policy Review* 23(2): 145-163.

Khan MM, Yoder RA (1998) *Expanded Program on Immunization in Bangladesh: Cost, Cost-Effectiveness, and Financing Estimates*. Special Initiatives Report No. 6. Bethesda, MD: Partnerships for Health Reform Project, Abt Associates Inc.

Khan MM, Khan SH, Walker D, Fox-Rushby J, Cutts F, Akramuzzaman SM (2004) Cost of delivering child immunization services in urban Bangladesh: a study based on facility-level surveys. *Journal of Health, Population and Nutrition* 22(4): 404-412

Khatun T (2001) District-based measurement of human poverty in Bangladesh. *Bangladesh Development Studies* 27(3): 91-109.

Killingsworth JR, Hossain N, Hendrick-Wong Y, Thomas SD, Rahman A, Begum T (1999) Unofficial fees in Bangladesh: price, equity and institutional issues. *Health Policy and Planning* 14: 152-163.

Kirgia JM, Lambo E, Sambo LG (2000) Are public hospitals in Kwazulu-Natal Province of South Africa technically efficient? *African Journal of Health Sciences* 26(1): 21-27.

Kirigia JM, Emrouznejad A, Sambo LG (2002) Measurement of technical efficiency of public hospitals in Kenya: using data envelopment analysis. *Journal of Medical Systems* 26(1): 39-45.

Kirigia JM, Emrouznejad A, Sambo LG, Munguti N, Liambila W (2004) Using data envelopment analysis to measure the technical efficiency of public health centers in Kenya. *Journal of Medical Systems* 28(2): 155-166.

Kirigia JM, Sambo LG, Scheel H (2001) Technical efficiency of public clinics in Kwazulu-Natal Province of South Africa. *East African Medical Journal* 78(3 Suppl): S1-13.

Koopmans TC (1951) An analysis of production as an efficient combination of activities. In Koopmans TC (Ed) *Activity Analysis of Production and Allocation.* Cowles Commission for Research in Economics, Monograph No. 13, New York: Wiley.

Kumbhakar SC, Lovell CAK (2000) *Stochastic Frontier Analysis.* Cambridge: Cambridge University Press.

Leibenstein H (1966) Allocative efficiency vs "X-efficiency". *American Economic Review* 56: 392-415.

Levin A, Howlader S, Ram S, Siddiqui SM, Razul I, Routh S (1999) *Case Study on the Costs and Financing of Immunization Services in Bangladesh.* Special Initiatives Report No. 21. Bethesda, MD: Partnerships for Health Reform Project, Abt Associates Inc.

Lobo CM, Euser L, Kamp J, Frijling BD, Severens JL, Hulscher ME, Grol RP, Prins A, van der Wouden JC (2003) Process evaluation of a multifaceted intervention to improve cardiovascular disease prevention in general practice. *European Journal of General Practice* 9(3):77-83

Loevinsohn B, Harding A (2005) Buying results? Contracting for health service delivery in developing countries. *Lancet* 366: 676-681.

Lovell CAK (1993) Production frontiers and productive efficiency. In Fried HO, Lovell AK, Schmidt SS (Eds) *The Measurement of Productive Efficiency: Techniques and Applications*. New York: Oxford University Press.

Lovell CAK, Walters LC, Wood LL (1993) Stratified models of education production using DEA and regression analysis. In: *Data Envelopment Analysis: Theory, Methods and Applications* Charnes A, Cooper WW, Lewin YA, Seiford LM (Eds) New York: Quorum Books.

Masiye F, Ndulo M, Roos P, Odegaard K (2002) A comparative analysis of hospitals in Zambia: a pilot study on efficiency measurement and monitoring. In: Seshamani V, Mwikisa CN, Odegaard K (Eds) *Zambia's Health Reforms Selected Papers 1995-2000.* Lund: Sweden.

Manca A, Palmer S (2005) Handling missing data in patient-level cost-effectiveness analysis alongside randomised clinical trials. *Applied Health Economics and Health Policy* 4(2): 65-75

Mason J, Freemantle N, Nazareth I, Eccles M, Haines A, Drummond M (2001) When is it cost-effective to change the behaviour of health professionals? *Journal of the American Medical Association* 286(23): 2988-2992.

Mazumder MA, Bhuiyan MA, Tunon C, Baqui AH, Chowdhury AI, Khan SA, Islam R (1997) *An Inventory of Health and Family Planning Facilities in Dhaka City*. Dhaka: MCH-FP Extension Project (Urban), Health and Population Extension Division, ICDDR,B.

McCarty T, Yaisawarng S (1993) Technical Efficiency in New Jersey School Districts." In Fried HO, Lovell AK, Schmidt SS (Eds) *The Measurement of Productive Efficiency: Techniques and Applications*. New York: Oxford University Press.

Meeusen W, van der Broeck J (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review* 18: 435-444.

Ministry of Health and Family Welfare (2003) *Bangladesh National Formulary* (2nd edition) Dhaka: Directorate of Drug Administration.

Mitra SA, Al-Sabir A, Cross A, Jamil K (2001) *Bangladesh Demographic and Health Survey 2000-2001*. Dhaka: National Institute of Population Research and Training; Dhaka, Bangladesh: Mitra and Associates; Calverton, Maryland: Macro International Inc.

Musgrove P (2000) A critical review of 'A Critical Review.' *Health Policy and Planning* 15(1): 110-115.

National Institute for Clinical Excellence (2001) *Technical Guidance for Manufacturers and Sponsors on Making a Submission to a Technology Appraisal*. London: National Institute for Clinical Excellence.

Newhouse JP (1994) Frontier estimation: How useful a tool for health economics? *Journal of Health Economics* 13: 317-322.

Nunamaker TR (1983) Measuring routine nursing service efficiency: a comparison of cost per patient day and data envelopment analysis models. *Health Services Research* 18(2 Pt 1):183-208.

Nunamaker TR (1985) Using data envelopment analysis to measure the efficiency of non-profit organizations: A critical evaluation *Managerial and Decision Economics* 6(1): 50-58.

Osei D, D'Almeida S, George MO, Kirigia JM, Mensah AO, Kainyu LH (2005) Technical efficiency of public district hospitals and health centres in Ghana: a pilot study. *Cost-Effectiveness and Resource Allocation* 3(9).

Overdyk FJ, Harvey SC, Fishman RL, Shippey F (1998) Successful strategies for improving operating room efficiency at academic institutions. *Anaesthesia and Analgesia* 86:896-906.

Owino W, Korir J (1997) Public health sector efficiency levels: estimation and policy implications. Institute of Policy Analysis and Research Discussion paper 7.

Pastor JT (1995) *How to Account for Environmental Effects in DEA: An Application to Bank Branches.* Working Paper, Departamento de Estadistica e Investigacion Operativa. 03071 Alicante, SPAIN: Universidad de Alicante.

Pavananunt K (2004) Hospital efficiency: a study of public community hospital in Thailand. In: Emrouznejad A, Podinovski V (Eds.) *Data Envelopment Analysis and Performance Management.* Coventry: Warwick Print.

Penaloza Ramos MC (2003) Evaluación de la Eficiencia en instituciones hospitalarias públicas y privadas con Data Envelopment Analysis (DEA). República de Colombia. Departamento Nacional de Planeación, Dirección de Estudios Económicos. Archivos de Economia. Documento 244.

Perry HB (1999) *Health for All in Bangladesh: Lessons in Primary Health Care for the Twenty-First Century.* Dhaka: University Press, Ltd.

Pinzon Martinez MJ (2003) Medición de eficiencia técnica relativa en hospitales públicos de baja complejidad mediante la metodología Data Envelopment Analysis (DEA). República de Colombia. Departamento Nacional de Planeación, Dirección de Estudios Económicos. Archivos de Economia. Documento 245

Pitt MM, Lee L-F (1981) Measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics* 9: 43-64.

Rahman M, DaVanzo J (1993) Gender preference and birth spacing in Matlab, Bangladesh. *Demography* 30: 315-332.

Raikou M, Briggs A, Gray A, McGuire A (2000) Centre-specific or average unit costs in multi-centre studies? Some theory and simulation *Health Economics* 9(3): 191-198.

Ramanathan TV, Chandra KS, Thupeng WM (2003) A comparison of the technical efficiencies of health districts and hospitals in Botswana. *Development Southern Africa* 20(2): 307-320.

Ramsey CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT (2000) Assessment of the learning curve in health technologies. A systematic review *International Journal of Technology Assessment in Health Care* 16(4) 1095-1108.

Rannan-Eliya RP, Somanathan A (2003) The Bangladesh health facility efficiency survey. In Yazbeck AS, Peters DH (Eds) Health Policy Research in South Asia: Building Capacity for Reform Washington DC: World Bank.

Reinhardt U (2003) Forward to Rice T. *The Economics of Health Reconsidered* Chicago: Health Administration Press.

Richmond J (1974) Estimating the efficiency of production. *International Economic Review* 32: 715-723.

Sahin I, Ozcan YA (2000) Public sector hospital efficiency for provincial markets in Turkey. *Journal of Medical Systems* 24(6): 307-320.

Sarkis J, Talluri S (2002) Efficiency measurement of hospitals: issues and extensions. *International Journal of Operations and Production Management* 22(3): 306-313.

Schmidt P (1976) On the statistical estimation of parametric frontier production functions. *Review of Economics and Statistics* 58: 238-239.

Sculpher MJ, Pang FS, Manca A, Drummond MF, Golder S, Urdahl H, Davies LM, Eastwood A (2004) Generalisability in economic evaluation studies in healthcare: a review and case studies. *Health Technology Assessment* 8(49).

Severens JL (2003) Value for money of changing healthcare services? Economic evaluation of quality improvement *Quality and Safety in Health Care* 12:366-371.

Sherman HD (1984) Hospital efficiency measurement and evaluation: empirical test of a new technique. *Medical Care* 22(10): 922-938.

Skinner J (1994) What do stochastic frontier cost functions tell us about inefficiency? *Journal of Health Economics* 13: 323-328.

Somanathan A, Hanson K, Dorabawila T, Perera B (2000) *Operating Efficiency in Public Sector Health Facilities in Sri Lanka: Measurement and Institutional Determinants of Performance.* Small Applied Research Paper No. 12. Bethesda, MD: Partnerships for Health Reform Project, Abt Associates Inc.

Soucat A, Levy-Bruhl D, De Bethune X, Gbedonou P, Lamarque JP, Bangoura O, Camara O, Gandaho T, Ortiz C, Kaddar M, Knippenberg R (1997) Affordability, cost-effectiveness and efficiency of primary health care: the Bamako Initiative experience in Benin and Guinea. *International Journal of Health Planning and Management* 12 (Suppl 1): S81-108.

Stevenson RE (1980) Likelihood functions for generalised stochastic frontier estimation. *Journal of Econometrics* 61: 57-66.

Stiglitz JE (1987) The causes and consequences of the dependence of quality on price. *Journal of Economic Literature* 25: 1-48.

Stone PW, Schackman B., Neukermans CP, Olchanski N, Greenberg D, Rosen A, Neumann PJ (2005) A synthesis of cost-utility analysis literature in infectious disease. *Lancet Infectious Diseases* 383-391.

Street A (2003) How much confidence should we place in efficiency estimates? *Health Economics* 12: 895-907.

Tan-Torres Edejer T, Baltussen R, Adam T Hutubessy R, Acharya A, Evans DB, Murray CJL (2003) *Making Choices in Health: WHO Guide to Cost-Effectiveness Analysis*. Geneva: World Health Organization.

Task Force 8 (Community and Stakeholder Participation) (1997a) Report on Client and Fieldworker Consultations Using Participatory Rural Appraisal Method: Status of Existing Services. Dhaka: MOHFW.

Task Force 8 (Community and Stakeholder Participation) (1997b) Report on Client and Fieldworker Consultations Using Participatory Rural Appraisal Method: Utilization of Government Health Care Services. Dhaka: MOHFW.

Task Force 8 (Community and Stakeholder Participation) (1997c) Report on Client and Fieldworker Consultations Using Participatory Rural Appraisal Method: Why Poor Villagers Rarely Utilize Government Health Care Facilities. Dhaka: MOHFW.

Timmer CP (1971) Using a probabilistic frontier function to measure technical efficiency. *Journal of Political Economy* 79: 776-794.

Uddin MJ, Kabir H, Mahbub-ul-Alam, Uddin AHN, Ashraf A (2002) Assessment of the record-keeping and reporting system of the Bangladesh health and population sector programme at the union level. ICDDR,B Working Paper No. 155

UNICEF (1999) *Progotir Pathay*. Dhaka: UNICEF.

Valdmanis V (1992) Sensitivity analysis for DEA models: an empirical example using public vs. NFP hospitals. *Journal of Public Economics* 48: 185-205.

Valdmanis V, Kumanarayake L, Lertiendumrong J (2004) Capacity in Thai public hospitals and the production of care for poor and non-poor patients. *Health Services Research* 39(6 Pt 2): 2117-2134.

Valdmanis V, Walker D, Fox-Rushby JA (2003) Are vaccination sites in Bangladesh scale efficient? *International Journal of Technology Assessment in Health Care* 19(4): 692-697.

Varian HR (1992) *Microeconomic Analysis* (3rd edition) New York: Norton.

Vaughan JP, Karim E, Buse K (2000) Health care systems in transition III. Bangladesh, Part I. An overview of the health care system in Bangladesh. *Journal of Public Health Medicine* 22(1): 5-9.

Verstappen WH, van Merode F, Grimshaw J, Dubois WI, Grol RP, van der Weijden T. (2004) Comparing cost effects of two quality strategies to improve test ordering in primary care: a randomized trial. *International Journal of Quality Health Care* 16(5): 391-398.

Walker D, Fox-Rushby J (2000) Economic evaluation of communicable disease interventions in developing countries: a critical review of the published literature. *Health Economics* 9(8): 681-698.

Walker D, Khan S, Akramuzzaman SM, Khan M, Fox-Rushby J, Cutts F (2000) Cost-effectiveness analysis of measles control in Dhaka, Bangladesh. Report prepared for WHO HQ/98/454419 011638.

Walker D, Mosqueira NR, Penny ME, Lanata CF, Clark A, Sanderson C, Fox-Rushby JA (2004) Variation in the costs of delivering routine immunization services in Peru. *Bulletin of the World Health Organization* 82(9): 676-682.

Warner KE, Hutton RC (1980) Cost-benefit and cost-effectiveness analysis in health care. Growth and composition of the literature. *Medical Care* 18(11):1069-1084.

Wilke RJ, Glick HA, Polsky D, et al. (1998) Estimating country-specific cost-effectiveness from multinational clinical trials *Health Economics* 7(6): 481-494.

Wilson PW (1995) Detecting influential observations in deterministic nonparametric frontier models with multiple outputs. *Journal of Business and Economic Statistics* 11: 319-323.

World Bank (1993) *World Development Report 1993: Investing in Health.* Oxford: Oxford University Press.

World Bank (2005a) Country classification. Available at http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20420458~menuPK:64133156~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html Accessed November 2005.

World Bank (2005b) Bangladesh at a glance. Available at http://devdata.worldbank.org/AAG/bgd_aag.pdf Accessed November 2005

World Health Organization (2000) *World Health Report 2000. Health Systems: Improving Performance.* Geneva: WHO.

World Health Organization (2005) Immunization Financing Database. Available at: www.who.int/vaccines-surveillance/Vaccine_Financing/ or www.who.int/immunization_financing/en/ Accessed on November 2005.

Worthington AC (2004) Frontier efficiency measurement in health care: a review of empirical techniques and selected applications. *Medical Care Research and Review* 61(2): 135-170.

Yellen JL (1984) Efficiency wage models of unemployment. *American Economic Review* 74: 200-205

Zere E, McIntyre D, Addison T (2001) Technical efficiency and productivity of public sector hospitals in three South African provinces. *South African Journal of Economics* 69(2): 336-358.

Appendix 1: Khan et al. (2004)

J HEALTH POPUL NUTR 2004 Dec;22(4):404-412 © 2004 ICDDR,B: Centre for Health and Population Research
ISSN 1606-0997 $ 5.00+0.20

Cost of Delivering Child Immunization Services in Urban Bangladesh: A Study Based on Facility-level Surveys

M. Mahmud Khan[1,2], Suhaila H. Khan[2], Damian Walker[3], Julia Fox-Rushby[3], Felicity Cutts[4], and S.M. Akramuzzaman[5]

[1,2]Tulane University School of Public Health and Tropical Medicine, 1440 Canal Street, New Orleans, LA 70112, USA, [2]Public Health Sciences Division, ICDDR,B: Centre for Health and Population Research, [3]Health Policy Unit and [4]Infectious Disease Epidemiology Unit, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, UK, and [5]Clinical Sciences Division, ICDDR,B: Centre for Health and Population Research, GPO Box 128, Dhaka 1000, Bangladesh

ABSTRACT

This facility-based study estimated the costs of providing child immunization services in Dhaka, Bangladesh, from the perspective of healthcare providers. About a quarter of all immunization (EPI) delivery sites in Dhaka city were surveyed during 1999. The EPI services in urban Dhaka are delivered through a partnership of the Government of Bangladesh (GoB) and non-governmental organizations (NGOs). About 77% of the EPI delivery sites in Dhaka were under the management of NGOs, and 62% of all vaccinations were provided through these sites. The outreach facilities (both GoB and NGO) provided immunization services at a much lower cost than the permanent static facilities. The average cost per measles-vaccinated child (MVC), an indirect measure of number of children fully immunized (FIC = the number of children immunized by first year of life), was US$ 11.61. If all the immunization doses delivered by the facilities were administered to children who were supposed to be immunized (FVC), the cost per child would have been US$ 6.91. The wide gap between the cost per MVC and the cost per FVC implies that the cost of immunizing children can be reduced significantly through better targeting of children. The incremental cost of adding new services or interventions with current EPI was quite low, not significantly higher than the actual cost of new vaccines or drugs to be added. NGOs in Dhaka mobilized about US$ 15,000 from the local community to support the immunization activities. Involving local community with EPI activities not only will improve the sustainability of the programme but will also increase the immunization coverage.

Key words: Immunization; Costs and cost analysis; Health facilities; Non-governmental organizations; Community participation; Bangladesh

INTRODUCTION

The Expanded Programme on Immunization (EPI) aims to reduce morbidity and mortality from six vaccine-preventable diseases: tuberculosis, diphtheria, pertussis, tetanus, measles, and poliomyelitis. A fully-immunized child (FIC) receives six standard EPI antigens through eight vaccinations given in the first year of life. The recommended schedule is: one shot of Bacille Calmette Guerin (BCG) at birth, three doses of oral polio vaccine (OPV) together with three shots of diphtheria-pertussis-tetanus (DPT) at age 6, 10, and 14 weeks, and one shot of measles vaccine at age 9 months. Along with these six antigens, the routine EPI also included

Correspondence and reprint requests should be addressed to:
Dr. M. Mahmud Khan
Tulane University School of Public Health and Tropical Medicine
1440 Canal Street, #1900
New Orleans, LA 70112, USA
Email: khan@tulane.edu

two doses of tetanus toxoid (TT) for pregnant women and one dose of vitamin A for children at the time of the study. The main EPI programme (the routine EPI) is supplemented by other interventions, such as National Immunization Day (NID), mop-up after NID, acute flaccid paralysis (AFP) surveillance, and maternal and neonatal tetanus (MNT) surveillance.

