AN APPROACH TO DETERMINING INEFFICIENT WATER SERVICES PROVISION

Report to the
Water Research Commission

by

Stephen Hosking, Warren Brettenny, Siyathemba Magobiane, Gary Sharp & Jennifer Hosking
Departments of Economics and Statistics
Nelson Mandela Metropolitan University

WRC Report No. 2118/1/14
ISBN 978-1-4312-0540-0

APRIL 2014

Obtainable from

Water Research Commission
Private Bag X03
Gezina, 0031

orders@wrc.org.za or download from www.wrc.org.za

EXECUTIVE SUMMARY

In order to satisfy demand amongst water services users for greater efficiency in its provision, it is necessary that inefficiencies be identified. As a demonstration that there is such a demand, this study calculates willingness to pay amongst water service users for improved sanitation management by the Nelson Mandela Bay Municipality; the benefit being reduced sewage contamination of the Swartkops estuary. In 2012 the consumers of water service in this municipality were willing to pay an additional R3million to R10million per annum to improve the efficiency of the sanitation service (Appendix A).

How are the relevant inefficiencies to be identified? One way is to institute performance benchmarking aimed at efficiency in water service provision. The idea of benchmarking production and cost performance is not new (Chapter 1). Following high profile success during the 1970s it became standard practiced for decades in many parts of the private sector, even where there were differences in business type and output. These differences did not preclude firms from gaining insights into their own performance by observing the practices of others.

The embracement of performance benchmarking by firms and organisations in the public sector has been slower; two reasons being the uniqueness of public sector inputs and outputs, and the different (social) welfare perspective taken by public sector enterprises on hiring the inputs and delivering the outputs. Only by the 1990s had the idea of performance benchmarking had gained significant popularity within the public sector. The change in attitude within the public sector was motivated by the acceptance that performance could indeed be enhanced through benchmarking and measuring production and cost performance, and by the realisation that performance benchmarking was a way public sector organisations and officials could be held accountable for the way they spent the public funds entrusted to them. In order to reinforce the accountability function of performance benchmarking it was made audit and/or funding requirement for many public sector organisations.

Performance benchmarking is a practical method of comparing an organisation's performance against other organisations that are considered to be the best-in-class. Once the best-in-class performer is found, the differences between the current method and the best-in-

class method are identified and policy makers can attempt to engineer adaptions (strategic interventions) aimed at bringing inferior performers more into line with their best-in-class counterparts. It can assist in identifying areas where productivity can be improved through reorganisation and where breakdowns and losses in service are reduced, *inter alia*. If not done formally (learning from others), it is invariably done informally and unofficially by those charged with the responsibility for production.

Within the private sector corporate environment, performance benchmarking typically focuses attention on a specific process and endeavours to find the best-in-class performers of that process. Within public sector enterprises it is more common to find targets set as performance benchmarks – as yardsticks generated in strategic planning workshops.

The main aim of this study was to select and apply a method for assessing relative efficiency performance among South African municipal water service suppliers. In addressing this aim it investigated the feasibility of applying to South African municipal water service providers one of the parametric or non-parametric statistical performance benchmarking measures. Two forms of parametric estimation were considered, corrected ordinary least squares (COLS) and stochastic frontier analysis (SFA), and one form of non-parametric estimation, data envelopment analysis (DEA). These methods compare the statistics describing the performance of a particular water service provider against external performance standards or norms (benchmark statistics) as revealed by an analysis of the statistics of other (municipal) providers.

The COLS parametric method of generating an efficiency performance measure is a deterministic one (Chapter 2). A functional form is decided upon and then this function estimated using ordinary least squares corrected (adjusted) to a frontier position within the context of the sample data. A drawback of the deterministic frontier model is that it embeds all measurement error and possible stochastic variation in the dependent variable (output or cost) into the calculated economic (in)efficiency measure, and does not allow for the possibility of statistical noise or measurement error. The SFA parametric method overcomes this drawback of the COLS model by allowing for the possibility of statistical noise and measurement error in the model, and for this reason is considered the superior parametric approach, but the data requirements for its successful application are not always satisfied (Chapter 2).

Instead of imposing a specific functional form to the frontier (as done following a parametric methodology), the (non-parametric) DEA method of generating performance benchmarks uses the data itself to construct a quasi-convex hull (frontier) of linear segments around the data (Chapter 3). This hull represents the production frontier against which measures of relative efficiency are determined.