EPI has reduced morbidity and mortality from vaccine-preventable diseases in Bangladesh, but little is known about costs and effectiveness of urban EPI. A comprehensive review in 1998 and two studies on the cost-effectiveness of the Bangladesh EPI have pointed out the need for collecting cost information from urban areas (1-4). Unlike rural Bangladesh, urban EPI is delivered through a partnership between the public sector and the private sector. In fact, the private service providers, especially NGOs, play such an important role in urban EPI that estimates based on national-level expenditure or cost data will be a significant underestimate of total costs if the contribution of NGOs is not included. However, the exact level of involvement of NGOs in EPI delivery was not known at the time of the study. The national-level data do not include all the costs incurred by NGOs and, therefore, an attempt to estimate the costs of urban EPI will be extremely useful for calculating the actual cost of immunization in Bangladesh.

MATERIALS AND METHODS

Study design and sampling

This facility-based study estimated the costs of providing routine EPI services from the perspective of EPI service providers. A comprehensive list of all the facilities involved in the delivery of EPI services in Dhaka city was used as the sampling frame to select a random sample of facilities. The then Urban Health Programme of ICDDR,B prepared the list to better understand the supply environment of primary healthcare services in Dhaka city (5). Information contained in the list was used for stratifying the EPI delivery sites by type (static and outreach) and location (zone within Dhaka city). For the classification of the EPI sites by type, health centres operating one day or less per week were defined as outreach sites, while all others were categorized as static sites. From each of the strata defined, 25% of the facilities, chosen at random, generated a sample of 132 EPI delivery sites. The classification of health facilities by ownership (government/NGO) could not be carried out prior to drawing of the sample due to lack of information. Since the study selected a large proportion (25%) of all EPI sites, the results of the survey should indicate the relative importance of the Government of Bangladesh (GoB) and NGO service providers in urban Dhaka.

Data collection

Facility-based data were collected from the EPI delivery sites for 1999. Two approaches were followed for collecting data on the use of resources, costs, and number of immunizations delivered. The first approach obtained information on the use of resources and the number of vaccinations administered from the record-keeping and accounting books of the facility. The second approach interviewed facility staff to obtain relevant additional information. In most cases, the manager or the vaccinator of the facility was interviewed. To ensure that the enumerators collect all the relevant data from the health facilities, a structured questionnaire was designed. The cost part of the instrument collected data on all the resources used in the process of delivering EPI services, including donated items, volunteer time, resources provided through other health activities, and space provided by the communities. The resources reviewed included a comprehensive list of capital and recurrent items. The capital items of EPI included vehicle, equipment, e.g. refrigerator, cold boxes, etc., furniture, e.g. tables, chairs, etc., and training of facility staff to increase human capital endowment (long-term training leading to a diploma or a degree). The recurrent items of EPI included salary (salaries and benefits of manager, vaccinator, physician, etc.), rent (rent, utilities, operation, and maintenance), vaccines, supplies, e.g. syringe, ice-pack, etc., transport, and recurrent training (short-term training for maintaining skills and knowledge of the service providers). For obtaining the annualized value of land and buildings, the study collected information on the current rent for all facilities. If the facility was owned by the service provider rather being rented from others, e.g. GoB facilities, the rent value for the facility was imputed at the average rent for sites of the same type (static/outreach) and location (zone).

Capital costs were annualized using a discount rate of 5%, and the economic life of all EPI-relevant capital items was assumed to be five years. For health-sector cost-effectiveness analysis, most researchers prefer using a low discount rate of 3-5%. Since a number of EPI costing studies used a 5% discount rate, using the same rate will allow an easy comparison of results of the study with prior studies. For non-exclusive resources, such as resources used in delivery of other primary healthcare services as well, costs were apportioned to EPI based on the proportion of time spent by the service providers on EPI activities. Cost data obtained were for 1999. All the local currency values were converted into US dollars using the 1999 exchange rate of US$ 1.00=Tk 49.50 (6). For costing the vaccines, the 1997 UNICEF prices were inflated by a factor of 2.5% per year. The survey collected information on other variables relating to EPI service-delivery, such as days of operation of the facility per year, hours of operation per day, and number of vaccines delivered per year. This study did not collect any information on household-level costs, such as travel costs of the mother and child to the EPI delivery facility.

RESULTS

EPI delivery sites and EPI sessions

The EPI delivery sites were usually located in or near residential areas of urban Dhaka. A typical static facility was located in a large building with multiple rooms providing health and non-health services to the population in the area. The types of services delivered include: maternal and child health services, curative care, family planning, microcredit activities, literacy sessions, etc. A typical outreach facility was located in a much smaller building in a residential area not well-connected to other parts of the city by main roads. Outreach sites do not have resident EPI staff, and teams travel there from other static sites.

Of the 132 sites surveyed by the study, less than a quarter were GoB-run facilities, and about 60% of all the sites were NGO-run outreach centres. In 1999, 38% of 11,028 EPI sessions in the surveyed sites were organized by the government static sites, 3% by the government outreach sites, 29% by the NGO static sites, and 31% by the NGO outreach sites. On average, the EPI delivery sites organized 84 (range 12-288) EPI sessions per site per year. NGOs played a very important role in the delivery of EPI services in urban Dhaka. About 77% of the EPI delivery sites in Dhaka city were under the management of NGOs, and these sites organized 60% of the EPI sessions. The predominance of NGOs in the delivery of EPI in urban Bangladesh is in sharp contrast to the delivery structure in rural areas, where it is almost exclusively a publicly-run programme.

Cost of EPI services

The cost of EPI service-delivery by various cost items is shown in Table 1. The total annual cost of routine EPI services in the surveyed EPI delivery sites was US$ 467,171. The capital cost constituted 24% of the total cost. Since EPI is a labour-intensive programme, personnel cost constituted 51% of the total cost. Table 1 shows that about 53% of the total EPI cost in urban Dhaka was due to the activities of NGOs. If we consider cost allocation within the GoB and NGO structures, about half of all EPI costs in the NGO sector was due to service-delivery through the outreach sites, while it was only 8% for the government sector outreach sites. This indicates the emphasis NGOs assign on delivering EPI services from outreach sites rather than from static sites.

Table 2 reports the average cost per facility by ownership-type and facility-type categories. The average cost of running an EPI facility was US$ 3,500 per year in Dhaka city in 1999. However, the costs varied significantly by ownership type, i.e. whether the facilities were run by NGOs or GoB. In general, the static sites were more expensive to organize than the outreach sites for both the GoB and NGO sectors. The average cost of running a static and an outreach delivery site was about US$ 7,500 and US$ 2,100 respectively. The cost of running a GoB static site was US$ 8,300 compared to US$ 6,500 for NGOs. NGOs needed less money to run the outreach sites—US$ 1,300 per site per year—compared to US$ 2,900 for the government sites. The NGO outreach sites had a much lower salary cost, as they usually had only vaccinators to provide services. As expected, the permanent static sites used capital items much more intensively than the outreach sites. On average, the capital cost of the static sites was about 30% of the total EPI cost and only about 5% for outreach sites.

Effectiveness of delivery structure

Table 3 presents a number of effectiveness or output measures of urban EPI. The surveyed EPI delivery sites provided 508,188 vaccinations through 11,028 EPI sessions in 1999. The distribution of the number of vaccinations administered was as follows: BCG 10%, DPT 24%, OPV 30%, measles 7%, vitamin A 13%, and TT 15%. The highest number of vaccinations was due

Table 1. Total annual cost (US$) of immunization in surveyed sites in 1999

Cost	GoB static (n=24)	GoB outreach (n=6)	NGO static (n=22)	NGO outreach (n=80)	Total cost (n=132)	% of total cost
Capital cost						
Vehicle	0	0	212	300	512	0.11
Equipment	1,913	127	4,341	2,687	9,067	1.94
Furniture	46,392	106	2,584	1,071	50,154	10.74
Training (non-recurrent)	26,397	0	21,998	4,400	52,794	11.30
Subtotal	74,702	233	29,134	8,457	112,526	24.20
Recurrent cost						
Salary	90,740	13,480	73,622	59,308	237,149	51
Rent	7,585	507	9,143	3,624	20,860	4.47
Vaccine	24,904	2,669	29,895	29,568	87,036	18.63
Supplies	944	148	1,129	1,560	3,781	0.81
Training (recurrent)	380	202	970	1,846	3,398	0.73
Transport	1,078	144	631	568	2,421	0.52
Sub-total	125,631	17,150	115,390	96,474	354,645	75.80
Total cost	200,333	17,383	144,524	104,931	467,171	

GoB static: Government-run static sites; GoB outreach: Government-run outreach sites; NGO static: NGO-run static sites; NGO outreach: NGO-run outreach sites

to the delivery of OPV, and the lowest was for measles. DPT and OPV doses were supposed to be delivered together, but the number of DPT doses delivered was about 19% lower than that of OPV. This probably indicates the relative difficulty of delivering injectables compared to an oral vaccine. Table 3 indicates that 34% of all vaccinations was carried out by the government static sites, 4% by the government outreach sites, 38% by the NGO static sites, and 24% by the NGO outreach sites. The static sites immunized more children and women compared to the outreach sites for all six antigens in the routine EPI. On average, 46 vaccinations were provided per EPI session organized or about 12 vaccinations per hour of session. A number of delivery sites reported zero vaccinations during their EPI sessions. An EPI session providing no vaccination at all indicates the presence of slack time of EPI delivery staff due to lack of demand. All the EPI delivery sites also reported a significant wastage of vaccines, and the wastage rates were used for estimating the total cost of immunization.

Table 2. Average cost (US$) per facility by type and ownership of facility in 1999

Cost	GoB static	GoB outreach	NGO static	NGO outreach	Average cost/site
Mean capital cost					
Vehicle	0	0	10	4	4
Equipment	80	21	197	34	69
Furniture	1,933	18	117	13	380
Training (non-recurrent)	1,100	0	1,000	55	400
Sub-total	3,113	39	1,324	106	852
Range	12-44,548	11-59	26-13549	6-4435	6-44548
Mean recurrent cost					
Salary	3,781	2,247	3,346	741	1,797
Rent	316	85	416	45	158
Vaccine	1,038	445	1,359	370	659
Supplies	39	25	51	20	29
Training (recurrent)	16	34	44	23	26
Transport	45	24	29	7	18
Sub-total	5,235	2,858	5,245	1,206	2,687
Range	1,493-47,958	808-4,141	556-13,829	278-6,427	278-47,958
Total (mean)	8,347	2,897	6,569	1,312	3,539
Range	1,601-49,507	838-4,187	592-23,235	312-6,670	312-49,507

Table 3. Total annual number of vaccinations delivered by type and ownership of facility in 1999

Vaccine	GoB static (n=24)	GoB outreach (n=6)	NGO static (n=22)	NGO outreach (n=80)	Total A	Vaccination/ facility B=A/132	Vaccination/ session C=A/11,028	Vaccination/ hour of session D=C/4*
BCG	18,276	1,500	17,580	14,256	51,612	391	4.68	1.17
DPT	47,352	5,424	45,780	25,884	124,440	943	11.28	2.82
OPV	59,892	6,240	53,760	33,912	153,804	1,165	13.95	3.49
Measles	15,228	1,668	11,424	8,988	37,308	283	3.38	0.85
Vitamin A	12,672	2,712	31,092	19,368	65,844	499	5.97	1.49
TT	21,876	2,364	31,284	19,656	75,180	570	6.82	1.71
Total	175,296	19,908	190,920	122,064	508,188	3,850	46.08	11.52

*The number of facilities surveyed was 132, and these facilities organized 11,028 sessions during 1999. Since the average duration of a session was 4.0 hours, total hours of sessions can be calculated by multiplying the number of sessions by 4

BCG=Bacille Calmette Guerin; DPT=Diphtheria-pertussis-tetanus; OPV=Oral polio vaccine; TT=Tetanus toxoid

Average cost of delivering EPI

Using the numbers reported in Tables 1 and 3, we can calculate the average cost per unit of output produced by the EPI delivery sites. Table 4 reports the average costs per unit of various outcome measures. The average cost per EPI session in 1999 was about US$ 42, while the average cost per dose administered, excluding vitamin A and tetanus toxoid, was US$ 1.18. Since measles is the last vaccine a child should get in the EPI schedule, the number of children immunized against measles can be used as an indirect measure of fully-immunized children. The average cost per measles-vaccinated child (MVC) was US$ 11.61, and the average cost was lower for the NGO facilities compared to that for the government facilities. We do not have any information on the number of children fully immunized by 12 months of life (FIC). In our sample, the estimated number of children immunized against BCG, DPT, OPV, and measles was 51,612, 41,480, 51,268, and 37,308 respectively. Since the number of children immunized against measles was lower compared to other vaccinations, we can use MVC as a rough guide of FIC. Therefore, US$ 11.61 may be considered an approximation of per FIC cost in urban Dhaka.

Table 4 also reports a hypothetical number, cost per FVC, and cost of providing all the EPI vaccinations to all infants without incomplete vaccinations (some children receiving only few vaccines) or double-dosing. This hypothetical cost per FVC is simply the total cost of providing three doses of DPT, three doses of OPV,

Table 4. Average cost (US$) per unit of output in 1999

Cost	GoB static	GoB outreach	NGO static	NGO outreach	Average cost
Cost per session	48	60	46	31	42
Cost per hour of session	11.56	17.25	7.92	8.00	10.06
Cost per dose (without TT and vitamin A)	1.40	1.13	1.02	1.06	1.18
Cost per dose (with TT and vitamin A)	1.14	0.87	0.76	0.86	0.92
Cost per FVC	8.07	6.81	6.19	6.16	6.91
Cost per MVC (without TT and vitamin A)	12.93	10.07	11.50	9.80	11.61
Cost per MVC (with TT)	13.05	10.17	11.65	9.97	11.75
Cost per MVC (with vitamin A)	13.03	10.32	12.51	11.51	12.38
Cost per MVC (with TT and vitamin A)	13.16	10.42	12.65	11.67	12.52

Cost per session and cost per hour of session include TT and vitamin A

Tables with numbers of EPI sessions and hours of EPI session not shown

Cost per dose (without TT and vitamin A)=[total cost–TT vaccine cost–vitamin A vaccine cost–transport cost*2/6–supply cost*3/10] / [total dose–TT dose–vitamin A dose]

Cost per MVC (without TT and vitamin A)=[total cost–TT vaccine cost–vitamin A vaccine cost–transport cost*2/6–supply cost*3/10] / [measles dose]

Cost per MVC (with TT)=[total cost–vitamin A vaccine cost–transport cost*1/6–supply cost*1/10]/[measles dose]

Cost per MVC (with vitamin A)=[total cost–TT vaccine cost–transport cost*1/6–supply cost*2/10]/[measles dose]

FVC=Fully vaccinated child; MVC=Measles-vaccinated child one dose of BCG, and one dose of measles vaccines to a child. FVC was computed in two steps: first, cost per specific antigen was calculated, and then FVC was computed (Tables 5 and 6 in Appendix). The average estimated cost per FVC is only about US$ 6.91, implying that many children received partial immunizations (lower completion rate due to drop-outs), and some might have received the same vaccines more frequently than the EPI schedule suggests. The cost per MVC (US$ 11.61), in general, should be close to the hypothetical cost per FVC (US$ 6.91) in the absence of significant partial vaccinations or double-dosing. The high cost of MVC compared to the hypothetical minimum cost indicates that the system (for both GoB and NGOs) can be made much more effective if children are identified and vaccinated in a timely manner without significant mistargeting or double-dosing. For the purpose of estimating the costs without mistargeting or double-dosing, it is not necessary to identify the mistargeted cases. If the number of children receiving measles vaccination were fully immunized, we can calculate the total vaccination cost for the cohort. The ratio of this hypothetical cost and actual cost may be used as a measure of degree of mistargeting by both GoB and NGOs.

Cost per vaccinated child, either the cost per MVC or the hypothetical cost per FVC, can be used as a measure of efficiency of the EPI delivery system. Table 4 indicates that the cost per MVC was the highest (US$ 12.93) for the government static sites and was the lowest (US$ 9.80) for the NGO outreach sites. Between the government and the NGO delivery structures, the NGO static facilities were more cost-effective (US$ 11.50) than the government static facilities (US$ 12.93). The NGO outreach sites were also more cost-effective than the government outreach sites (US$ 9.80 and US$ 10.07 per MVC respectively). If the cost of delivering TT vaccines is included with other vaccines, the average cost per MVC increases by about 14 cents. If the cost of distributing vitamin A is added, the average cost per MVC increases by 77 cents. Therefore, adding these other services with the traditional vaccine does not increase the cost per child significantly. The incremental cost of adding a new vaccine will be slightly higher than the cost of the vaccine itself. The additional cost of administering the vaccine or distribution of vitamins appears relatively low.

Financing of EPI

The EPI activities of the Ministry of Health and Family Welfare (MoHFW), GoB, are supported by a donor consortium comprising GoB, World Bank, United Nations Children's Fund, World Health Organization, U.S. Agency for International Development, Japanese International Cooperation Agency, and Department for International Development-UK. Additional donor involvement was found in the surveyed EPI delivery sites of Dhaka City Corporation (DCC), such as Norweigian Aid, Swedish International Development Agency, Ford Foundation, Action Aid, etc. These additional sources of support can be categorized into three groups: (a) agencies providing both monetary and logistical (vaccines, supplies, training) support, (b) agencies providing only monetary support, and (c) agencies providing only logistical support. The resources received by all EPI service implementers from the EPI Headquarters were vaccines, supplies, EPI-related training, and some capital equipment. If we exclude these common resources, the additional resources that NGOs mobilized for EPI were about US$ 177,460 for the surveyed facilities. If we project this cost for urban Bangladesh, the additional resources mobilized by NGOs for EPI services become US$ 1.4 million. Since these resources do not show up in the macro-level cost accounting of EPI, the cost of delivering EPI is usually underestimated. Furthermore, NGOs in Dhaka were able to generate about US$ 15,000 (of US$ 177,460) from local community resources. This was estimated from the resources used by the NGO outreach sites where most space (rent) and furniture were provided by the local community, such as a room in private households, schools, pharmacy, cultural clubs, etc. Thus, even the poor communities of the city can potentially support some EPI activities.

DISCUSSION

EPI is one of the most cost-effective health interventions with high potential benefits and low costs (3,4,7-12). Most cost studies of EPI used national- or regional-level secondary data without supplementing information by collecting facility-level data. This study estimated the cost of delivering EPI in urban Bangladesh using facility-based surveys. The survey results indicate that the secondary data sources would have underestimated the urban EPI costs by at least 40-50%. The NGO outreach-delivery structure is highly dependent on community-level resources, and none of these are accounted for in the secondary data. Even the government delivery system solicits additional resources from the communities around their outreach sites. Despite the underestimation of costs, EPI remains a highly cost-effective intervention. If we use cost per MVC as a measure of cost per fully-immunized child, the cost remains less than US$ 15 per child. This excludes the societal costs of vaccination that were not assessed in this study.

An important conclusion of this study is that it is feasible to generate a significant amount of local resources for delivering EPI services. All the NGO outreach sites mobilized resources from the communities in which they work. Therefore, it is feasible to generate some local resources even from poor regions for conducting immunization services. Involving the local community with EPI activities not only will improve the sustainability of the programme but will also help increase rates of immunization coverage. Furthermore, in the absence of community involvement, GoB and NGOs would have to supply these resources, especially if emphasis is put on the delivery of EPI through static sites. The additional resources generated by NGOs included resources from local communities and from additional donor agencies. The estimated additional resource generated by NGOs in urban Bangladesh was about US$ 1.4 million per year. If we add this cost with the estimates of Levin *et al.* (4), total cost of EPI for Bangladesh becomes about US$ 31 million, about 6% higher than their estimate. Although it is not a very significant increase in total cost, it is important to derive the actual resource use in the EPI programme for planning and policy analysis.