The above described methods were applied to information collected about a number of different municipal suppliers of water services (see Chapter 4 and Appendix C). As a result, various efficiency indices could be calculated for 39 selected local municipalities and 12 selected metropolitan and district municipalities (see Chapters 5 and 6).

The non-parametric indices calculated (using the DEA methodology) were preferred to the parametric indices on the ground that the DEA method is more closely linked to the principles of a frontier than those based on estimating parameters of pre-specified models. It was also found that the validity of the DEA derived indices could be improved by applying a bias correction procedure and that a modified method could be applied to rank (maximally) equivalently efficient municipalities. The modified method for calculating the DEA frontier is one where the DEA efficiency index is calculated by excluding the municipality from the DEA evaluation, with reference to a re-formed frontier excluding the particular municipality under consideration. The excluded municipality is then re-introduced and its efficiency estimate is recalculated against the new frontier (formed in its absence). When the efficiency indices are recalculated in this way, all the maximally efficiency estimates, which were 1, are larger than 1, because the relevant municipality falls beyond the frontier which was constructed in its absence. This method serves to enable a ranking of otherwise equivalent (maximally) efficient water service providers, while at the same time provides a mechanism for identifying outliers. Efficiencies recalculated in this manner are termed "super-efficiencies" and potential outliers are identified by having super-efficiency indices above 18. The presence of outliers may indicate a need to reconsider the data for consistency and accuracy. A couple of outliers were identified in the data set used.

Applying the DEA methodology to a combined set of data, including metropolitan, district and local municipalities, it was found that the metropolitan municipalities outperformed the local municipalities. The average efficiency index for metropolitan municipalities was calculated to be 0.986, while the average of the local municipalities was calculated to be 0.717; a difference that is probably due to the different operating environments of the respective groups of municipal water providers.

When efficiencies were calculated on the basis of the DEA estimates for the local municipalities as a separate group, their average efficiency estimate was calculated to be 0.7291, indicating a moderate overall efficiency level. The province with the highest average value of the local municipality efficiency index estimate was the Eastern Cape (0.9033), and the province with the lowest average estimate was the Northern Cape, with an average efficiency index estimate of 0.6364. The local municipality with the lowest estimated efficiency index level in the whole data set was the Mogalakwena local municipality in Limpopo, with an estimated efficiency index of 0.2850. The provinces with only a single observed local municipality (Limpopo, North West and KwaZulu-Natal) were excluded in the assessment of average efficiency.

When the efficiencies indices for the metropolitan and district municipalities were estimated for this group alone, the two metropolitan municipalities in the Eastern Cape were calculated to be relatively inefficient, namely Nelson Mandela Bay and Buffalo City. Within the district municipalities, there were a large number of WSA's who were performing efficiently. The average efficiency index level of the district and metropolitan municipalities was found to be 0.8397, indicating a reasonably high average level of efficiency, but it should also be borne in mind that the relative small of the sample size of the metropolitan and district municipalities had the effect of inflating the estimates of efficiency.

The study concluded that the municipal water services providing sector is suited to efficiency performance measuring and benchmarking (Chapter 7). There are a large number of separate water service providers and their performances differ, but there is sufficient commonality between their inputs and outputs to reject at a general level the uniqueness and welfare counter arguments against the usefulness of efficiency performance measuring and benchmarking. This conclusion is a qualified one, though, because there are many challenges that need to be resolved before the efficiency indices generated can be relied upon to guide policy and regulatory effort to improve efficiency performance. The most important of these challenges is data adequacy and authenticity.

This study was able to access relevant (and the best available) data for just over 50 municipalities in order to demonstrate how to rate their efficiency performance. With a bigger sample size and data covering more variables a statistically more reliable and preferred model could be estimated, and also efficiency performance indices.

Given that the parametrically derived estimates incorporated (average) input price information and the non-parametrically derived estimates did not, it would have been expected that differences would emerge between the two sets of efficiency indices. These differences were found, but they do not alter the conclusion drawn by this study – that either of the frontier estimation methodologies can serve as a platform for estimating relative efficiency performance across South Africa's water service providing municipalities. The case for incorporating or excluding input price information is not overwhelming in practice, but in theory, the input prices are key determinants of the cost function.