If the average costs of delivering different types of services are considered, it is clear that the outreach facilities (both government and non-government) are more cost-effective than the static facilities. The NGO-outreach sites delivered EPI services at the lowest average cost, probably due to the externality created by community participation, using capital items less intensively and having minimal staff providing services. It is usually assumed that the public sector must organize and deliver preventive services, especially in poor countries where the demand for preventive services is expected to be low. The fact that NGOs delivered 62% of all immunizations in urban Dhaka clearly demonstrates no inherent disadvantage of NGOs compared to the public sector in providing immunization services. Furthermore, NGOs in Dhaka delivered EPI services at a lower cost than the government sites, which suggests that NGOs can successfully organize and deliver preventive services in a poor community and, in the case of urban Dhaka, they were more efficient than the GoB.

Another important finding of the study is that the incremental cost of adding services should not be significantly higher than the actual cost of new vaccines or drugs to be added. The new vaccine will obviously increase the cost of acquiring the commodities and supplies, but the current delivery structure has enough slack in the system to be able to deliver the new vaccine without employing additional personnel or other inputs. For example, the number of vaccine doses delivered, including the distribution of vitamin A capsules, was less than 12 per hour of EPI session in urban Dhaka. This number can be increased by 50% without changing the size of the facilities or the number of personnel involved with delivery.

This study also indicates that the current EPI delivery structure could be made more efficient. Apart from the wastage of vaccines and slack time of personnel, better targeting of children alone should significantly lower the average cost of EPI. If the completion rate of vaccination can be improved and double-dosing avoided, cost per MVC should decline to about US$ 7. The estimated cost per MVC was US$ 11.61, indicating that perfect targeting can reduce the cost per FVC by about 60%. However, no system can be 100% efficient in terms of targeting or completion rates, but it should be possible to reduce the cost per MVC by at least US$ 2-3 by better managing the delivery structure, training providers, and mobilizing the community. Better use of existing human resources and vaccines should reduce the cost per FVC even further without increasing the service-delivery costs.

One of the important aspects of the EPI delivery structure identified by the study is the complex nature of the system in urban Dhaka. The predominance of NGOs in the delivery of EPI in urban Bangladesh is in sharp contrast to the EPI delivery structure in rural areas, where it is almost exclusively a publicly-run programme. Despite the high degree of involvement of the private sector in urban EPI, the delivery structure has remained relatively inefficient. Therefore, sub-contracting health activities to the private sector, by itself, may not improve efficiency in the delivery of EPI. It is important to identify the factors affecting the efficiency of NGO and government facilities, including the payment mechanisms adopted by the contracting arrangement.

ACKNOWLEDGEMENTS

This study was conducted by the Health Economics Programme (HEP) of ICDDR,B: Centre for Health and Population Research with funding from Health Systems and Economics Unit and WHO/London School of Hygiene and Tropical Medicine. The DFID funding through HEP (currently Health Systems and Economics Unit) allowed the study to conduct a much more comprehensive study than the initial research proposal envisaged. Our cordial thanks to the EPI officials, DCC officials, NGO officials, private providers, UNICEF (Dhaka) officials, and EPI delivery site staff for their cooperation in providing information. We also thank the Health Systems and Economics Unit field interviewers for their effort and dedication during data collection. A full report of this study with detailed calculations of the estimations and the data collection instrument is available at HEP, ICDDR,B.

REFERENCES

1. Bangladesh. Ministry of Health and Family Welfare. Expanded Programme on Immunization. EPI programme review 1998. Dhaka: Expanded Programme on Immunization. Directorate General Health Services. Ministry of Health and Family Welfare, Government of Bangladesh, 1998. 31 p.

2. Bangladesh. Ministry of Health and Family Welfare. Expanded Programme on Immunization. National coverage evaluation survey, Bangladesh 1998. Survey report on routine EPI and NID coverage and supplement on progress in AFP surveillance. Dhaka: Expanded Programme on Immunization, Directorate General of Health Services. Ministry of Health and Family Welfare, Government of Bangladesh. 1998. 11 p.

3. Khan MM, Yoder RA. Expanded Programme on Immunization in Bangladesh: cost, cost-effectiveness, and financing estimates. Bethesda, MD: Partners for Health Reform Project, 1998. 61 p. (PHR technical report no. 24).

4. Levin A, Howlader S, Ram S, Siddiqui SM, Razul I, Routh S. Case study on the costs and financing of immunization services in Bangladesh. Bethesda, MD: Partners for Health Reform Project, 1999. 82 p. (Special initiatives report no. 21).

5. Mazumder MA, Bhuiyan MA, Tunon C, Baqui AH, Chowdhury AI, Khan SA, Islam R. An inventory of health and family planning facilities in Dhaka city. Dhaka: MCH-FP Extension Project (Urban). Health and Population Extension Division. International Centre for Diarrhoeal Disease Research, Bangladesh, 1997. 121 p.

6. Bangladesh. Ministry of Finance. Bangladesh economic review 2003. Dhaka: Economic Adviser's Wing, Finance Division, Ministry of Finance, Government of Bangladesh, 2003. 56 p.

7. Brenzel L, Claquin P. Immunization programs and their costs. *Soc Sci Med* 1994;39:527-36.

8. Jha P, Bangoura O, Ranson K. The cost-effectiveness of forty health interventions in Guinea. *Health Policy Plan* 1998;13:249-62.

9. Phonboon K, Shepard DS, Ramaboot S, Kunasol P, Preuksaraj S. The Thai expanded programme on immunization: role of immunization sessions and their cost-effectiveness. *Bull World Health Organ* 1989;67:181-8.

10. Shepard DS, Robertson RL, Cameron CS, 3rd, Saturno P, Pollack M, Manceau J *et al.* Cost-effectiveness of routine and campaign vaccination strategies in Ecuador. *Bull World Health Organ* 1989;67:649-62.

11. World Bank. Investing in health: the world development report 1993. Washington, DC: World Bank, 1993. 329 p.

12. Pegurri E, Fox-Rushby JA, Walker D. The effects and costs of expanding the coverage of immunization services in developing countries: a systematic literature review. *Vaccine* 2004. (In press).

Appendix

Table 5. Average cost (US$) per specific antigen in 1999

Vaccine	GoB static	GoB outreach	NGO static	NGO outreach	Average cost
BCG	1.79	1.82	1.23	1.03	1.39
DPT	0.73	0.54	0.52	0.62	0.63
OPV	0.63	0.55	0.49	0.52	0.55
Measles	2.20	1.72	1.93	1.71	1.98
Vitamin A	2.41	1.04	0.97	1.42	1.38
TT	1.41	1.09	0.65	0.70	0.90

Cost per specific antigen=[capital cost/6+salary/6+rent/6+recurrent training/6+transport cost/6+supply cost*1/10+specific vaccine cost]/no. of specific vaccine doses administered

Supply cost multiplied by 2/10 if vaccine is injectable

GoB static: Government-run static sites; GoB outreach: Government-run outreach sites; NGO static: NGO-run static sites; NGO outreach: NGO-run outreach sites

BCG=Bacille Calmette Guerin; DPT=Diphtheria-pertussis-tetanus; OPV=Oral polio vaccine; TT=Tetanus toxoid

Table 6. Costs (US$) of specific antigens* in 1999

Vaccine	GoB static	GoB outreach	NGO static	NGO outreach	Total cost
BCG	3,492.10	277.47	2,422.82	2,050.59	8,242.99
DPT	5,605.08	476.07	4,891.88	3,453.77	14,426.80
OPV	8,772.95	1,011.09	7,067.38	5,093.21	21,944.64
Measles	4,231.01	410.37	2,915.01	2,801.80	10,358.18
Vitamin A	1,308.62	372.84	11,273.04	15,068.46	28,022.95
TT	1,494.05	121.57	1,324.49	1,100.54	4,040.65
Total	24,903.81	2,669.40	29,894.61	29,568.38	87,036.21

*Costs for antigens included doses administered and doses wasted

Appendix 2: Valdmanis et al. (2003)

ARE VACCINATION SITES IN BANGLADESH SCALE EFFICIENT?

Vivian Valdmanis

Damian Walker

Julia Fox-Rushby

London School of Hygiene & Tropical Medicine

Abstract

Objectives: The overall aim of this study is to discern whether and to what degree vaccination sites exhibit constant returns to scale.
Methods: Data Envelopment Analysis is used to compare all the facilities in the sample in terms of input costs used to produce multiple outputs. The application considers the Expanded Program on Immunization (EPI), which operated in Dhaka City, Bangladesh, during 1999.
Results: A preponderance of EPI sites were determined to be operating at increasing returns to scale.
Conclusions: Our findings question the applicability of cost-effectiveness analyses that assume constant returns to scale.

Keywords: Scale, Efficiency, Immunization, Data envelopment analysis, Bangladesh

Compared with other health interventions, vaccinations are judged to be one of the most cost-effective ways of improving and maintaining child health, especially in low-income countries (16). This view has been held for a considerable time (e.g., 15) and may help to explain the increase in global coverage of the Expanded Program of Immunization (EPI) from an average of 5 percent at its inception in 1974 to the current average of 80 percent (4). Many cost and cost-effectiveness analyses of EPI country programs in low-income countries have been evaluated at a given level of production (e.g., 11), used only a few providers (e.g., 4), or aggregated and averaged at a country level (e.g., 1;14). Even when studies estimated the costs of increasing coverage rates or predicted country-wide estimates of costs from a small study, most have assumed a linear function to "scale-up" programs (10). For example, if the unit cost per fully vaccinated child is \$20, the increase in expanding vaccination services for another fifty children is assumed to be \$1,000.

That such constant returns to scale exist is doubted. For example, England et al. (5) have hypothesized that many impediments exist to scaling up measles control in West and Central Africa and suggested that considerable investment would be needed in management and health systems before expansion. In reviewing the cost profiles of immunization programs from accounting-based cost studies, some investigators have found that the proportion of fixed costs indicates the likely existence of economies of scale (e.g., 8).

If average costs and incremental cost-effectiveness ratios did change with production, then assuming constant returns to scale would produce biased estimates of any change in production, and the bigger the expected change, the larger the bias. Even if size were accounted for, there is no notion of best practice benchmarking (2) or knowledge of how this might change by setting. In this study, both of these issues are addressed by a novel

We acknowledge and thank Dr. Suhaila Kahn at ICDDR.B for her organization of the data collection and Dr. Mahmud Khan at Tulane University who was one of the principal investigators on the same study. Vivian Valdmanis, Damian Walker, and Julia Fox-Rushby are members of the Health Economics and Financing Program, which is supported by funds from the United Kingdom Department for International Development (DFID).

application of data envelopment analysis. Three objectives are pursued: first, the cost of delivering routine vaccination services from the perspective of the providers is determined; second, the outputs of vaccination sites for each provider in terms of the number of doses of each type of vaccine is assessed; third, the scale efficiency of the vaccination sites as well as factors that explain variation in scale efficiency are evaluated.

THE BANGLADESH EPI

The EPI in Bangladesh was established in 1979 and became fully operational in 1985. It aimed to reduce morbidity and mortality from six vaccine-preventable diseases. Therefore, a fully vaccinated child received six standard EPI antigens against diphtheria, pertussis, and tetanus (DPT), tuberculosis (TB), polio, and measles through eight vaccinations. Pregnant women were also given vaccinations to prevent maternal and neonatal tetanus. Since 1985, vaccination coverage has increased from 2 percent for all antigens to a reported 92 percent for BCG and 62 percent for measles (16). However, immunization coverage rates were much lower in urban compared with rural areas. Therefore, in 1988, the United States Agency for International Development implemented a program to strengthen vaccination services in urban areas in conjunction with the array of government and nongovernment funders and providers of service.

METHODS

Data envelopment analysis (DEA) was used to allow comparison of all the clinics in the sample in terms of input costs used in the production of multiple outputs. DEA is a nonparametric, deterministic approach using linear programming techniques that defines a "best practice" production frontier. Firms lying on the production frontier are considered to be operating at the best practice or in other words, provide a benchmark à la Birch and Gafni (2). However, it should be noted that the measure of efficiency is considered to be relative rather than absolute, as no a priori information exists as to what should be considered as absolute efficiency. The benchmark clinics, that is, those that are technically and scale efficient, reflect the best practice for the given sample of clinics.

A benefit of this DEA approach is that, by identifying best practice by a "local" standard, it may be assumed that given certain productive characteristics (as well as environmental ones) best practice can be feasibly reproduced at the less-efficient clinics. Another benefit of the DEA approach used here is that the overall technical efficiency (TE^{CRS}) measure can be decomposed into pure technical efficiency (TE^{VRS}) and scale efficiency (SE). In other words, TECRS = TEVRS × SE.

Whereas there have been a plethora of other related studies applying DEA to the health care sector using quantities of inputs in their natural units to produce outputs (see 13 for a review), we specified the objective as minimizing input costs given outputs (6;7). As the objective of this study is to determine scale effects, the definition of the cost minimizing technology used here was applicable.

The technology was initially constructed under constant returns to scale and strong disposability of costs (as costs increase, outputs must increase, ceteris paribus) TE^{CRS}. Allowances can be made in the restraints to allow for variable returns to scale TE^{VRS}. Furthermore, we determined the type of scale inefficiencies by using a third model TE^{NIRS}. In all these cases, we followed the definitions given by Färe, Grosskopf, and Lovell (7) and solved similar linear programming problems. We used the DEAP program by Coelli (3) for the computations.

The technology is said to be operating at a cost- as well as scale-efficient level if $TE^{CRS} = TE^{VRS}$. However, if they were not equal, the extent to which inefficiency was caused due to operating at the wrong scale was assessed. Determining the type of scale inefficiency (either increasing or decreasing returns to scale) required the solution of a third linear programming problem, referred to as nonincreasing returns to scale technology (NIRS). To define the type of scale inefficiency that is operating here, we compared the solutions of the three linear programming problems. If $\frac{TE^{CRS}}{TE^{VRS}} < 1$, $TE^{CRS} = TE^{NIRS}$ then increasing returns to scale exist. If $\frac{TE^{CRS}}{TE^{VRS}} < 1$, but, $TE^{NIRS} > TE^{CRS}$, then decreasing returns to scale exist. Such models allowed for the impact of scale effects on the EPI clinics to be evaluated.

However, deviations from the best practice frontier may be due to independent factors that may be out of the managers' or policy makers' direct control. Therefore, the measures of efficiency were analyzed by using a variety of statistical tests, in conjunction with other environmental factors that may affect scale efficiency.

DATA AND RESULTS

Our sample was obtained by means of a 1999 cost analysis of EPI services undertaken in a random sample of 25 percent of the facilities (132 of 511) providing EPI services in Dhaka City Corporation. To be parsimonious, five outputs (the amount of doses given for DPT, TB, polio, measles, and TT in 1999) and one input (total program costs of the EPI by site) were specified. Only program sites with full information were included. The final data set consisted of 117 of a possible 132 total clinics. Hence, 89.3 percent of all clinics sampled were included. The type of missing data that resulted in sites being excluded from the sample included ownership form, type of vaccination site, duration of operation, as well as some of the outputs. The descriptive statistics are given in Table 1.

Turning next to our efficiency results given in Table 2, we found that overall efficiency (TE CRS) was only 0.33. In other words, if program sites were technically efficient and operated at the correct scale, costs on average could have been reduced by 67 percent without sacrificing the current level of outputs produced. By decomposing this overall measure into pure technical efficiency (TE VRS) and scale efficiency, we found that more of the overall inefficiency was due to sites incurring too much cost in producing the array of vaccinations rather than operating at the wrong size. However, both sources of this overall inefficiency must be addressed for these sites to become less wasteful of scarce resources.

Given the findings that the sites in this sample exhibited variable returns to scale, the types of diseconomies of scale were examined next. Table 3 shows that the majority of the program sites exhibited increasing returns to scale (suggesting that they are too small), 17 program sites exhibit decreasing returns to scale (suggesting that they are too large), and only six program sites were the "right" size.

In Tables 4 and 5, we assessed whether differences in efficiency followed systematic patterns due to factors beyond managerial control. Table 4 displays statistically significant differences between the efficiency of two ownership forms, and shows that scale efficiency is relatively greater in government-owned program sites. As outreach sites were statistically significantly less scale efficient than fixed sites, we infer that satellite sites are too small given the best practice frontier.

Although the EPI program has been in existence in Dhaka City Corporation since 1988, not all sites began providing EPI services at the same time. Table 5 shows that the length of time a program site has been in operation is positively correlated with scale efficiency.

DISCUSSION/POLICY IMPLICATIONS

The sites in our sample were, on average, relatively inefficient both in terms of technical inefficiency as well as scale inefficiency. To become technically efficient, program sites

Table 1. Descriptive Statistics of Outputs and the Inputs

Variable	Mean	SD	Min	Max
BCG	257.40	304.94	1	1,680
DPT	578.57	685.54	1	3,264
Polio	707.42	842.91	1	3,756
Measles	190.28	210.83	1	960
TT	390.03	443.37	1	2,208
Total costs	2,600.31	4,972.79	238	45,716

Table 2. Descriptive Statistics of Efficiency Measures

Measure	Mean	SD	Min	Max
TE CRS	0.33	0.26	0.001	1.00
TE VRS	0.50	0.29	0.012	1.00
Scale	0.64	0.27	0.007	1.00

Table 3. Returns to Scale in Vaccination Sites

Types of returns to scale	Number of vaccination sites
Increasing	95
Constant	6
Decreasing	17

Table 4. Selected Statistics between Ownership and Type of Clinics and Efficiency

	Mean Scale efficiency score	F-test ($p > F$)	Median test ($p > Z$)	Kruskal-Wallis ($p > \chi^2$)
Government ($N = 25$)	0.77			
NGO ($N = 92$)	0.60	8.82 (.003)	2.47 (.01)	9.77 (.002)
Fixed ($N = 35$)	0.79			
Outreach ($N = 82$)	0.57	19.73 (.0001)	3.81 (.0001)	17.80 (.0001)

NGO, not government owned.

Table 5. Correlation Coefficients for Time Since EPI Clinic Opened and Total Cost and Scale

Variables	Correlation coefficient	$p > \|r\|$
Time/scale	0.34	(.0001)
Total costs/scale	0.16	(.08)

EPI, Expanded Program on Immunization.

would have had to decrease their costs by an average of 50 percent, and if they had been operating at the right size, costs could have been reduced by a further 36 percent. Sites that were relatively more inefficient, on average, were not government-owned satellites. Therefore, the governmentally owned sites, perhaps due to more centralized control, appeared to be better at long-term planning. We also found that sites that had been practicing longer were relatively more scale efficient, which is perhaps attributable to a learning curve effect.

The presence of pure technical inefficiencies suggests that, if such cost data were used as the numerator of a cost-effectiveness ratio, a cost-effectiveness analysis would not reflect the minimum efficient point of production at a given level. However, to ascertain whether this outcome is likely to be the case, researchers need to begin using a larger sample size of provider units for costing, especially if results are intended for use beyond the geographical focus of an evaluation.

Our evidence provides empirical support to Jacobs and Baladi's (9) contention that assuming constant returns to scale might not be realistic. The presence of increasing returns has two particular implications. First, that this intervention cannot be treated as perfectly divisible within a population and retain the same level of incremental cost-effectiveness. Second, it suggests that, if constant returns to scale are assumed when increasing returns to scale exist, an intervention is likely to be overprovided in that form. Finally, the potential learning effect raises questions about how relevant it is to transfer cost-effectiveness ratios over time or across countries as levels of technology differ (12). Therefore, we conclude that ignoring the possible existence of technical inefficiencies and variable returns to scale would make the generalizability of cost-effectiveness ratios suspect and could worsen rather than improve the allocation of resources.

REFERENCES

1. Barnum HN, Tarantola D, Setiady IF. Cost-effectiveness of an immunization programme in Indonesia. *Bull World Health Org*. 1980;58:499-513.
2. Birch S, Gafni A. Cost-effectiveness/utility analysis. Do current decision rules lead us to where we want to be? *J Health Econ*. 1992;1:279-296.
3. Coelli TJ. *A guide to DEAP version 2.1: A data envelopment analysis computer program*. CEPA Working Paper No. 8/96. ISBN 1 86389 4969. Department of Econometrics, University of New England; 1996.
4. Cutts FT, Phillips M, Kortbeek S, et al. Door-to-door canvassing for immunization program acceleration in Mozambique: Achievements and costs. *Int J Health Serv*. 1990;20:717-725.
5. England S, Loevinsohn B, Melgaard B, et al. *The evidence base for interventions to reduce mortality from vaccine-preventable diseases in low and middle-income countries*. CMH Working Paper Series, Paper No. WG5 10. Geneva, Switzerland: Commission on Macroeconomics and Health; 2001.
6. Färe R, Grosskopf S. A non-parametric cost approach to scale efficiency. *Scand J Econ*. 1985;87:594-604.
7. Färe R, Grosskopf S, Lovell CAK. *Production functions*. Cambridge: Cambridge University Press; 1994.
8. Gilson L. *Quality and cost of primary health care in rural Tanzania*. PhD thesis. London: School of Hygiene and Tropical Medicine; 1992.
9. Jacobs P, Baladi JF. Biases in cost measurement for economic evaluation studies in health care. *Health Econ*. 1996;5:525-529.
10. Kumaranayake L, Watts C. Costs of scaling up HIV program activities to a national level for sub-Saharan Africa: Methods and estimates. *S Afr J Econ*. 1999;68:1012-1032.
11. Levine OS, Ortiz E, Contreras R, et al. Cost-benefit analysis for the use of Haemophilus influenzae type b conjugate vaccine in Santiago, Chile. *Am J Epidemiol*. 1993;137:1221-1228.
12. Ramsey CR, Grant AM, Wallace SA, et al. Assessment of the learning curve in health technologies. A systematic review. *Int J Technol Assess Health Care*. 2000;16:1095-1108.

13. Seiford LM. A bibliography of data envelopment analysis. In: Charnes A, Cooper WW, Lewin AY, Seiford LM, eds. *Data envelopment analysis: Theory, methodology and application*. Boston: Kluwer Academic Publishers, 1998:437-469.
14. Shepard DS, Sanoh L, Coffi E. Cost-effectiveness of the expanded programme on immunization in the Ivory Coast: A preliminary assessment. *Soc Sci Med*. 1986;22:369-377.
15. Walsh JA, Warren KS. *Strategies for primary health care: Technologies for the control of disease in the developing world*. Chicago: The University of Chicago Press; 1982.
16. World Bank. *World development report 1993. Investing in health*. Oxford: Oxford University Press; 1993.

Appendix 3: Dervaux et al. (2003)

International Journal of Health Care Finance and Economics, 3, 287–299, 2003

Parameters of Control when Facing Stochastic Demand: A DEA Approach Applied to Bangladeshi Vaccination Sites

B. DERVAUX AND H. LELEU

Department of Health Economics, LABORES, CNRS and Catholic University of Lille, 60 Boulevard Vauban, BP 109 59016 Lille Cedex, France

V. VALDMANIS AND D. WALKER

Health Policy Unit, London School of Hygiene and Tropical Medicine Keppel Street, London UK WC1E 7HT

An aim of vaccination programs is near-complete coverage. One method for achieving this is for health facilities providing these services to operate frequently and for many hours during each session. However, if vaccine vials are not fully used, the remainder is often discarded, considered as waste. Without an active appointment schedule process, there is no way for facility staff to control the stochastic demand of potential patients, and hence reduce waste. And yet reducing the hours of operation or number of sessions per week could hinder access to vaccination services. In lieu of any formal system of controlling demand, we propose to model the optimal number of hours and sessions in order to maximize outputs, the number and type of vaccines provided given inputs, using Data Envelopment Analysis (DEA). Inputs are defined as the amount of vaccine wastage and the number of full-time equivalent staff, size of the facility, number of hours of operation and the number of sessions. Outputs are defined as the number and type of vaccines aimed at children and pregnant women. This analysis requires two models: one DEA model with possible reallocations between the number of hours and the number of sessions but with the total amount of time fixed and one model without this kind of reallocation in scheduling. Comparing these two scores we can identify the "gain" that would be possible were the scheduling of hours and sessions modified while controlling for all other types of inefficiency. By modeling an output-based model, we maintain the objective of increasing coverage while assisting decision-makers determining optimal operating processes.

Keywords: data envelopment analysis, vaccination programs, Bangladesh

JEL classification: D2, I1

1. Introduction

Resource constraints in making health care allocations were highlighted at the recent Earth Summit in Johannesburg reported by Stulman et al. (2002) who asserted that "more work is necessary to assess the efficiency of resource utilization in specific countries". Therefore, eliminating current inefficiencies in health care programs may yield health and monetary gains. As part of this efficiency/effectiveness role of services in health care, vaccinations against preventable illnesses are reported to be one of the most cost-effective health care interventions provided (World Development Report, 1993).

"In 1983, the World Bank developed the concept that absence of health was a main obstacle to the economic development of poor countries and indicated that vaccination would be a first step to improved economies." (Rappuoli, Miller and Falkow, 2002). One aim of vaccination programs operating in low and middle-income countries is near-complete coverage. Specifically, the Commission on Macroeconomics and Health of the World Health Organization (WHO) expects a target of 80% coverage (England et al., 2001).

Both governmentally run clinics and private for-profit or non-profit non-governmental organizations (NGOs) also provide vaccinations. However, very little has been reported in the literature (either refereed or "gray") on the relative productivity of these sectors in providing vaccine or preventative services.[1] Property rights, however, may affect organizational goals that may lead to inefficient use of inputs, inefficient scheduling, or both.

The Expanded Program on Immunization (EPI) in Bangladesh was established in 1979 and became fully operational in 1985. A fully vaccinated child receives six standard EPI antigens against diphtheria-pertussis-tetanus (DPT), oral polio vaccine (OPV), and Bacillus of Calmette and Guérin (BCG) to ward off tuberculosis (TB) and measles. In addition, women of childbearing age also receive a course of vaccinations against tetanus (TT). The complete vaccination schedule includes eight vaccinations administered at five contacts with health care staff. The program has been very successful as evidenced by an increase in vaccination coverage rates from 2% for all antigens to a reported 92% for BCG and 62% for measles. However, more recently, coverage rates appear to have reached a plateau, with only 59% of children under the age of 1 year of age having received a full course of vaccines. The Government of Bangladesh (GoB)'s stated objective is "to increase coverage with a full series of routine vaccines gradually to at least 90% in all districts by 2005" (Walker et al., 2000).

Along with the attempt to increase vaccination rates, there is the other commensurate objective of reducing vaccine waste. Overall wastage rates[2] in developing countries have been estimated to be around 50% by the United Nations Children's Fund (UNICEF) and World Health Organization (WHO, 1999). Vaccine wastage is important as it can show program errors. For example, it can highlight that too many drops of OPV or the wrong dosage for other vaccines is used; cold-chain failures or poor logistics; and false reporting of more vaccinations administered than vaccine received. There are also economic implications associated with wastage. If wastage can be reduced without affecting coverage, it can result in significant fund savings for programs. This is especially true for very poor countries, which do not typically have budgetary flexibility to expand program financing.

In Bangladesh, vaccine wastage rate is also high (estimated to be around 40% for DPT during 1998–1999). Further, there is no set acceptable wastage rate that can be applied universally, however normal rates of wastage can be expected (EPI/WHO, 1983) from between 25% to 50%.

In view of this problem with vaccine waste, the Government of Bangladesh (GoB) has proposed the following solutions:

- The promotion of an open vial policy[3] for DTP, TT, and OPV;
- A reduction in the number of vaccination sessions.

On the supply side, one method for achieving complete vaccination coverage is for health facilities providing these services to operate frequently and to stay open for longer hours during each session. This permits individual flexibility in getting children vaccinated. However, scheduling of appointments is not done, rather patients enter on a walk-in basis for their vaccinations. Again, we stress that one problem with this policy is that if vaccines are not fully used during a session, they are often discarded and considered waste. Simply reducing the hours of operation or number of sessions per week to possibly reduce waste could hinder access to vaccination services. Without an active appointment schedule process however, there is no way for clinic staff to control the stochastic demand of potential patients. In lieu of any formal system of controlling arrivals for vaccinations via a scheduling system, we propose, in this paper, to model the optimal number of hours and number of sessions in order to maximize the vaccines delivered—using Data Envelopment Analysis (DEA).

In terms of an economic model, inefficiency can be viewed in two equivalent ways. First, we can measure inefficiency comparing the observed output(s) to an optimal set of output(s) that results from a maximizing output while maintaining input constant. Second, we can compare the observed input with an optimal input basket that comes from a minimization of input for a constant output. Here we opt for the former and hence our aim in this paper is to ascertain how clinics and health centers could maximize the numbers of delivered vaccines while maintaining wastage and other inputs constant. To accomplish this, we employ two models, one DEA model, which allows for the reallocation between the number of hours and the number of sessions and one model without this same reallocation. By taking the ratio of the two scores obtained via the DEA models the "gain" that can be made if the scheduling of hours and sessions are modified can be measured. In essence, we control for productive (in)efficiency (the conversion of inputs into outputs) and just focus on the time related inputs. Further, by modeling an output-based model, we maintain the objective of increasing coverage while assisting decision-makers determining optimal operating processes. In the next section of the paper, we describe the models employed here. A description of the data and the results are presented in Section 3. Section 4 concludes the paper with discussion and policy implications.

2. Methods

To illustrate our methodology, we use vaccination centers operating in Bangladesh in 1999. Specifically, we employ data envelopment analysis (DEA), as described by Farrell (1957) and Färe, Grosskopf and Lovell (1994), which is particularly relevant for this analysis because of its ability to employ multiple inputs and outputs and does not require an a priori specification of a cost or profit function. This work, however, differs from typical DEA studies in that we do not formally address inefficiency *per se*. The reason for this approach is because deviations from the frontier may be caused by *both* managerial errors (a typical reason for inefficiency) and the stochastic nature of demand. Since scheduling appointments is not feasible for the population in our sample, because they do not, in general, own clocks or watches nor do they have reliable sources of transportation, we can only address the issue of efficiency subject to time/availability constraints while maintaining as much flexibility as possible. Further we demonstrate an alternative application of the DEA framework where the objectives are not strictly maximizing outputs or minimizing inputs.

By using this approach, we also address the policy proposal of finding the optimal frequency of sessions and hours/session of immunization clinics as suggested in the "Expanded Program on Immunization National Plan of Action 2001–2005". This proposal has been put forward due to the potential high costs due to wastage.

As stated above, we are more interested in applying the DEA methodology in order to assess optimal scheduling rather than an efficiency study of productive technologies. Therefore, rather than using input costs to construct an iso-cost line we use time in order to impose an "iso-time" constraint. Before, however, specifying the relevant DEA models to be solved, we describe the microeconomic underpinnings to our approach.

In figure 1, we illustrate the impact of the possible trade-off between the number of session and the time per session on efficiency as evaluated via DEA. As we opt for output oriented models, we present the analysis in the output space. Here we consider two outputs Y1 and Y2 that define, for example, two types of vaccines. We also consider three observations, a, b and c, with different schedules of sessions defined by the length per session (L) and the number of sessions (s) and different levels of output but with equal total hours of operation (Iso-time $= L * s$) and with equal use of other relevant inputs (x^0). In a traditional analysis without reallocation among inputs (our first DEA model), each of the three observations lies in separate output production sets ($P(x^0, L, s)$) since they differ in the scheduling

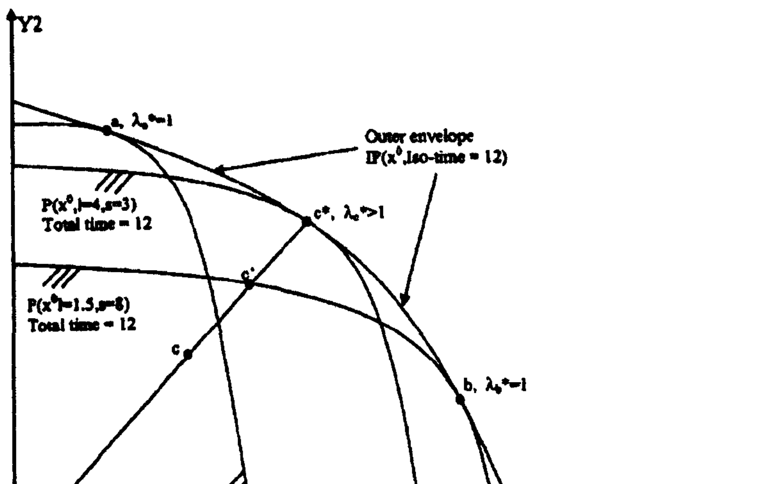


Figure 1. Inclusion of the trade off between the number of sessions and the time per session in DEA evaluation.

of sessions. We consider here two technically efficient observations (a and b) lying on the boundary of their output production sets and one inefficient observation c lying in the interior of its output production set. Without any trade-off between the length and the number of sessions, the technical inefficiency of observation c is measured as usual as the ratio of the two distances $0c'/0c$. Considering now the trade off possibility between the length of session and the number of sessions (our second DEA model), production plans are gauged against a new output production set ($PR(x^0, L, s)$)[4] defined by the iso-time input, which indicates all the possible schedules with a total time of 12 hours. The frontier of this new production set is simply defined as the outer envelop of all initial production sets with the same total hours of operations. Clearly, observations a and b are still efficient since they also lie on the frontier of the new production set. Hence, they have an optimal scheduling of their sessions and no trade-off between the length of session and the number of sessions is necessary to maintain their technical efficiency. This trade-off will be formally defined in our models by the λ variable that indicates simultaneously the possible increase in the length of session and the possible decrease in the number of sessions in order to maintain the total hours of operations constant: $(\lambda * L) * ((1/\lambda) * s)$ = iso-time. Since no reallocation is needed for observations a and b, we have the optimal $\lambda_a^* = 1$ and $\lambda_b^* = 1$. On the contrary, even if observation c removes its technical inefficiency by reaching the initial frontier at c', it will still not be at optimal practice under feasible reallocations between the length of session and the number of sessions. Observation c^* produces more of the two outputs while maintaining the same total hours of operations and all other inputs constant. Hence, the total technical inefficiency under feasible reallocation is measured by the distance $0c^*/0c$. In order to assess the net effect of the reallocation, we define the gain as the ratio of the two technical efficiency measures $(0c^*/0c)/(0c'/0c) = (0c^*/0c')$. It represents the net gain by increasing in the length of sessions (from 1.5 hours to 4 hours) and by decreasing the number of sessions (from 8 to 3). For observation c, we have the optimal $\lambda c^* = 8/3 > 1$.

From this simple illustration, several properties of our models can be highlighted. First, the two models are nested since the frontier of the output production set under feasible reallocation is the outer envelop of the initial frontiers. Therefore, technical inefficiency measures are ordered under both models and the net gain from reallocation is always positive ($(0c^*/0c')$ is always greater than or equal to one in an output DEA framework). Second, some observations may be efficient under both models (as a and b in our illustration) and they are situated at tangency points of the frontiers under both models. A simple analogy is to that of efficient observations under both constant and variable returns to scale in traditional DEA models. Third, although it is not the case in our illustration, the outer envelop can include parts of the initial frontiers but not only at one point. Therefore, the optimal value of the λ variable may not be unique and a post-optimal analysis is required to compute lower and upper bounds on the feasible trade-off.[5]

Next we formally specify the two DEA models. The first model is given by:

$$\max_{h_{j'}, z_j} \; h_{j'}$$

s.t.

$$\sum_{j=1}^{J} z_j \; vaccine_{mj} \geq h_{j'} \cdot vaccine_{mj'} \quad \forall m = 1, \ldots, M \tag{1}$$

$$\sum_{j=1}^{J} z_j\ waste_{mj} \le waste_{mj'} \quad \forall m = 1, \dots, M \tag{2}$$

$$\sum_{j=1}^{J} z_j\ time_j \le time_{j'} \tag{3}$$

$$\sum_{j=1}^{J} z_j\ sessions_j \le sessions_{j'} \tag{4}$$

$$\sum_{j=1}^{J} z_j\ other\ inputs_{nj} \le other\ inputs_{nj'} \quad \forall n = 1, \dots, N \tag{5}$$

$$\sum_{j=1}^{J} z_j = 1 \tag{6}$$

Where J is the number of observations in the sample, M is the number by type of vaccines provided and N is the number of others relevant inputs. The z's are the intensity variables that construct the convex production frontier. Intuitively, for an observation j', the model seeks a referent observation, constructed as a convex combination of all observed observations via the z's variables, ensuring that this reference provides the same or more of each of the M vaccines (constraint 1). This must be accomplished while incurring the same or less wastage on the M vaccines (constraint 2), using the same or less length of time per session and the same or fewer number of sessions (constraints 3 and 4) while consuming the same or less of all others inputs (constraint 5). The referent observation is the one that satisfies all these constraints and that maximizes the provision of all vaccines compared to the evaluated observation j'. The measure is given by h'_j, which indicates that for the observation j' the maximal percentage of additional vaccines that could be made while maintaining all the inputs constant. We note that in this model, time and sessions cannot be substituted for each other.

In the second model, we change the constraints on these two time inputs so that we must now solve:

$$\max_{h_{j'}, z_j, \lambda} \quad h'_{j'}$$

s.t.

$$\sum_{j=1}^{J} z_j\ vaccine_{mj} \ge h'_{j'} \cdot vaccine_{mj'} \quad \forall m = 1, \dots, M \tag{1'}$$

$$\sum_{j=1}^{J} z_j\ waste_{mj} \le waste_{mj'} \quad \forall m = 1, \dots, M \tag{2'}$$

$$\sum_{j=1}^{J} z_j\ time_j \le \lambda \cdot time_{j'} \tag{3'}$$

$$\sum_{j=1}^{J} z_j\ sessions_j \le \frac{1}{\lambda} \cdot sessions_{j'} \tag{4'}$$

$$\sum_{j=1}^{J} z_j \textit{ other inputs}_{nj} \leq \textit{other inputs}_{nj'} \quad \forall n = 1, \ldots, N \tag{5'}$$

$$\sum_{j=1}^{J} z_j = 1 \tag{6'}$$

The "λ" variable in both time and sessions constraints 3′ and 4′, forces these two inputs to be direct inverses of each other, therefore total time is held fixed for each clinic in our sample. This allows us to construct the iso-time curve and analyze each clinic's scheduling practice holding both total time and technology fixed. Note that the two models are nested since the first model is included in the second model, with λ forced to 1. Thus, it is possible to have either $h_j \leq h'_{j'}$ or the ratio $h'_{j'}/h_{j'} \geq 1$. In both models, $h_{j'}$ or $h'_{j'}$ gives the total inefficiency of the clinic evaluated, hence the $h'_{j'}/h_{j'}$ ratio provides the net gains that can be achieved if scheduling were to be optimized while controlling for inefficiency. This is the main strength of our modeling approach—while we cannot clearly interpret the productive inefficiency measured (specifically if it is due to managerial non-performance or stochastic demand effects), we cannot ignore it. Rather, we evaluate the possible efficiency gains from an optimal schedule of sessions controlling for production inefficiency. In other words, the h'/h ratio gives the percentage of additional vaccine allowed by an optimal schedule of the sessions even if the clinic is not fully efficient. We stress that this is a relative measure similar to the usual relativity of measures given in DEA. The key to our analyses therefore is the ratio of the two scores rather than one measure or the other.