The study recommends that there take place discussions with and between the various water service authorities (municipalities) on how the data problem can and should be addressed, and on how the problem of measured inefficiency should be addressed once it is identified. It also recommends (prefers) the adoption of the DEA method for rating technical efficiency over the SFA one, because DEA more closely relates to the primal production optimisation problem being modelled, viz. maximising output with a given level of inputs (and cost outlay).

The ultimate rationale for statistical efficiency performance benchmarking is to trigger a process of improving efficiency in service delivery. Literature on this topic suggests that the way the signal is sent that efficiency should be improved has to be managed very carefully, and should take into account that there are limits on the scope for engineering change, for example, through threats of withholding budgets. If managed insensitively, efficiency performance measuring and benchmarking can become associated with negative signalling and undermine employee morale.

ACKNOWLEDGEMENTS

The financial support of the Water Research Commission is gratefully acknowledged, as was the guidance provided by a reference group assembled and managed by Mr Jay Bhagwan of the WRC. Without the support and commitment of numerous municipalities the research would not have been possible. They are sincerely thanked. In addition there were many other people who were part of the Benchmarking Initiative, such as Mr Jim Gibson that contributed to our understanding of the research problem and Mr Geoff Chandler of East London (formerly employed by Border Technikon, now Walter Sisulu University) and Ms Judy Keir who assisted with various aspects of data generation and compiling the final report. The advice of Mr Kevin Jacoby and Ms Judy Clarence, residents of Cape Town and Breede Valley municipalities respectively are also gratefully acknowledged.

The members of the WRC Reference Group that advised this study were:

Ms Katie Gull	Palmer Development Group (Member)
Dr Dawie Mullins	Conningarth Economists (Member)
Prof Hamanth Kasan	Rand Water (Member)
Ms Antonino Manus	City of Johannesburg (Member)
Mr Mark Bannister	Department of Water Affairs
Mr Ronald Gillmer	uThukela Water (Member)
Mr IM Mahomed	National Treasury (Member)
Mr William Moraka	SALGA (Member)
Mr Tawanda Nyandoro	Rand Water (Member)

The following capacity building was achieved through Project K5/2118

Student name	Employment	Degree (year awarded)	Title of thesis
Siyathemba Magobiane	Student, Nelson Mandela Metropolitan University	M Com (2012)	Willingness to Pay for water quality changes in the Swartkops estuary
Warren Brettenny	Lecturer, Nelson Mandela Metropolitan Univesity	PhD (not yet submitted – more material to be added)	Generating indices by which to monitor the efficiency of the provision of South African municipal water services

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

ANC	African National Congress
COLS	Corrected Ordinary Least Squares
CVM	Contingent Valuation Method
DEA	Data Envelopment Analysis
DWAF	Department of Water Affairs and Forestry
KPA	Key Performance Area
KPI	Key Performance Indicator
LAWA	Landerarbeitsgemeinschaft Wasser
NBI	National Benchmarking Initiative
OFWAT	Office of Water Services
OLS	Ordinary Least Squares
OMBI	Ontario Municipal Benchmarking Initiative
SALGA	South African Local Government Association
SFA	Stochastic Frontier Analysis
TWTP	Total Willingness To Pay
UK	United Kingdom
USA	United States of America
WSA	Water Service Authority
WTP	Willingness To Pay

CHAPTER 1

AN OVERVIEW OF PERFORMANCE BENCHMARKING IN THE SOUTH AFRICAN PUBLIC SECTOR

1.1 Background

Performance benchmarking in the public sector has endured a rough ride to prominence. In the 1970s generating interest in public sector performance comparisons was a difficult undertaking. At that time the typical arguments used by government officials against performance benchmarking were that their organisations were unique and thus any comparisons would ultimately be of no use (Ammons, 1999) and that public sector organisations generated additional welfare benefits over private firms, ones for which they were not reimbursed.

The uniqueness argument is that one cannot learn by comparing the activities of different service providing organisation within the public sector, because their inputs and outputs are too different. The uniqueness-makes-it-incomparable argument weakens (but does not collapse) as the outputs and inputs of production become more similar. In the case of public and private water service providing enterprises, the outputs and inputs are very similar, but that said, there almost always remains some degree of uniqueness in the inputs and outputs, because of differing raw water scarcity, topography, number, type and distribution of customers served, and so on.