The λ variable can be either greater than, equal to or less than 1. The interpretation of this variable is that if λ is less than one there are too few sessions but too many hours per session. The converse is true if λ is greater than one (too many sessions but too few hours per session). If λ is equal to one, then there is an optimal repartitioning of the number of sessions and the hours per sessions (independently of whether the clinic is globally efficient or inefficient). Note that since λ is unconstrained (except a natural positivity constraint), it may lead from a theoretical point of view to any scheduling that may be irrealistic or simply not feasible. Hence, time restrictions seem to be required to ensure that the length of sessions times the number of sessions is belong to some bounds (e.g. duration *session < 168 hours a week). Nevertheless, it is not necessary in our context since we opt for variable returns to scale DEA models. Therefore, the constraint on activity variables i.e. $\sum_{j=1}^{J} z_j = 1$ ensures that the optimal values for the reference set associated to the time and session constraints are linear combinations of observed values among the sample observations. So, we avoid possible infinite extensions of observed production plans as in constant returns to scale models for example. Thus, it is not necessary to include bounds on the total time since optimal values are constrained to be lower (at most equal) to observed values in the sample.

Finally, we stress two technical points that are important to implement these models. First, note that in the second model with the λ variable leads to a nonlinear program. Whereas some problems for the estimation may arise, the nonlinearities are not excessive and recent programming solvers handle them easily. Second, as stated above, multiple solutions for the λ variable may arise in program 2 leading to the same optimal solution. A post-optimal analysis is required to compute lower and upper limits on the λ variable. It simply involves two additional programs with a new objective function that seeks to maximise and minimize λ respectively and by including a new constraint forcing the h' variable to be equal to its optimal value h'^*. In case of multiple solutions, we adopt the 'philosophy' of the DEA approach which always evaluates an observation under its best possible light and we keep the nearest solution from the initial scheduling of evaluated vaccine's sites.

3. Data and Results

A cost-effectiveness analysis of measles control has been undertaken in DCC (Walker et al., 2000). This study was based on a stratification of health centers by zone and type of site (fixed or satellite (outreach)), from which 132 sites were randomly selected representing 25% of all the EPI delivery sites. We further selected only clinics for our final sample that did not have missing data for inputs, outputs, or time related variables. The final sample size we study here consists of 117 clinics. These included 35 (30%) fixed and 82 (70%) outreach delivery sites. Of the fixed sites, 19 (54.3%) were operated by the government and 16 (45.7%) by Non-Governmental Organizations (NGOs). Of the satellite sites, 16 (17.4%) were operated by the government and 76 (82.6%) by NGOs. Because this earlier study also collected data on the total costs of delivering EPI activities and site- and antigen-specific wastage rates, it provided an opportunity to assess the role of scheduling on the efficiency of EPI provision and therefore forms the basis of the data analyzed in this paper.

Inputs are defined as the number of full time equivalent medical staff, size of the facility dedicated to the EPI (in squared meters), the number of hours of operation, and the number of sessions. The level of wastage by vaccine type is also given as a constraint. Outputs are defined as the number and type of five vaccines aimed at children less than 5 years of age and pregnant women. Table 1 contains the descriptive statistics of inputs and outputs.

We next turn to the results on the possible efficiency gains and the observed and optimal scheduling of clinic hours which are presented in Table 2.

The results presented in Table 2 illustrate that 60% of the clinics operate with an optimal scheduling of vaccination sessions. The rest of the clinics are equally split between two groups resulting in 20% of the total number of clinics. The post optimal analysis reveals that 16% of observations have multiple solutions for the λ variable. As stated above, we keep in this case, the closest value compared to the initial scheduling of sessions. The first group characterized by $\lambda < 1$ may be interpreted as clinics with too long but too few sessions per month. The second group with $\lambda > 1$ is characterized by clinics with too short but too many sessions per month. Within these two groups, clinics not employing optimal scheduling may increase the total number of vaccines by 10% for those with $\lambda < 1$ and by 19% for clinics receiving scores of $\lambda > 1$, which is statistically significantly different from clinics operating with relatively optimal schedules ($\lambda = 1$). Here, we stress again that these potential gains are solely due to the scheduling of sessions once we have controlled for other sources of inefficiencies. The wastage of vaccines incurred by the non-optimizing clinics is also statistically significantly greater than the clinics operating with optimal scheduling.

Table 1. Descriptive statistics of inputs and outputs.

Variable	Mean	Std. Dev.	Minimum	Maximum
INPUTS				
Labor	3.11	3.47	1	20
Facility size dedicated to EPI	368.53	406.81	30	2.700
BCG waste	126.12	142.09	0.50	701.46
DPT waste	147.75	141.42	0.50	746.66
OPV waste	215.49	263.72	0.00	1795.50
Measles waste	90.22	100.89	0.50	465.91
TT waste	140.57	133.05	0.50	515.88
Monthly sessions	6.30	5.33	1	24
Hours/session	4.14	1.44	1.5	8
OUTPUTS				
Number of BCG vaccines	259.59	305.31	1	1680
Number of DPT vaccines	581.67	687.67	1	3264
Number of OPV vaccines	710.70	845.78	1	3756
Number of measles vaccine	190.98	211.60	1	960
Number of TT vaccines	389.87	445.27	1	2208

Table 2. Results by optimal scheduling of sessions.

	$\lambda < 1$	$\lambda = 1$	$\lambda > 1$	K-W tests[a]
Number of clinics	23	71	23	
%	20%	60%	20%	
Efficiency gains	10%	0%	19%	***
(% of total vaccines)				
Wastes	49%	28%	44%	***
(% of total vaccines)				
Number of hours/session				
Observed	4.70	4.18	3.47	***
Optimal	4.05	4.18	4.19	ns
Difference	0.65	0.00	−0.72	***
Number of sessions/month				
Observed	3.83	6.70	7.52	**
Optimal	4.41	6.70	6.06	ns
Difference	−0.58	0.00	+1.46	***

[a]K-W stand for the Kruskal-Wallis statistics to test if several samples come from the same population.
** and *** means significant at the 5% level and the 1% level respectively.

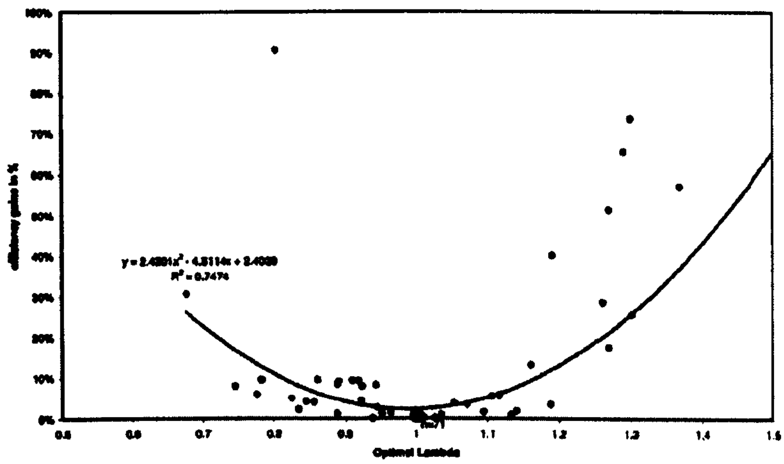


Figure 2. Efficiency gains and optimal trade off for sessions' scheduling.

Note however that even optimal clinics incur wastage, but their rates are significantly lower.

Clinics classified as having a $\lambda < 1$ operate 16% more hours per session but 13% fewer sessions per month than the optimal level. This relationship is just the converse for those clinics categorized as having a $\lambda > 1$: they have to increase the time of session by 17% while decreasing the number of sessions by 24%. Further the differences between observed and optimal time is statistically significantly different among the clinics based on scheduling, but there is no statistically significant difference for optimal operations vis-à-vis scheduling. We also provide a graphical depiction of the relationship between λ and efficiency gains in percent. This is given in figure 2. We observe a positive relationship between the optimal scheduling and the efficiency of clinics. The U-shape comes from the ordering of the optimal λ: as the value of λ is less than unity and decreases, clinics can increasingly gain more with the same dynamic occurring for an increasing λ for clinics with a $\lambda > 1$. However, we note that efficiency gains are not symetric in the two cases and gains would be higher for clinics reducing the number of sessions ($\lambda > 1$).

We are also interested in organizational features such as private non-governmental organizations (NGOs) which in this sample are non-profit. Another organizational aspect of these clinics that may lead to less than optimal scheduling would be whether they were fixed or satellite clinics. In Table 3 we present the main characteristics of clinics along their ownership and their type. In Table 4 we present the mean results of relative gains, scheduling partitioning (λ) and differences in scheduling the time and number of sessions by organizational form.

From an organizational perspective we are also interested if information can be gleaned by assessing whether the combination of ownership and type of clinic (fixed versus satellite)

Table 3. Activity by ownership and type of clinics.

Organizational Variable	N	Total Vaccines (average)	No. of Hours	No. of Sessions	Total Wastage
NGO/fixed	16	3305	5.9	11.3	32%
Government/fixed	19	4077	3.9	13.9	36%
NGO/satellite	76	1342	3.9	3.5	33%
Government/satellite	6	2867	3.5	4.0	35%

Table 4. Results by ownership and type of clinics.

Organizational Variable	N	Efficiency Gains	Mean λ	Dif. in Hours	Dif. in Sessions
NGO/fixed	16	2.2%	1.04	−0.12	0.21
Government/fixed	19	8.7%	1.15	−0.41	1.21
NGO/satellite	76	3.8%	0.98	0.12	−0.10
Government/satellite	6	2.4%	1.08	−0.23	0.23
K-W tests*		ns	***	***	***

*K-W stand for the Kruskal-Wallis statistics to test if several samples come from the same population.
** and *** means significant at the 5% level and the 1% level respectively.

also perform differently in terms of the scheduling of vaccinations. From Table 3, we note that globally fixed clinics perform twice the number of vaccinations of satellite centers. While there is no significant difference by amount of total vaccines provided by fixed clinics, non-governmental satellite centers perform half as many vaccinations than governmental centers. Turning next to the scheduling issue, we note no significant differences among satellites centers even though non-governmental fixed clinics have longer sessions. In terms of wastage, there is no statistically significant difference among all types of clinics.

From Table 4, we observe that on average all four organizational forms could have achieved gains in efficiency from more optimal scheduling but there is no global significant differences among them. Government/fixed clinics currently are the least efficient (relatively) to the other three organizational forms since they could gain the most from moving from its observed position to an optimal position (8.7%). Observed NGO clinics perform relatively better than the other form in the scheduling of their sessions ($\lambda = 1.04$ and $\lambda = 0.98$ for fixed and satellite respectively). Again, the most important improvement in scheduling is related to governmental fixed clinics. On average, they have to reduce the number of sessions by 1.21 by month and to increase the length of sessions.

4. Discussion

The purpose of this paper was to assess if the 117 vaccination clinics operating in Dhaka, Bangladesh meet the objectives of maximizing vaccines produced. We accomplished this objective by employing DEA techniques to two separate models–one without constraints on the two time elements (hours per session and number of sessions) and another with time constraints. By taking the ratio of the resulting scores by clinic, we are able to determine what gains could be made in terms of increasing the number of vaccines provided if clinics had been operating with an optimal schedule. We avoid making any statements regarding firm inefficiency, in a productive sense, since we hypothesize that the distance a clinic is from the frontier may be due to the stochastic nature of the demand for these clinic services and not just managerial inefficiency.

The underlying reason for pursuing this line of inquiry is because the Government of Bangladesh has recommended that one way to reduce waste would be to reduce some vaccination sessions. In a similar vein to this stringent policy, we found that optimality could be attained if in some cases, *both* number of sessions and hours of operation per session were altered.

To summarize our results, we found that optimality of scheduling was, on average, around seven sessions with each session lasting four hours per month. If optimality had been met, gains (i.e. amount and type of vaccines provided) could have been achieved from between 10 to 20%. In other words, the clinics not operating with optimal schedules could increase the number of vaccines provided while reducing the waste incurred. However, this is not to say that zero waste is a possibility since some waste is inevitable. We found that optimally operated clinics did incur waste (28%) but this rate was significantly lower than clinics that did not operate optimal schedules. It was also demonstrated that the relationship between λ and efficiency gains exhibited a type of scale diseconomy with clinics that were opened for more sessions but for fewer hours per session than was optimal, incurred the higher degree of inefficiency. Analyzing our results by organizational form of the clinics, we further found that government/fixed clinics tended to have too many sessions, which may be one reason for the focus the GoB policy. However, we did find that fixed site clinics, especially those that are publicly owned provided more total vaccines on average without statistically significantly more wastage. This may also be due to the attenuation of property rights in governmental clinics, whereas they may be less concerned with efficiency of scheduling and more concerned with overall access and provision of vaccines. This hypothesis is borne out further since government/fixed clinics provided more total vaccines than the other three organizational types. In addition, NGO clinics were more inclined to operate at more optimal scheduling than governmental clinics for both type of centers (fixed and satellite). One reason may be that despite the non-profit status of the private clinics in our sample and the stated objective of maximizing vaccines provided, they may be more closely monitored by the relative funding agencies.

Whereas we focused on what could be termed avoidable waste, the WHO's revised policy statement on "The use of opened multi-dose vials of vaccine in subsequent immunization sessions" suggests that the revised policy has the potential to reduce vaccine wastage rates by up to 30%, resulting in annual savings worldwide of $40 million (US) in vaccine costs (WHO, 2000). Note however that this figure represents current use of vaccines which are not as expensive as future vaccines for the prevention of diseases such as *Hepatitis B* and *Haemophilus Influenzae Type B (HIB)* in which cases waste of these vaccines would incur higher costs. Further research combining scheduling as well as an open vial policy is the next natural step towards addressing the GoB's concerns of vaccine waste.

Acknowledgments

Vivian Valdmanis and Damian Walker are members of the Health Economics and Financing Program, which is supported by funds from the UK Department for International Development.

Notes

1. In a search using the Medline search engine and the Sigle search engine for "gray literature" we only found one article that was applicable to our research question.
2. Vaccine wastage is the proportion of vaccine supplied, but not administered to children, usually stated as a rate and is calculated as: vaccine wastage rate = ([doses supplied − doses administered]/doses supplied) * 100.
3. Open vials can be open and shut and therefore reusable for longer periods of time.
4. As usual, $P(x)$ stands for a traditional output production set while in our context $RP(x)$ stands for the output production set with feasible reallocations among parts of the inputs.
5. We thank an anonymous referee for pointing out this relevant point.

References

England, S. et al. (2001). The Evidence Base for Interventions to Reduce Mortality from Vaccine-Preventable Diseases in Low and Middle-Income Countries. CMH Working Paper Series. Vol. Paper No. WG5: 10.

Expanded Program on Immunization National Plan of action 2001–2005. Government of Bangladesh.

Färe, R., S. Grosskopf and C. A. K. Lovell. (1994). *Production Frontiers*. Cambridge: Cambridge University Press.

Farrell, M. J. (1957). "The Measurement of Productive Efficiency." *Journal of the Royal Statistical Society*, Series A (General) 120, 253–281.

Rappouli, R, H.I. Miller and S. Falkow. (2002). "The Intangible Value of Vaccination." *Science* 297, 937–939.

Walker, D., S. Khan, S. M. Akramuzzaman et al. (2000). *Cost-Effectiveness of Measles Control in Urban Dhaka, Bangladesh*. Geneva: Final Report to WHO.

World Bank. (1993). *World Development Report 1993. Investing in Health*. Oxford: Oxford University Press.

World Health Organization. Expanded Programme on Immunization (1999). WHO-UNICEF policy statement on the use of vaccine vial monitors in immunization services. Geneva: WHO, (document WHO/V&B/99.18); www.who.int/vaccines-documents/DocsPDF99/www9928.pdf

World Health Organization. (2000). WHO Policy Statement: The Use of Opened Multi-Dose Vials of Vaccine in Subsequent Immunization Sessions. Geneva: Department of Vaccines and Biologicals, WHO, (document WHO/V&B/00.09); www.who.int/vaccines-documents/DocsPDF99/www9924.pdf

B. Dervaux is a health economist and department head of CRESGE as well as a member of CNRS.

H. Leleu is a health economist, also with CNRS.

V. Valdmanis is a lecturer in health economics at the London School of Hygiene and Tropical Medicine.

D. Walker is a research fellow and doctoral student in health economics at the London School of Hygiene and Tropical Medicine.

SFA software

Stochastic frontiers can be estimated using a different range of multi-purpose econometric software which can be adapted for the desired estimation. This software includes well-known statistical packages such as SPSS, Shazam, GAUSS and SAS. However, the two most commonly used packages for estimating stochastic production frontiers and inefficiency are FRONTIER 4.1 (Coelli 1996b) and LIMDEP (Greene 1995). A more detailed review of both packages is provided by Sena (1999).

FRONTIER 4.1 is a single purpose package specifically designed for the estimation of stochastic production frontiers (and nothing else), while LIMDEP is a more general package designed for a range of non-standard, i.e. non-OLS, econometric estimation. An advantage of the former package is that estimates of efficiency are produced as a direct output from the package. The user is able to specify the distributional assumptions for the estimation of the inefficiency term in a programme control file. In LIMDEP, the package estimates a one-sided distribution, but the separation of the inefficiency term from the random error component requires additional programming.

FRONTIER is able to accommodate a wider range of assumptions about the error term than LIMDEP (see Table 1), although it is unable to model exponential distributions. Neither package can include gamma distributions. Only FRONTIER is able to estimate an inefficiency model as a one-step process. An inefficiency model can be estimated in a two-way process using LIMDEP. However, this may create bias as the distribution of the inefficiency estimates is pre-determined through the distributional assumptions used in its generation.

Table 1: Distributional assumptions allowed in the software

Distribution	FRONTIER	LIMDEP
Time invariant firm specific inefficiency		
Half-normal distribution	✓	✓
Truncated normal distribution	✓	✓
Exponential distribution	✗	✓
Time variant firm specific distribution		
Half-normal distribution	✓	✗
Truncated normal distribution	✓	✗
One-step inefficiency model	✓	✗

DEA software

Most of the general-purpose mathematical optimisation software can be adapted to solve DEA problems. Examples of programme code for DEA models have already been published and are readily adaptable, e.g. Olesen and Petersen (1995) present the GAMS code for a DEA model that can be adapted to suit most analyses. Emrouznejad (2000) developed a SAS programme for different DEA models, including options for input- and output-orientation orientated CRS and VRS models. The Emrouznejad (2000) programme can be downloaded from http://deazone.com/software/index.htm#sasdea

These general programmes offer the possibility of a wide range of applications using non-specialist DEA software. However, there are several DEA-specific programmes that provide a variety of interesting facilities. Seven of the most common ones are listed below:

1. DEAP 2.1 (Coelli 1996a);
2. DEA-Solver Professional 4.0;

3. EMS 1.3;
4. Frontier Analyst 3.1.5;
5. IDEAS 6.1;
6. OnFront 2.02;
7. Warwick DEA 1.0;

The key features of the different software packages are summarised in Table 2.

Table 2: Key features of the packages

Package	Malmquist index	Weight restrictions	Input / output orientation	Multi-stage DEA	CRS / VRS	NIRS
DEA Solver		✓	✓		✓	✓
DEAP	✓		✓	✓	✓	
EMS	✓	✓	✓		✓	✓
Frontier Analyst		✓	✓		✓	
IDEAS			✓		✓	
OnFront	✓		✓		✓	✓
Warwick DEA	✓*	✓	✓		✓	

* This can be carried out by doing some modifications to the programme

Of the packages described above, EMS and DEAP are free.

References

Barr RS (2004) DEA Software Tools and Technology: A State-of-the-Art Survey. In Cooper WW, Seiford LM, Zhu J (Eds) *Handbook on Data Envelopment Analysis* Boston: Kluwer Academic Publishers.

Coelli TJ (1996a) *A Guide to DEAP Version 2.1: A Data Envelopment Analysis Computer Program*, CEPA Working Paper No. 8/96, Department of Econometrics, University of New England.

Coelli TJ (1996b) *A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation.* CEPA Working Paper No. 7/96, Department of Econometrics, University of New England.

Emrouznejad A (2000) An extension to SAS/OR for decision system support. Indianapolis: Proceedings of the SAS User Group International Conference.

Greene W (1995) *LIMDEP (Version 7): User's Manual and Reference Guide.* New York: Econometric Software Inc.

Hollingsworth B (1997) A review of data envelopment analysis software. *The Economic Journal* 107(443): 1268-1270.

Olesen O, Petersen N (1996) A Presentation of GAMS for DEA, *Computers and Operations Research* 23(4): 323-339.

Sena V (1999) Stochastic frontier estimation: a review of the software options. *Journal of Applied Econometrics* 14(5): 579-586.

Appendix 5: Facility survey form for vaccination delivery units

Date:
Name of interviewer:
Name of respondent, designation:
Name of facility, address:

1. Type of vaccination delivery unit:

Key: more than one day per week = 1, one day per week = 2, other (specify) = 3

2. How many days/sessions held per month?

3. How many hours is one session?

4. Address of EPI site:

    a. Zone

    b. Ward

5. From which year routine EPI services started in this delivery site:

6. Activities, other than routine vaccination services, delivered at this unit:

7. Which organization manages this EPI unit:

8. Which organization or agency funds (monetary and . or logistical) this EPI unit:

9. Building and construction:

    a. What kind of facilities are there for immunization services:
    Key: room = 1, corridor = 2, veranda = 3, other (specify) = 4

    b. Type of construction material:
    Wall:

    Floor:

    Roof

    c. What is the size of this facility (square feet)?
    Length:

    Width:

    Total:

    d. What is the size of the area used just for routine EPI (square feet)?
    Length:

    Width:

10. Vehicles:

Type	Quantity	Unit cost	% of use for routine EPI
Jeep			
Pickup			
Bicycle			
Motorcycle			
Other (specify)			
Other (specify)			

11. Travel:

Activity	Collection of vaccine stock	EPI meetings	Social mobilisation for EPI	Other activities (specify)
Number of trips per month				
Distance per trip (km)				
How? Key: bus=1, baby-taxi=2, unit's vehicle=3, on foot=4,bicycle=5, motorcycle=6, other (specify)=7				
Expenditure per round trip				
Duration per round trip				

12. Equipment:

Type	Quantity	Unit cost	% of use for routine EPI
Refrigerator			
Cold box			
Steam steriliser			
Vaccine carrier			
Other (specify)			
Other (specify)			

13. Supplies used for vaccination activities:

Type	Quantity	Unit cost
Brush		
Card (child)		
Card (mother)		
Carry bag		
Cotton (roll)		
Day carrier		
Dial thermometer		
Duster		
Icepack 0.4 litre		
Icepack 0.6 litre		
Indent form book		
Mixing syringe 5ml		
Moni flag		
Monthly report book		
Needle 18 gauge (box of 12)		
Needle 23 gauge (box of 12)		
Needle 26 gauge (box of 12)		
Pamphlet		
Plastic bowl		
Poster		
Register book (child)		
Register book (mother)		
Soap		
Soap box		
Steriliser bag		
Syringe 0.05ml (box of 10)		
Syringe 0.5ml (box of 10)		
Syringe 1ml (box of 10)		
Table cover		
Tally sheet		
Timer		
Others(specify)		

14. Furniture:

Type	Quantity	Unit cost	% of use for routine EPI
Table			
Chair			
Bench			
Other (specify)			
Other (specify)			

15. Staff:

a. Salaries

Type	Quantity	Salary (including all benefits)	% of time for n routine EPI
Doctor			
Paramedic			
Health assistant			
Nurse			
Counsellor			
Vaccinator			
Health educator			
Health worker			
Driver			
Cleaner			
Guard			
Other (specify)			
Other (specify)			

b. Training

Type	Long-term			Short-term		
	Number of type of staff	Number of sessions	Type of training	Number of type of staff	Number of sessions	Type of training
Doctor						
Paramedic						
Health assistant						
Nurse						
Counsellor						
Vaccinator						
Health educator						
Health worker						
Other (specify)						

16. Expenditure for the following items:

Item	Amount
Gas	
Electricity	
Water	
Telephone	
Postage	
Printing	
Repairs and maintenance	
Other (specify)	
Other (specify)	

17. Vaccines:

State the range of vaccines provided at this site (OPV, DPT, BCG, measles, other (specify)	State the number of doses in each vial of the vaccines	State the total number of doses administered of each vaccine at this site	State the number of vials of each vaccine in stock at the beginning of the July 1998	State the number of vials of each vaccine supplied with between July 1998 – June 1999	State the number of vials of each vaccine at the end of June 1999
BCG					
DPT					
OPV					
Measles					
TT					
Other (specify)					
Other (specify)					

Appendix 6: Location, type of ownership and type of vaccination delivery unit

Serial #	Location (zone)	Type of ownership (GoB or NGO)	Type of vaccination delivery unit (fixed or outreach)
1	1	NGO	fixed
2	1	NGO	outreach
3	1	NGO	outreach
4	1	GoB	fixed
5	1	NGO	outreach
6	1	NGO	outreach
7	1	GoB	fixed
8	1	NGO	outreach
9	1	NGO	fixed
10	2	GoB	outreach
11	2	GoB	fixed
12	2	GoB	fixed
13	2	GoB	fixed
14	2	GoB	fixed
15	2	GoB	fixed
16	3	GoB	fixed
17	3	NGO	outreach
18	3	GoB	fixed
19	3	NGO	outreach
20	4	NGO	outreach
21	4	NGO	outreach
22	4	GoB	fixed
23	4	NGO	fixed
24	4	NGO	outreach
25	4	NGO	outreach
26	4	NGO	fixed
27	4	NGO	fixed
28	4	NGO	outreach
29	4	NGO	fixed
30	4	NGO	outreach
31	4	NGO	outreach
32	4	NGO	outreach
33	4	NGO	outreach
34	4	NGO	outreach
35	4	NGO	outreach
36	5	NGO	outreach
37	5	GoB	fixed
38	5	GoB	fixed
39	5	NGO	fixed
40	5	GoB	fixed
41	5	NGO	fixed
42	5	GoB	fixed
43	5	NGO	outreach
44	5	NGO	outreach
45	5	NGO	outreach
46	6	NGO	outreach
47	6	NGO	outreach
48	6	NGO	outreach
49	6	NGO	fixed
50	6	NGO	outreach
51	6	NGO	outreach
52	6	NGO	fixed
53	6	NGO	outreach
54	7	NGO	outreach
55	7	NGO	outreach

56	7	NGO	fixed
57	7	NGO	fixed
58	7	NGO	outreach
59	7	NGO	outreach
60	7	NGO	outreach
61	7	GoB	outreach
62	7	NGO	outreach
63	7	NGO	outreach
64	7	NGO	outreach
65	7	NGO	outreach
66	7	NGO	outreach
67	7	NGO	outreach
68	7	NGO	outreach
69	7	NGO	outreach
70	7	NGO	outreach
71	8	NGO	outreach
72	8	NGO	outreach
73	8	NGO	outreach
74	8	NGO	outreach
75	8	NGO	outreach
76	8	NGO	outreach
77	8	NGO	outreach
78	8	NGO	outreach
79	8	GoB	fixed
80	8	NGO	outreach
81	8	NGO	outreach
82	8	NGO	fixed
83	8	NGO	fixed
84	8	GoB	outreach
85	8	NGO	outreach
86	9	NGO	outreach
87	9	NGO	outreach
88	9	NGO	outreach
89	9	NGO	outreach
90	9	NGO	outreach
91	9	NGO	outreach
92	9	NGO	outreach
93	9	NGO	fixed
94	9	NGO	outreach
95	9	NGO	outreach
96	9	NGO	outreach
97	9	NGO	outreach
98	9	NGO	outreach
99	9	NGO	outreach
100	9	NGO	outreach
101	9	GoB	outreach
102	10	NGO	outreach
103	10	NGO	outreach
104	10	NGO	outreach
105	10	GoB	fixed
106	1	NGO	outreach
107	1	GoB	fixed
108	9	NGO	outreach
109	9	GoB	outreach
110	9	NGO	outreach

Appendix 7: Data for model DEA1

Serial #	BCG	DPT	OPV	Measles	TT	Total cost
1	324	1104	1404	300	1884	2382
2	72	168	192	24	36	1085
3	216	612	708	120	372	1572
4	900	984	1704	720	720	3981
5	348	360	2400	300	600	954
6	360	384	2400	324	648	918
7	456	1380	1380	336	384	2965
8	108	168	168	84	396	924
9	240	1164	1164	300	1548	970
10	972	3264	3420	792	1248	4063
11	228	240	240	72	144	1668
12	444	1644	1644	504	924	3389
13	84	816	1296	480	552	3377
14	48	804	696	132	384	2117
15	432	1896	2460	696	876	3251
16	480	1200	1500	480	600	1901
17	804	3012	3444	432	432	6532
18	1680	1800	2400	960	960	1561
19	576	3036	3756	576	1308	4398
20	24	48	60	12	132	474
21	396	804	996	324	612	1207
22	156	300	300	60	264	2588
23	804	900	516	516	960	3796
24	72	588	744	156	2208	1730
25	132	468	540	96	216	702
26	120	360	360	60	420	566
27	252	888	924	36	156	4546
28	1	216	252	36	192	747
29	912	2616	2736	312	1308	2779
30	192	252	312	144	396	531
31	960	720	840	240	360	829
32	1	36	72	36	12	238
33	60	84	84	24	48	448
34	132	360	360	108	324	1102
35	36	36	84	48	972	519
36	600	1800	2184	384	612	2659
37	576	1608	1608	540	816	12022
38	600	1404	1884	432	564	15077
39	84	348	552	204	216	6313
40	864	1788	2520	636	1164	7172
41	720	1860	1860	360	1920	2271
42	540	744	744	636	552	3166
43	216	312	312	204	384	906
44	72	156	156	1	192	716
45	216	192	192	108	312	928
46	180	516	684	168	252	2793
47	36	84	84	1	168	1204
48	1	60	60	60	12	833
49	84	144	180	36	48	1026
50	24	60	120	36	120	1651
51	24	108	132	24	48	1030
52	180	516	684	168	252	2793
53	60	108	168	60	108	1094
54	60	84	108	24	36	540
55	84	324	420	96	84	648
56	456	1176	1332	156	600	5357

57	36	180	240	24	72	4347
58	48	48	60	12	72	434
59	48	108	168	48	120	519
60	192	336	336	84	36	2217
61	108	264	360	96	240	3599
62	336	984	1188	204	480	3086
63	1	216	216	12	108	708
64	144	420	588	168	240	677
65	84	240	420	96	96	560
66	84	180	312	48	48	577
67	60	144	228	48	12	375
68	36	84	84	1	24	359
69	60	276	276	48	60	561
70	900	432	420	360	420	672
71	120	144	144	36	204	441
72	60	204	264	60	120	1382
73	324	660	756	300	600	1261
74	24	60	96	36	36	483
75	96	240	240	120	240	1260
76	1	168	168	1	60	2135
77	48	144	144	24	72	1177
78	1	96	96	12	36	416
79	492	1704	2100	396	768	7087
80	60	132	132	1	12	626
81	180	420	420	180	480	1144
82	396	1476	1752	276	1596	7483
83	300	1560	1560	300	1440	3662
84	60	216	240	240	228	768
85	12	84	84	1	96	485
86	300	360	300	240	180	738
87	72	96	96	12	168	724
88	48	204	204	48	240	759
89	24	132	132	24	216	745
90	1	72	72	1	24	296
91	864	900	840	240	240	1131
92	480	240	240	240	240	912
93	300	828	936	108	288	2100
94	48	48	48	12	360	745
95	840	540	360	360	240	686
96	156	420	420	36	180	9785
97	36	132	156	60	120	1010
98	480	240	240	240	240	1090
99	36	48	48	1	24	325
100	720	300	300	180	300	707
101	336	1512	2028	516	384	3165
102	192	168	408	192	192	2008
103	120	96	240	240	96	1053
104	36	60	96	24	36	911
105	864	1440	2400	960	1020	2422
106	36	156	216	60	24	718
107	228	1200	1500	300	420	2742
108	1	24	48	1	24	582
109	24	168	192	24	252	2532
110	108	324	324	240	252	875

Appendix 8: Data for model DEA2

Serial #	BCG	DPT	OPV	Measles	TT	Labour	Facility	Hours
1	324	1104	1404	300	1884	20	600	154
2	72	168	192	24	36	2	79	79
3	216	612	708	120	372	1	1620	216
4	900	984	1704	720	720	2	252	252
5	348	360	2400	300	600	2	405	405
6	360	384	2400	324	648	1	432	432
7	456	1380	1380	336	384	2	600	450
8	108	168	168	84	396	2	75	75
9	240	1164	1164	300	1548	1	270	270
10	972	3264	3420	792	1248	2	9000	1296
11	228	240	240	72	144	2	13068	225
12	444	1644	1644	504	924	2	3267	324
13	84	816	1296	480	552	2	4500	720
14	48	804	696	132	384	1	990	135
15	432	1896	2460	696	876	2	4019	162
16	480	1200	1500	480	600	1	300	300
17	804	3012	3444	432	432	10	2500	144
18	1680	1800	2400	960	960	1	600	200
19	576	3036	3756	576	1308	10	2000	144
20	24	48	60	12	132	6	113	113
21	396	804	996	324	612	4	120	120
22	156	300	300	60	264	3	6750	164
23	804	900	516	516	960	6	816	144
24	72	588	744	156	2208	5	959	338
25	132	468	540	96	216	2	270	270
26	120	360	360	60	420	1	108	108
27	252	888	924	36	156	13	12500	100
28	1	216	252	36	192	6	864	216
29	912	2616	2736	312	1308	9	450	450
30	192	252	312	144	396	1	216	216
31	960	720	840	240	360	1	9000	135
32	1	36	72	36	12	3	2613	70
33	60	84	84	24	48	5	900	900
34	132	360	360	108	324	3	1350	270
35	36	36	84	48	972	5	7200	360
36	600	1800	2184	384	612	7	7500	900
37	576	1608	1608	540	816	2	3000	500
38	600	1404	1884	432	564	2	4019	375
39	84	348	552	204	216	4	1000	240
40	864	1788	2520	636	1164	1	4019	525
41	720	1860	1860	360	1920	3	4096	1500
42	540	744	744	636	552	2	180	180
43	216	312	312	204	384	1	750	375
44	72	156	156	1	192	1	200	200
45	216	192	192	108	312	1	240	240
46	180	516	684	168	252	1	560	192
47	36	84	84	1	168	1	300	300
48	1	60	60	60	12	1	288	288
49	84	144	180	36	48	1	144	144
50	24	60	120	36	120	1	100	100
51	24	108	132	24	48	1	225	225
52	180	516	684	168	252	6	10800	180
53	60	108	168	60	108	1	225	225
54	60	84	108	24	36	2	94	94
55	84	324	420	96	84	1	180	180
56	456	1176	1332	156	600	8	600	600

57	36	180	240	24	72	10	1000	405
58	48	48	60	12	72	2	324	324
59	48	108	168	48	120	2	252	252
60	192	336	336	84	36	3	405	405
61	108	264	360	96	240	2	980	504
62	336	984	1188	204	480	8	2700	2700
63	1	216	216	12	108	1	535	248
64	144	420	588	168	240	2	843	843
65	84	240	420	96	96	2	300	300
66	84	180	312	48	48	2	843	506
67	60	144	228	48	12	2	162	81
68	36	84	84	1	24	1	4247	33
69	60	276	276	48	60	2	400	400
70	900	432	420	360	420	2	320	320
71	120	144	144	36	204	1	288	288
72	60	204	264	60	120	1	3267	2400
73	324	660	756	300	600	4	315	315
74	24	60	96	36	36	3	160	160
75	96	240	240	120	240	1	1100	270
76	1	168	168	1	60	2	490	270
77	48	144	144	24	72	1	375	375
78	1	96	96	12	36	2	450	200
79	492	1704	2100	396	768	9	3600	1800
80	60	132	132	1	12	1	320	320
81	180	420	420	180	480	4	225	225
82	396	1476	1752	276	1596	7	3375	864
83	300	1560	1560	300	1440	14	1633	270
84	60	216	240	240	228	3	216	216
85	12	84	84	1	96	1	300	300
86	300	360	300	240	180	1	150	150
87	72	96	96	12	168	1	270	270
88	48	204	204	48	240	1	35	35
89	24	132	132	24	216	1	1012	675
90	1	72	72	1	24	1	2700	506
91	864	900	840	240	240	1	216	72
92	480	240	240	240	240	1	7840	225
93	300	828	936	108	288	3	980	64
94	48	48	48	12	360	1	653	326
95	840	540	360	360	240	1	150	60
96	156	420	420	36	180	6	150	30
97	36	132	156	60	120	1	5227	60
98	480	240	240	240	240	1	2613	272
99	36	48	48	1	24	1	9801	100
100	720	300	300	180	300	1	180	180
101	336	1512	2028	516	384	2	1012	675
102	192	168	408	192	192	3	750	225
103	120	96	240	240	96	4	700	700
104	36	60	96	24	36	3	200	200
105	864	1440	2400	960	1020	2	1000	180
106	36	156	216	60	24	2	252	252
107	228	1200	1500	300	420	3	675	675
108	1	24	48	1	24	1	1600	337
109	24	168	192	24	252	2	1125	337
110	108	324	324	240	252	2	2500	150