A welfare argument against extending private sector type performance benchmarking to the public sector is that public sector enterprises contribute to social welfare in many additional ways to private sector ones. Public sector enterprises frequently hire more workers than is efficient in order to create more employment and render additional services (unreported output) for which they are not reimbursed by those who benefit. As a result costs are inflated and the value of the output is under-recorded (implying that the so-called inefficiency is intentional). For this welfare argument to be sustained in the case of water services providing institutions, both the unreported output and the deliberate excess employment numbers need to be established.

The successes in applying performance benchmarking in the private sector have been impressive. The successful use of benchmarking by the Xerox Corporation in the early 1980s showed that organisations need not be alike for them to be able to learn from each other (Ammons *et al.*, 2001). In this case, the Xerox Corporation learnt valuable warehousing and distribution lessons from a benchmarking project undertaken with the catalogue merchant L. L. Bean (Dorsch and Yasin, 1998). The disparity in the two business types, in this case, did not preclude them from learning valuable practices from each other (Ammons, 1999). The use of performance benchmarking in this manner undermines the "uniqueness argument" (Ammons *et al.*, 2001, pg. 100) used by public sector officials to rule out their involvement. As could be expected performance benchmarking and measurement in the public sector began gaining ground in the early 1990s (Bruder and Gray, 1994) and is currently employed by many public sector organisations as a tool to improve performance, create more accountability and meeting external (funder and audit) requirements (Tillema, 2010). Chapter 1 will review the application and history of performance benchmarking in South Africa, with particular reference to water services.

1.2 Performance Benchmarking in the Public Sector

1.2.1 Introduction

Bruder and Gray (1994) define benchmarking as a practical method of comparing an organisation's performance against other organisations that are considered to be best-in-class. More formally, benchmarking is defined as "a continuous, systematic process for evaluating the products, services, and work processes of organisations that are recognised as representing best practices for the purpose of organisational improvement" (Spendolini, 1992; Dorsch and Yasin, 1998).

Simply put, benchmarking is the process of identifying the practices of other organisations which lead to best-in-class performance and then learning from them (Folz, 2004). There has been extensive use of benchmarking in the private sector (Dorsch and Yasin, 1998) sometimes with much success. The example of the Xerox Corporation's use of benchmarking to improve productivity and reduce machine defects and service labour costs by substantial amounts is often cited (Mittelstaedt 1992). The transfer of this success from the private sector encouraged to the public sector only happened gradually (Dorsch and Yasin, 1998). Benchmarking initiatives in the public sector only gained critical mass toward the end of the 1990s, but similar type practises had already been utilised in local governments and the public sector for some time before this (Bowerman and Ball, 2000). Some contend that benchmarking type practices were being used local authorities even before they became popular through their use in the private sector (Bowerman and Ball, 2000: 22).

Ammons *et al.* (2001) identifies three different types of benchmarking practices used to improve and measure performance in the public sector:

- *Corporate style benchmarking*:
  Corporate style benchmarking focuses attention on a specific process and endeavours to find the best-in-class performers of that process. Once the best-in-class performer is found, the differences between the current method and the best in class method are identified and adaptations are made to bring the inferior process in line with its best-in-class counterpart. As the name suggests, this type of benchmarking technique is typical of private sector organisations and far less common in the public sector.
- *Targets as benchmarks*:
  Targets as benchmarks is like strategic planning where the desired targets, such as targets for sanitation and water supply, represent the yardstick against which organisations and their progress can be measured. This form of public sector benchmarking is the most widely publicised of the three.
- *Comparison of performance statistics*:
  Comparing performance statistics entails the comparison of one's own statistics against external performance standards or norms (benchmark statistics) as revealed by a number of sections within an organisation or by a number of different organisations. The benchmarks are compiled by collecting performance statistics for a number of differing functions within the organisation or across different organisations. The key objective of this type of benchmarking is the tabulation and subsequent comparison of performance statistics. This type of benchmarking practice is the most commonly used one in the public sector.

Although using benchmarking in the public sector can improve the cost-effectiveness of services provided and employee morale (Bruder and Gray, 1994), it will not do so on its own (Ammons *et al.*, 2001: 109). It must be part of a systematic and continuous process of monitoring and managing (Mittelstaedt, 1992). The process of benchmarking in the public sector can be broken down into seven steps (Bruder and Gray 1994, the interested reader is referred to their article for a detailed description). They start with identifying performance criteria and end with ascribing accountability and linking incentives to performance.