Appendix 9: Data for model DEA3

Serial #	Total vaccines	Labour	Facility	Hours
1	5016	20	600	154
2	492	2	79	79
3	2028	1	1620	216
4	5028	2	252	252
5	4008	2	405	405
6	4116	1	432	432
7	3936	2	600	450
8	924	2	75	75
9	4416	1	270	270
10	9696	2	9000	1296
11	924	2	13068	225
12	5160	2	3267	324
13	3228	2	4500	720
14	2064	1	990	135
15	6360	2	4019	162
16	4260	1	300	300
17	8124	10	2500	144
18	7800	1	600	200
19	9252	10	2000	144
20	276	6	113	113
21	3132	4	120	120
22	1080	3	6750	164
23	3696	6	816	144
24	3768	5	959	338
25	1452	2	270	270
26	1320	1	108	108
27	2256	13	12500	100
28	697	6	864	216
29	7884	9	450	450
30	1296	1	216	216
31	3120	1	9000	135
32	157	3	2613	70
33	300	5	900	900
34	1284	3	1350	270
35	1176	5	7200	360
36	5580	7	7500	900
37	5148	2	3000	500
38	4884	2	4019	375
39	1404	4	1000	240
40	6972	1	4019	525
41	6720	3	4096	1500
42	3216	2	180	180
43	1428	1	750	375
44	577	1	200	200
45	1020	1	240	240
46	1800	1	560	192
47	373	1	300	300
48	193	1	288	288
49	492	1	144	144
50	360	1	100	100
51	336	1	225	225
52	1800	6	10800	180
53	504	1	225	225
54	312	2	94	94
55	1008	1	180	180
56	3720	8	600	600

57	552	10	1000	405
58	240	2	324	324
59	492	2	252	252
60	984	3	405	405
61	1068	2	980	504
62	3192	8	2700	2700
63	553	1	535	248
64	1560	2	843	843
65	936	2	300	300
66	672	2	843	506
67	492	2	162	81
68	229	1	4247	33
69	720	2	400	400
70	2532	2	320	320
71	648	1	288	288
72	708	1	3267	2400
73	2640	4	315	315
74	252	3	160	160
75	936	1	1100	270
76	398	2	490	270
77	432	1	375	375
78	241	2	450	200
79	5460	9	3600	1800
80	337	1	320	320
81	1680	4	225	225
82	5496	7	3375	864
83	5160	14	1633	270
84	984	3	216	216
85	277	1	300	300
86	1380	1	150	150
87	444	1	270	270
88	744	1	35	35
89	528	1	1012	675
90	170	1	2700	506
91	3084	1	216	72
92	1440	1	7840	225
93	2460	3	980	64
94	516	1	653	326
95	2340	1	150	60
96	1212	6	150	30
97	504	1	5227	60
98	1440	1	2613	272
99	157	1	9801	100
100	1800	1	180	180
101	4776	2	1012	675
102	1152	3	750	225
103	792	4	700	700
104	252	3	200	200
105	6684	2	1000	180
106	492	2	252	252
107	3648	3	675	675
108	98	1	1600	337
109	660	2	1125	337
110	1248	2	2500	150

Appendix 10: Data for model SFA1-6

Serial #	Total vaccine	Total vaccine weighted by price	Total vaccine weighted by DALY (000,000s)	Labour	Size of EPI (sq ft)	Total hours
1	5016	348	9938	20	600	154
2	492	40	1115	2	79	79
3	2028	158	4499	1	1620	216
4	5028	422	10237	2	252	252
5	4008	342	4351	2	405	405
6	4116	349	4666	1	432	432
7	3936	327	9949	2	600	450
8	924	62	1850	2	75	75
9	4416	313	9862	1	270	270
10	9696	794	23889	2	9000	1296
11	924	70	1921	2	13068	225
12	5160	416	12909	2	3267	324
13	3228	282	7687	2	4500	720
14	2064	164	5603	1	990	135
15	6360	539	15255	2	4019	162
16	4260	355	9956	1	300	300
17	8124	685	19644	10	2500	144
18	7800	641	16456	1	600	200
19	9252	756	21493	10	2000	144
20	276	17	487	6	113	113
21	3132	249	6941	4	120	120
22	1080	79	2316	3	6750	164
23	3696	275	8953	6	816	144
24	3768	221	6644	5	959	338
25	1452	117	3376	2	270	270
26	1320	93	2821	1	108	108
27	2256	182	5352	13	12500	100
28	697	53	1609	6	864	216
29	7884	607	17853	9	450	450
30	1296	95	2646	1	216	216
31	3120	239	5865	1	9000	135
32	157	15	402	3	2613	70
33	300	23	659	5	900	900
34	1284	97	2962	3	1350	270
35	1176	54	1581	5	7200	360
36	5580	460	12829	7	7500	900
37	5148	419	12800	2	3000	500
38	4884	406	10825	2	4019	375
39	1404	122	3262	4	1000	240
40	6972	563	14770	1	4019	525
41	6720	483	14569	3	4096	1500
42	3216	273	8188	2	180	180
43	1428	110	3282	1	750	375
44	577	38	1108	1	200	200
45	1020	73	2033	1	240	240
46	1800	149	4070	1	560	192
47	373	23	673	1	300	300
48	193	20	659	1	288	288
49	492	40	1060	1	144	144
50	360	27	665	1	100	100
51	336	28	785	1	225	225
52	1800	149	4070	6	10800	180
53	504	40	1050	1	225	225
54	312	25	646	2	94	94

55	1008	87	2414	1	180	180
56	3720	288	8132	8	600	600
57	552	45	1216	10	1000	405
58	240	16	422	2	324	324
59	492	38	999	2	252	252
60	984	83	2384	3	405	405
61	1068	83	2266	2	980	504
62	3192	255	7152	8	2700	2700
63	553	43	1387	1	535	248
64	1560	130	3515	2	843	843
65	936	81	1962	2	300	300
66	672	57	1324	2	843	506
67	492	44	1077	2	162	81
68	229	18	507	1	4247	33
69	720	60	1864	2	400	400
70	2532	194	4945	2	320	320
71	648	44	1248	1	288	288
72	708	57	1597	1	3267	2400
73	2640	207	5987	4	315	315
74	252	22	567	3	160	160
75	936	73	2254	1	1100	270
76	398	31	1008	2	490	270
77	432	34	1018	1	375	375
78	241	20	637	2	450	200
79	5460	445	12516	9	3600	1800
80	337	27	765	1	320	320
81	1680	126	3860	4	225	225
82	5496	400	11561	7	3375	864
83	5160	381	11948	14	1633	270
84	984	87	2723	3	216	216
85	277	19	585	1	300	300
86	1380	116	3517	1	150	150
87	444	28	805	1	270	270
88	744	53	1670	1	35	35
89	528	35	1112	1	1012	675
90	170	13	433	1	2700	506
91	3084	243	6704	1	216	72
92	1440	113	2959	1	7840	225
93	2460	196	5556	3	980	64
94	516	25	755	1	653	326
95	2340	187	5322	1	150	60
96	1212	93	2760	6	150	30
97	504	40	1191	1	5227	60
98	1440	113	2959	1	2613	272
99	157	11	307	1	9801	100
100	1800	131	3100	1	180	180
101	4776	419	11598	2	1012	675
102	1152	98	2196	3	750	225
103	792	76	1918	4	700	700
104	252	21	508	3	200	200
105	6684	568	14358	2	1000	180
106	492	44	1214	2	252	252
107	3648	308	8760	3	675	675
108	98	7	167	1	1600	337
109	660	45	1354	2	1125	337
110	1248	107	3360	2	2500	150

Appendix 11: Results of model DEA1

Serial #	Overall efficiency	Technical efficiency	Scale efficiency	Type of returns to scale
1	0.502	0.75	0.669	drs
2	0.131	0.304	0.432	irs
3	0.331	0.402	0.822	irs
4	0.294	0.304	0.966	irs
5	0.962	0.962	1	-
6	1	1	1	-
7	0.391	0.395	0.989	drs
8	0.29	0.435	0.667	irs
9	1	1	1	-
10	0.673	1	0.673	drs
11	0.129	0.239	0.541	irs
12	0.407	0.418	0.974	drs
13	0.243	0.268	0.906	irs
14	0.316	0.348	0.91	irs
15	0.504	0.529	0.952	drs
16	0.542	0.565	0.96	irs
17	0.391	0.601	0.651	drs
18	1	1	1	-
19	0.592	1	0.592	drs
20	0.174	0.592	0.293	irs
21	0.571	0.65	0.879	irs
22	0.098	0.163	0.601	irs
23	0.266	0.272	0.978	irs
24	0.726	1	0.726	drs
25	0.566	0.756	0.749	irs
26	0.534	0.806	0.662	irs
27	0.164	0.176	0.929	irs
28	0.246	0.477	0.516	irs
29	0.791	1	0.791	drs
30	0.624	0.863	0.723	irs
31	0.961	1	0.961	irs
32	0.246	1	0.246	irs
33	0.161	0.619	0.26	irs
34	0.275	0.414	0.663	irs
35	1	1	1	-
36	0.579	0.587	0.986	drs
37	0.113	0.115	0.98	drs
38	0.081	0.083	0.973	irs
39	0.054	0.078	0.696	irs
40	0.224	0.29	0.772	drs
41	0.689	1	0.689	drs
42	0.327	0.341	0.957	irs
43	0.443	0.565	0.784	irs
44	0.199	0.464	0.428	irs
45	0.296	0.444	0.667	irs
46	0.16	0.209	0.765	irs
47	0.09	0.25	0.358	irs
48	0.117	0.324	0.361	irs
49	0.121	0.308	0.391	irs
50	0.055	0.172	0.322	irs
51	0.09	0.277	0.323	irs
52	0.16	0.209	0.765	irs
53	0.106	0.275	0.388	irs
54	0.134	0.512	0.262	irs
55	0.432	0.681	0.634	irs
56	0.186	0.19	0.977	irs

57	0.036	0.078	0.462	irs
58	0.146	0.626	0.233	irs
59	0.212	0.579	0.367	irs
60	0.129	0.203	0.635	irs
61	0.066	0.112	0.594	irs
62	0.272	0.292	0.932	irs
63	0.254	0.501	0.507	irs
64	0.552	0.772	0.715	irs
65	0.427	0.724	0.59	irs
66	0.309	0.616	0.502	irs
67	0.363	0.855	0.424	irs
68	0.197	0.759	0.26	irs
69	0.41	0.704	0.583	irs
70	1	1	1	-
71	0.386	0.785	0.491	irs
72	0.127	0.256	0.497	irs
73	0.479	0.553	0.865	irs
74	0.124	0.532	0.234	irs
75	0.192	0.312	0.615	irs
76	0.066	0.152	0.433	irs
77	0.103	0.265	0.388	irs
78	0.192	0.666	0.289	irs
79	0.206	0.208	0.992	drs
80	0.178	0.488	0.365	irs
81	0.357	0.458	0.779	irs
82	0.171	0.224	0.764	drs
83	0.355	0.395	0.899	drs
84	0.508	0.666	0.763	irs
85	0.144	0.566	0.255	irs
86	0.529	0.706	0.749	irs
87	0.174	0.433	0.402	irs
88	0.224	0.461	0.487	irs
89	0.176	0.432	0.408	irs
90	0.203	0.883	0.23	irs
91	0.703	0.79	0.889	irs
92	0.449	0.561	0.801	irs
93	0.334	0.37	0.901	irs
94	0.283	0.473	0.598	irs
95	0.976	1	0.976	irs
96	0.036	0.051	0.71	irs
97	0.12	0.299	0.399	irs
98	0.376	0.469	0.801	irs
99	0.127	0.784	0.162	irs
100	0.76	0.828	0.919	irs
101	0.415	0.423	0.982	irs
102	0.155	0.229	0.678	irs
103	0.371	0.486	0.763	irs
104	0.063	0.285	0.219	irs
105	0.654	1	0.654	drs
106	0.192	0.449	0.427	irs
107	0.376	0.39	0.964	irs
108	0.045	0.415	0.109	irs
109	0.063	0.135	0.465	irs
110	0.451	0.59	0.765	irs

Appendix 12: Results of model DEA2

Serial #	Overall efficiency	Technical efficiency	Scale efficiency	Type of returns to scale
1	1	1	1	-
2	0.346	0.5	0.693	irs
3	0.357	1	0.357	irs
4	1	1	1	-
5	0.969	1	0.969	drs
6	1	1	1	-
7	0.579	0.61	0.948	drs
8	0.826	0.871	0.948	drs
9	1	1	1	-
10	0.907	1	0.907	drs
11	0.126	0.5	0.252	irs
12	0.568	0.572	0.994	irs
13	0.272	0.5	0.544	irs
14	0.601	1	0.601	irs
15	1	1	1	-
16	1	1	1	-
17	1	1	1	-
18	1	1	1	-
19	1	1	1	-
20	0.17	0.31	0.55	irs
21	1	1	1	-
22	0.258	0.333	0.775	irs
23	0.935	0.937	0.998	irs
24	1	1	1	-
25	0.387	0.5	0.774	irs
26	0.715	1	0.715	irs
27	0.434	0.508	0.854	irs
28	0.127	0.173	0.736	irs
29	1	1	1	-
30	0.417	1	0.417	irs
31	0.744	1	0.744	irs
32	0.093	0.487	0.191	irs
33	0.022	0.2	0.109	irs
34	0.201	0.333	0.603	irs
35	0.417	0.423	0.986	irs
36	0.198	0.203	0.98	drs
37	0.447	0.5	0.893	irs
38	0.414	0.5	0.827	irs
39	0.175	0.25	0.701	irs
40	1	1	1	-
41	0.456	1	0.456	drs
42	1	1	1	-
43	0.312	1	0.312	irs
44	0.195	1	0.195	irs
45	0.346	1	0.346	irs
46	0.303	1	0.303	irs
47	0.111	1	0.111	irs
48	0.106	1	0.106	irs
49	0.234	1	0.234	irs
50	0.225	1	0.225	irs
51	0.116	1	0.116	irs
52	0.206	0.307	0.67	irs
53	0.146	1	0.146	irs
54	0.169	0.5	0.337	irs
55	0.434	1	0.434	irs
56	0.388	0.389	0.997	irs

57	0.054	0.1	0.538	irs
58	0.053	0.5	0.107	irs
59	0.114	0.5	0.227	irs
60	0.183	0.333	0.55	irs
61	0.098	0.5	0.195	irs
62	0.096	0.125	0.768	irs
63	0.129	1	0.129	irs
64	0.155	0.5	0.31	irs
65	0.22	0.5	0.44	irs
66	0.08	0.5	0.16	irs
67	0.277	0.5	0.555	irs
68	0.169	1	0.169	irs
69	0.156	0.5	0.311	irs
70	0.58	0.637	0.91	drs
71	0.188	1	0.188	irs
72	0.114	1	0.114	irs
73	0.438	0.445	0.984	irs
74	0.081	0.333	0.243	irs
75	0.191	1	0.191	irs
76	0.084	0.5	0.167	irs
77	0.105	1	0.105	irs
78	0.052	0.5	0.104	irs
79	0.138	0.146	0.945	drs
80	0.106	1	0.106	irs
81	0.4	0.401	0.998	irs
82	0.313	0.46	0.679	drs
83	0.644	0.644	1	-
84	0.322	0.412	0.783	irs
85	0.069	1	0.069	irs
86	0.626	1	0.626	irs
87	0.142	1	0.142	irs
88	1	1	1	-
89	0.14	1	0.14	irs
90	0.04	1	0.04	irs
91	1	1	1	-
92	0.286	1	0.286	irs
93	0.735	0.94	0.782	irs
94	0.233	1	0.233	irs
95	1	1	1	-
96	1	1	1	-
97	0.32	1	0.32	irs
98	0.286	1	0.286	irs
99	0.047	1	0.047	irs
100	0.849	1	0.849	irs
101	0.476	0.5	0.952	irs
102	0.169	0.333	0.508	irs
103	0.141	0.25	0.565	irs
104	0.067	0.333	0.2	irs
105	1	1	1	-
106	0.145	0.5	0.291	irs
107	0.434	0.434	1	-
108	0.021	1	0.021	irs
109	0.127	0.5	0.255	irs
110	0.299	0.5	0.598	irs

Appendix 13: Results of model DEA3

Serial #	Overall efficiency	Technical efficiency	Scale efficiency	Type of returns to scale
1	0.73	0.766	0.953	drs
2	0.258	0.5	0.516	irs
3	0.26	1	0.26	irs
4	1	1	1	-
5	0.569	0.58	0.981	irs
6	0.677	1	0.677	irs
7	0.425	0.5	0.85	irs
8	0.503	0.569	0.884	irs
9	1	1	1	-
10	0.622	1	0.622	drs
11	0.101	0.5	0.202	irs
12	0.403	0.5	0.806	irs
13	0.207	0.5	0.414	irs
14	0.382	1	0.382	irs
15	0.931	0.936	0.994	drs
16	0.902	1	0.902	irs
17	0.878	0.897	0.979	irs
18	1	1	1	-
19	1	1	1	-
20	0.094	0.31	0.302	irs
21	1	1	1	-
22	0.148	0.333	0.444	irs
23	0.534	0.536	0.995	irs
24	0.268	0.277	0.969	drs
25	0.276	0.5	0.551	irs
26	0.603	1	0.603	irs
27	0.351	0.461	0.762	irs
28	0.072	0.167	0.432	irs
29	0.766	1	0.766	drs
30	0.35	1	0.35	irs
31	0.577	1	0.577	irs
32	0.042	0.485	0.086	irs
33	0.019	0.2	0.093	irs
34	0.114	0.333	0.343	irs
35	0.076	0.2	0.381	irs
36	0.154	0.154	0.999	-
37	0.33	0.5	0.66	irs
38	0.333	0.5	0.665	irs
39	0.133	0.25	0.533	irs
40	0.894	1	0.894	irs
41	0.287	0.333	0.862	irs
42	0.862	0.876	0.985	irs
43	0.183	1	0.183	irs
44	0.165	1	0.165	irs
45	0.254	1	0.254	irs
46	0.245	1	0.245	irs
47	0.079	1	0.079	irs
48	0.042	1	0.042	irs
49	0.179	1	0.179	irs
50	0.176	1	0.176	irs
51	0.088	1	0.088	irs
52	0.199	0.276	0.719	irs
53	0.132	1	0.132	irs
54	0.143	0.5	0.286	irs
55	0.312	1	0.312	irs
56	0.292	0.294	0.994	irs

57	0.035	0.1	0.352	irs
58	0.04	0.5	0.08	irs
59	0.098	0.5	0.196	irs
60	0.125	0.333	0.374	irs
61	0.08	0.5	0.16	irs
62	0.078	0.125	0.625	irs
63	0.078	1	0.078	irs
64	0.131	0.5	0.261	irs
65	0.165	0.5	0.331	irs
66	0.056	0.5	0.113	irs
67	0.175	0.5	0.35	irs
68	0.141	1	0.141	irs
69	0.103	0.5	0.206	irs
70	0.427	0.5	0.855	irs
71	0.141	1	0.141	irs
72	0.091	1	0.091	irs
73	0.397	0.407	0.975	irs
74	0.07	0.333	0.209	irs
75	0.12	1	0.12	irs
76	0.053	0.5	0.105	irs
77	0.079	1	0.079	irs
78	0.036	0.5	0.072	irs
79	0.107	0.111	0.967	irs
80	0.068	1	0.068	irs
81	0.334	0.353	0.946	irs
82	0.158	0.158	0.999	-
83	0.391	0.401	0.976	drs
84	0.213	0.333	0.639	irs
85	0.059	1	0.059	irs
86	0.487	1	0.487	irs
87	0.101	1	0.101	irs
88	0.852	1	0.852	irs
89	0.068	1	0.068	irs
90	0.022	1	0.022	irs
91	1	1	1	-
92	0.185	1	0.185	irs
93	0.692	0.896	0.772	irs
94	0.066	1	0.066	irs
95	1	1	1	-
96	0.864	1	0.864	irs
97	0.191	1	0.191	irs
98	0.185	1	0.185	irs
99	0.038	1	0.038	irs
100	0.558	1	0.558	irs
101	0.349	0.5	0.699	irs
102	0.12	0.333	0.361	irs
103	0.063	0.25	0.251	irs
104	0.058	0.333	0.174	irs
105	0.892	0.896	0.995	drs
106	0.098	0.5	0.196	irs
107	0.318	0.333	0.955	irs
108	0.013	1	0.013	irs
109	0.05	0.5	0.099	irs
110	0.195	0.5	0.391	irs