A critical part of benchmarking in the public sector is accountability. In a study of benchmarking activities in the Netherlands water industry, Braadbaart (2007) showed that that while benchmarking had an immediate effect on the transparency of the public sector organisations, improvement in performance was only evident after the results of the benchmarking initiatives were made public and a sense of accountability to the public was created.

1.2.2 The Use of Performance Benchmarking in the Public Sector

World-wide, public sector organisations and municipalities are under pressure to improve quality, efficiency and service provision – partly due to tight budgets, partly due to new environmental laws and partly due to local residents demanding improved services (Burke 2004). In response to this pressure public sector organisations have felt the need to produce evidence that they are 'delivering' – and this is why many have embraced benchmarking (Braadbaart, 2007). The purpose is to show that they are indeed delivering quality services to their residents.

In Canada, a large scale benchmarking project has been undertaken in Ontario state. The Initiative, named the "Ontario Municipal Benchmarking Initiative (OMBI)", collects in excess of 850 measures across 37 municipal service areas (OMBI, 2011). This benchmarking initiative provides performance measures for numerous municipalities in Ontario, and provides reports on various sections of local government ranging from water and waste-water services to taxation.

OMBI provides a platform for local authorities to measure their progress and performance in various areas in comparison with others, and uses this information to further maximise the efficiency of their services provision. It is an example of a large scale public sector benchmarking initiative which has been well developed and now provides participants with valuable service. The reports of each benchmarking project are available to the public, so as to create a sense of accountability in each of the sectors and municipalities (OMBI, 2011).

In the Netherlands, the water industry participates in a nation-wide benchmarking initiative managed by VEWIN (Braadbaart, 2007). VEWIN, the Dutch association of water supply companies, employs professional consultants to conduct the benchmarking practice and the results are made available publicly. In this way, as with the OMBI, each association is made accountable to the public (Braadbaart, 2007).

These are just two of a multitude of benchmarking initiatives and programmes being conducted around the world. Benchmarking initiatives in the in the USA and UK public sectors are also prominent. In England and Wales local authorities are required to benchmark their services (Triantafillou, 2007). In the USA numerous public sector benchmarking initiatives have taken place such as the North Carolina Benchmarking Project and the ICMA Centre for Performance Measurement (Ammons *et al.*, 2001), to name only two.

1.2.3 Using Performance Benchmarking in South Africa

Discontent with service delivery in South Africa has been evident in recent years from the incidence of public protests with poor service delivery and poor governance (Tait and Marks, 2011). In a recent analysis of community protests in South Africa by Karamoko and Jain (2011) the main contributors to public protests were found to be lack of sufficient housing and electricity, no access to clean water, lack of service delivery (in general) and inadequate sanitation systems.

The study of Karamoko and Jain (2011) showed that the regularity of protests has decreased slightly since 2010, but the proportion of violent protests has increased each year. Sebugwawo (2011) contends that social activists believe that the unrest not only stems from poor service delivery, but also from the local government not engaging the public in political processes. Along a similar vein, Tait and Marks (2011) argue that the inactivity of the governments in response to the protests has caused a reduction in the faith the public has in the government to provide services (and in general).

South Africa's Minister of National Planning, Trevor Manuel, has openly acknowledged there are real problems at the local government service delivery level "we must accept that despite the adequate allocation of funding, we fail to deliver quality service to the poor" and he has encouraged the members of the South African parliament to demand better information from government officials (Sapa 2011). How is greater accountability to be s created within the public service sector? As is the case in several other countries the performance of municipalities will have to be measured and monitored, this information will have to be shared with the public and employment/deployment contracts will have to be linked to this performance.

There is evidence that the governing political party in South Africa has serious intentions to move in this direction. The Eastern Cape section of the African National Congress (ANC) announced their intention to require all deployed leaders to sign performance contracts with members in (Sebugwawo, 2011). Unless this performance is properly defined it will lack substance. This is where benchmarking performance will play a vital role.

However, benchmarking in the public sector in South Africa is not a new endeavour. There are three examples of benchmarking type initiatives which have been and are being undertaken in South Africa with the focus specifically being on the provision of water and sanitation services. These initiatives and the rationale behind their implementation in the South African public sector are discussed in later sections.