Appendix 14: Results for models SFA1-6

Serial #	SFA1	SFA2	SFA3	SFA4	SFA5	SFA6
1	0.603	0.276	0.602	0.273	0.572	0.240
2	0.426	0.117	0.488	0.093	0.421	0.137
3	0.611	0.225	0.626	0.207	0.606	0.230
4	0.721	0.934	0.720	0.935	0.713	0.999
5	0.678	0.730	0.686	0.730	0.574	0.395
6	0.709	0.859	0.712	0.867	0.627	0.453
7	0.666	0.605	0.675	0.574	0.683	0.723
8	0.547	0.224	0.555	0.144	0.522	0.228
9	0.732	0.842	0.716	0.753	0.736	0.999
10	0.700	0.670	0.701	0.393	0.711	0.647
11	0.360	0.118	0.424	0.087	0.326	0.146
12	0.676	0.434	0.680	0.360	0.688	0.508
13	0.581	0.243	0.615	0.191	0.587	0.250
14	0.637	0.258	0.650	0.252	0.661	0.323
15	0.711	0.806	0.714	0.716	0.715	0.901
16	0.725	0.825	0.723	0.856	0.733	0.999
17	0.679	0.579	0.686	0.546	0.682	0.582
18	0.774	0.989	0.763	0.999	0.772	0.999
19	0.698	0.659	0.698	0.608	0.697	0.625
20	0.224	0.058	0.268	0.038	0.179	0.048
21	0.666	0.671	0.670	0.557	0.664	0.731
22	0.391	0.136	0.443	0.100	0.362	0.146
23	0.623	0.363	0.629	0.294	0.628	0.351
24	0.608	0.288	0.586	0.205	0.563	0.255
25	0.549	0.268	0.581	0.256	0.553	0.325
26	0.627	0.242	0.627	0.208	0.620	0.313
27	0.423	0.368	0.483	0.298	0.408	0.413
28	0.305	0.055	0.372	0.049	0.293	0.059
29	0.685	0.686	0.680	0.816	0.685	0.999
30	0.588	0.241	0.601	0.226	0.573	0.273
31	0.649	0.692	0.656	0.620	0.617	0.761
32	0.143	0.043	0.244	0.039	0.141	0.037
33	0.155	0.033	0.223	0.032	0.142	0.039
34	0.453	0.113	0.496	0.093	0.446	0.118
35	0.338	0.055	0.324	0.026	0.231	0.043
36	0.571	0.131	0.598	0.119	0.566	0.196
37	0.667	0.417	0.674	0.338	0.679	0.465
38	0.662	0.387	0.673	0.324	0.657	0.404
39	0.461	0.123	0.524	0.120	0.457	0.131
40	0.725	0.701	0.720	0.542	0.720	0.612
41	0.648	0.532	0.643	0.323	0.642	0.464
42	0.685	0.623	0.692	0.613	0.701	0.851
43	0.556	0.207	0.581	0.185	0.559	0.213
44	0.447	0.107	0.472	0.089	0.412	0.115
45	0.542	0.192	0.559	0.172	0.520	0.208
46	0.620	0.231	0.640	0.235	0.620	0.251
47	0.336	0.072	0.367	0.055	0.289	0.067
48	0.227	0.037	0.348	0.047	0.287	0.066
49	0.438	0.089	0.499	0.092	0.426	0.115
50	0.402	0.066	0.456	0.060	0.357	0.075
51	0.335	0.063	0.414	0.065	0.337	0.081
52	0.420	0.165	0.485	0.137	0.399	0.211
53	0.413	0.094	0.476	0.096	0.394	0.108
54	0.327	0.070	0.404	0.058	0.301	0.076
55	0.557	0.185	0.598	0.204	0.568	0.254
56	0.571	0.315	0.589	0.367	0.561	0.438

57	0.213	0.026	0.293	0.031	0.195	0.034
58	0.214	0.044	0.273	0.036	0.173	0.040
59	0.349	0.091	0.417	0.085	0.321	0.098
60	0.421	0.158	0.487	0.161	0.429	0.200
61	0.434	0.134	0.486	0.113	0.415	0.127
62	0.449	0.206	0.497	0.199	0.442	0.260
63	0.402	0.076	0.463	0.073	0.416	0.091
64	0.491	0.303	0.541	0.266	0.488	0.285
65	0.462	0.172	0.524	0.176	0.442	0.186
66	0.348	0.091	0.431	0.086	0.314	0.081
67	0.404	0.094	0.485	0.084	0.388	0.090
68	0.256	0.251	0.336	0.224	0.233	0.228
69	0.391	0.131	0.464	0.129	0.413	0.169
70	0.628	0.463	0.637	0.423	0.608	0.467
71	0.446	0.125	0.477	0.107	0.411	0.125
72	0.314	0.239	0.393	0.115	0.310	0.111
73	0.595	0.399	0.611	0.388	0.595	0.493
74	0.233	0.049	0.326	0.048	0.222	0.058
75	0.480	0.109	0.525	0.100	0.487	0.119
76	0.289	0.055	0.366	0.050	0.301	0.067
77	0.349	0.087	0.419	0.083	0.354	0.100
78	0.221	0.033	0.306	0.030	0.234	0.041
79	0.546	0.173	0.577	0.182	0.543	0.255
80	0.313	0.066	0.391	0.065	0.308	0.076
81	0.539	0.280	0.563	0.253	0.538	0.345
82	0.589	0.186	0.596	0.164	0.573	0.235
83	0.589	0.178	0.598	0.190	0.585	0.228
84	0.463	0.177	0.527	0.183	0.497	0.260
85	0.283	0.054	0.338	0.045	0.264	0.058
86	0.617	0.251	0.639	0.266	0.638	0.379
87	0.376	0.085	0.409	0.068	0.329	0.081
88	0.599	0.165	0.608	0.109	0.599	0.219
89	0.338	0.098	0.383	0.068	0.318	0.078
90	0.155	0.018	0.230	0.015	0.159	0.019
91	0.725	0.475	0.720	0.438	0.724	0.499
92	0.512	0.190	0.552	0.163	0.483	0.212
93	0.622	0.558	0.637	0.405	0.615	0.387
94	0.373	0.074	0.366	0.044	0.285	0.050
95	0.706	0.387	0.705	0.348	0.709	0.432
96	0.532	0.546	0.562	0.244	0.522	0.224
97	0.371	0.236	0.444	0.229	0.361	0.273
98	0.535	0.151	0.569	0.134	0.512	0.143
99	0.159	0.053	0.227	0.046	0.131	0.062
100	0.646	0.330	0.646	0.304	0.609	0.326
101	0.671	0.691	0.684	0.644	0.682	0.723
102	0.455	0.122	0.518	0.117	0.411	0.105
103	0.322	0.104	0.425	0.123	0.326	0.132
104	0.222	0.046	0.307	0.044	0.196	0.049
105	0.734	0.756	0.732	0.709	0.730	0.710
106	0.349	0.091	0.440	0.098	0.358	0.119
107	0.621	0.567	0.641	0.557	0.632	0.666
108	0.118	0.011	0.182	0.009	0.090	0.008
109	0.351	0.069	0.398	0.052	0.322	0.064
110	0.482	0.155	0.540	0.141	0.504	0.184

Appendix 15: Facility survey form for health centres

Name of health centre:

Ward:

Upazila:

District:

Personnel

Indicate the number and type of staff working at this health centre

Name	Designation	July 2001 salary		June 2002 salary		Period of employment during July 2001 – June 2002
		Basic	Allowances	Basic	Allowances	

Vehicles

Indicate the number of and type of functioning vehicles at this health centre

Type of mean of transport	Quantity	Was it functioning in 2001 (specify in which months it was functioning if it was functioning only in parts of the year)

Buildings

Building No.	Room No.	Who works in this room	Square feet

Furniture

Indicate the number of functioning items available in each of the rooms of the facility (including waiting areas)

Capital furniture	**R 1**	**R 2**	**R 3**	**R 4**	**R 5**	**R 6**	**R 7**	**R 8**	**R 9**	**R 10**	**R 11**	**R 12**	**R 13**	**R 14**	**R 15**
Chair															
Clock															
Cupboard															
Delivery bed															
Drip stand															
Examination bed															
Hand washing basin															
Labour table															
Benches															
Mattress															
Metal file cabinet															
Screen															
Secretarial table															
Table (wood)															
Other (specify)															

Capital equipment

Indicate the number of functioning items available in each of the rooms of the facility (including waiting areas)

Capital equipment	**R 1**	**R 2**	**R 3**	**R 4**	**R 5**	**R 6**	**R 7**	**R 8**	**R 9**	**R 10**	**R 11**	**R 12**	**R 13**	**R 14**
Airway tube														
Autoclave machine														
Breast pump														
Catheter (rubber)														
Centrifuge														
Cold Box														
Drip stand														
Forceps														
Height measuring scale														
Incinerator														
Instrument tray														
Kidney tray														
Microscope														
Needle holder														
Other (specify)														
Refrigerator														
Rubbish bin														
Scissors														
Sterilizer/stove 4 burner														
Stethoscope														
Suction machine														
Surgical blade														
Test tube														
Vaccine carrier														
Weighing scale														
Wheelchair														
Other (specify)														

Drugs

Drug name	State the total used at this site between July 2000 – June 2001	State the total used at satellite sites between July 2000 – June 2001 (if applicable)	State the amount of each drug that you had in stock at the beginning of the July 2000	State the amount of each drug that you were supplied with during 2000/2001	State the amount of each drug that you had at the end of June 2001

Medical supplies

Name of medical supply	State the total used at this site	State the total used at satellite sites	State the amount of each medical supply that you had in stock at the beginning of the January	State the amount of each medical supply that you were supplied with during 2001	State the amount of each medical supply that you had at the end of the 2001

Table 14: Expenditure on utilities, operating and maintenance between July 2000 – June 2001

#	Item	Total
1.	**Means of transport**	
	Petrol	
	Lubricant	
	Maintenance	
	Repairs	
	Insurance	
	Tire spare parts	
	Others…list:	
	Total	
2.	**Building**	
	Electricity	
	Water	
	Facility rent (if relevant)	
	Maintenance	
	Telephone	
	Charcoal	
	Kerosene	
	Cleaning	
	Others…list:	
	Total	
3.	**Equipment**	
	Spare parts	
	Repairs	
	Others…list:	
	Total	

Record review

What was the total number of visits to the health centre for the following services in 2001?

Type of visit	Patients									Period covered if missing data
	Female			Male			Total			
Age	< 1	1-4	+ 4	< 1	1-4	+ 4	< 1	1-4	+ 4	
Dysentery										
Diarrhoea										
Diarrhoea / no dehydration										
Diarrhoea / some dehydration										
Diarrhoea / severe dehydration										
Diphtheria										
Jaundice / Hepatitis										
Measles										
Meningitis										
Neonatal tetanus										
Pertussis (whooping cough)										
ARI										
Pneumonia										
Severe pneumonia										
Very severe pneumonia										
Total of all cases seen										

Appendix 16: Data for model DEA1

Serial #	Total cost	visits <1	visits 1-4	visits +4
1	7,651	167	744	6261
2	8,136	110	674	4023
3	6,317	292	909	5073
4	9,032	182	665	4259
5	5,983	156	580	2742
6	12,959	83	810	8368
7	12,048	1682	2158	10073
8	8,758	242	608	5596
9	14,642	306	1495	9904
10	6,877	344	778	2770
11	8,230	431	879	5318
12	4,707	429	1380	6914
13	9,023	988	2145	10925
14	11,192	883	1055	6982
15	10,381	686	2474	13536
16	13,438	535	1993	14032
17	4,686	247	1171	6365
18	7,322	627	1938	9385
19	12,347	1268	2790	20048
20	12,030	3121	4464	6919
21	11,017	878	1923	6286
22	10,021	2637	3183	9996
23	11,464	1411	3641	32010
24	6,357	343	2000	12586
25	7,075	159	1217	5399
26	8,050	947	1815	6154
27	5,941	327	1012	5095
28	6,049	719	999	7586
29	11,961	693	1537	4855
30	7,702	716	2865	11489
31	6,867	699	1416	9700
32	7,067	273	1223	4772
33	7,157	491	3574	18425
34	7,117	450	1644	9426

Appendix 17: Data for model DEA2

Serial #	staff costs ($)	drug costs ($)	total visits
1	2905	3556	7172
2	2711	3660	4807
3	3340	1315	6274
4	3854	3900	5106
5	2558	2314	3478
6	5871	6058	9261
7	4771	6380	13913
8	3588	4010	6446
9	5508	7775	11705
10	3389	3311	3892
11	3359	3768	6628
12	1376	2060	8723
13	2991	4278	14058
14	5039	269	8920
15	5077	2329	16696
16	6357	910	16560
17	2207	2113	7783
18	4342	1667	11950
19	6273	1139	24106
20	5454	2214	14504
21	4444	3681	9087
22	5445	2268	15816
23	5974	2020	37062
24	5799	1806	14929
25	3869	2191	6775
26	4503	3363	8916
27	2777	1975	6434
28	2261	2439	9304
29	6441	2780	7085
30	3694	2859	15070
31	3768	3181	11815
32	3701	1755	6268
33	3429	2524	22490
34	3608	2347	11520

Appendix 18: Data for model DEA3

Serial #	staff costs ($)	drug costs ($)	visits <1	visits 1-4	visits +4
1	2905	3556	167	744	6261
2	2711	3660	110	674	4023
3	3340	1315	292	909	5073
4	3854	3900	182	665	4259
5	2558	2314	156	580	2742
6	5871	6058	83	810	8368
7	4771	6380	1682	2158	10073
8	3588	4010	242	608	5596
9	5508	7775	306	1495	9904
10	3389	3311	344	778	2770
11	3359	3768	431	879	5318
12	1376	2060	429	1380	6914
13	2991	4278	988	2145	10925
14	5039	269	883	1055	6982
15	5077	2329	686	2474	13536
16	6357	910	535	1993	14032
17	2207	2113	247	1171	6365
18	4342	1667	627	1938	9385
19	6273	1139	1268	2790	20048
20	5454	2214	3121	4464	6919
21	4444	3681	878	1923	6286
22	5445	2268	2637	3183	9996
23	5974	2020	1411	3641	32010
24	5799	1806	343	2000	12586
25	3869	2191	159	1217	5399
26	4503	3363	947	1815	6154
27	2777	1975	327	1012	5095
28	2261	2439	719	999	7586
29	6441	2780	693	1537	4855
30	3694	2859	716	2865	11489
31	3768	3181	699	1416	9700
32	3701	1755	273	1223	4772
33	3429	2524	491	3574	18425
34	3608	2347	450	1644	9426

Appendix 19: Results for model DEA1

Serial #	Overall efficiency	Technical efficiency	Scale efficiency	Type of returns to scale
1	0.309	0.571	0.541	irs
2	0.185	0.5	0.37	irs
3	0.366	0.667	0.548	irs
4	0.18	0.444	0.404	irs
5	0.253	0.8	0.316	irs
6	0.24	0.365	0.657	irs
7	0.597	0.639	0.935	irs
8	0.24	0.5	0.481	irs
9	0.255	0.341	0.747	irs
10	0.311	0.667	0.466	irs
11	0.312	0.501	0.624	irs
12	0.796	1	0.796	irs
13	0.629	0.685	0.918	irs
14	0.374	0.476	0.787	irs
15	0.57	0.61	0.934	irs
16	0.385	0.453	0.85	irs
17	0.631	1	0.631	irs
18	0.64	0.725	0.883	irs
19	0.681	0.705	0.966	irs
20	1	1	1	-
21	0.418	0.488	0.856	irs
22	1	1	1	-
23	1	1	1	-
24	0.764	0.913	0.836	irs
25	0.341	0.571	0.596	irs
26	0.576	0.677	0.851	irs
27	0.473	0.8	0.591	irs
28	0.626	0.806	0.777	irs
29	0.331	0.429	0.771	irs
30	0.861	0.912	0.944	irs
31	0.7	0.862	0.812	irs
32	0.362	0.571	0.633	irs
33	1	1	1	-
34	0.548	0.666	0.824	irs

Appendix 20: Results of model DEA2

Serial #	Overall efficiency	Technical efficiency	Scale efficiency	Type of returns to scale
1	0.376	0.549	0.685	irs
2	0.27	0.548	0.493	irs
3	0.3	0.927	0.324	irs
4	0.202	0.472	0.428	irs
5	0.207	0.767	0.27	irs
6	0.241	0.311	0.774	irs
7	0.445	0.451	0.987	irs
8	0.274	0.474	0.578	irs
9	0.324	0.331	0.98	irs
10	0.175	0.55	0.318	irs
11	0.301	0.505	0.596	irs
12	0.967	1	0.967	irs
13	0.717	0.726	0.987	irs
14	1	1	1	-
15	0.521	0.696	0.748	irs
16	0.809	0.829	0.975	irs
17	0.538	0.856	0.628	irs
18	0.441	0.786	0.561	irs
19	1	1	1	-
20	0.425	0.651	0.652	irs
21	0.312	0.471	0.662	irs
22	0.463	0.665	0.696	irs
23	1	1	1	-
24	0.444	0.691	0.642	irs
25	0.273	0.669	0.408	irs
26	0.302	0.493	0.612	irs
27	0.354	0.82	0.432	irs
28	0.627	0.783	0.801	irs
29	0.175	0.461	0.38	irs
30	0.622	0.692	0.899	irs
31	0.478	0.592	0.808	irs
32	0.268	0.767	0.349	irs
33	1	1	1	-
34	0.493	0.717	0.687	irs

Appendix 21: Results of model DEA3

Serial #	Overall efficiency	Technical efficiency	Scale efficiency	Type of returns to scale
1	0.401	0.549	0.73	irs
2	0.276	0.548	0.504	irs
3	0.381	0.927	0.411	irs
4	0.209	0.472	0.442	irs
5	0.231	0.767	0.302	irs
6	0.265	0.32	0.828	irs
7	0.756	0.768	0.984	drs
8	0.291	0.474	0.614	irs
9	0.335	0.347	0.966	irs
10	0.248	0.55	0.45	irs
11	0.365	0.505	0.722	irs
12	1	1	1	-
13	0.884	0.909	0.972	drs
14	1	1	1	-
15	0.627	0.741	0.846	irs
16	0.838	0.873	0.96	irs
17	0.547	0.856	0.639	irs
18	0.619	0.845	0.732	irs
19	1	1	1	-
20	1	1	1	-
21	0.473	0.534	0.885	irs
22	0.938	0.946	0.991	irs
23	1	1	1	-
24	0.563	0.718	0.784	irs
25	0.348	0.669	0.521	irs
26	0.468	0.558	0.839	irs
27	0.412	0.82	0.503	irs
28	0.84	0.846	0.993	irs
29	0.297	0.494	0.6	irs
30	0.78	0.801	0.974	irs
31	0.57	0.614	0.928	irs
32	0.399	0.767	0.521	irs
33	1	1	1	-
34	0.528	0.729	0.725	irs