1.3 Performance Assessment and Benchmarking in South Africa

Performance assessment and benchmarking has been conducted in South Africa in varying forms. Initially the National Benchmarking Initiative (NBI) which ran from 2004 to 2007 sought to compare the performance of municipalities in South Africa. Performance assessment which focused on the provision of potable water and the removal of waste water was introduced by the Blue and Green Drop Certification process in 2008. Most recently a second benchmarking initiative, the Municipal Benchmarking Initiative (MBI) is set to begin in 2012. All of these initiatives focus on the provision of water related services.

WRC Pilot Initiative: 2003-2005

The National Benchmarking Initiative (NBI) was preceded by a pilot study conducted by the Water Research Commission (WRC) between July 2003 and February 2005. This pilot study allowed the WRC to set up an infrastructure and define the performance indicators to be used in a benchmarking initiative. During this project contact was made with all the participants through workshops and visits to the individual organisations. These contact sessions allowed for feedback, training and sharing of knowledge. The study provided a detailed description of all the performance indicators which were expected of the participating municipalities and a method to calculate the indicators. An alarming outcome of this pilot study was that even with a reduced list of required performance indicators, only one municipality was able to provide the necessary data before the termination of the project on 15 February 2005 (Pybus *et al.*, 2005).

National Benchmarking Initiative (NBI): 2004-2007

The National Benchmarking Initiative (NBI) was an initiative which aimed to benchmark the provision of water services by municipalities across South Africa. Three annual benchmarking projects were undertaken between 2005 and 2007 as a joint initiative between the then Department of Water Affairs and Forestry (DWAF), the Water Research Council (WRC) and the South African Local Government Association (SALGA). Learning from the success of Canadian benchmarking initiatives, the aim of this initiative was to use benchmarking to improve the provision of water services by the South African local water service providers. DWAF, SALGA and WRC (2005) identified four main challenges faced by the South African water sector in the short to medium term, and implemented this benchmarking initiative in an attempt to address them. The challenges faced by the water sector were identified as (DWAF, SALGA and WRC, 2005):

- increase the public access to sanitation and waters services while maintaining its affordability,
- ensure that all process and service provided are sustainable,
- create the capacity to ensure that the above two challenges are met, and
- enhance the performance of the water service providers so that the above challenges are met in the most efficient manner.

The successes of benchmarking initiatives in other countries suggest that benchmarking is a useful tool to meet and address these challenges, so as to improve the performance and enhance the sustainability of water service provision in South Africa (Moraka *et al.*, 2011).

The performance indicators which were monitored for the initiative were (DWAF, SALGA and WRC, 2005: 2):

- access to a basic water supply service,
- access to a basic sanitation service,
- drinking water quality,
- impact on the environment,
- strategic asset and water demand management,
- customer service standards,
- financial performance, and
- institutional effectiveness.

The level of performance in each of these areas was determined through various quantitative measurements taken for each indicator. The NBI, which ran over three consecutive years, had mixed results. Each initiative is briefly discussed in what follows.

2004/2005 Initiative

The report of the 2004/5 initiative is the *National Water Services Benchmarking Initiative:Promoting best practice. Benchmarking outcomes for 2004/5* (DWAF, SALGA and WRC, 2005). The plan for the NBI was to collect the data and information on performance measures in the months of July and August. These months coincide with the end of the municipal financial year (June 30). The municipalities which took part in the initiative were divided into three categories according to their size: metropolitan, district and local.

In this first implementation of the NBI a service provider was not appointed until September 2005. This failure resulted in a severe time constraint when collecting data from the participating municipalities – only 22 of the 40 municipalities were able to respond on time, and the responses were not adequately verified, nor complete.

The resultant report did not assess the individual performances of the participating Municipalities, but rather assessed the results in each category anonymously. It concluded that many of the municipalities indicated that they lacked the internal capacity to participate in the initiative adequately. It found that less than 30% of all the collected information was "reliable" or better, and the remainder either "estimates" or "not stated".

It was not possible to draw conclusions on performance based on data of this nature, but many lessons were learnt regarding capacity of municipalities to participate in benchmarking initiatives, as well as some of the hurdles to overcome during the implementation of the initiative.

The results of the benchmarking initiative and subsequent drawbacks were discussed in detail in four breakaway sessions during the annual benchmarking conference. The results of these sessions were provided in detail in DWAF, SALGA and WRC (2005).

2005/2006 Initiative

The report of the 2005/6 initiative is the *National Water Services Benchmarking Initiative: Promoting best practice. Benchmarking outcomes for 2005/2006* (DWAF, SALGA and WRC, 2007). The implementation of the second round of the NBI yielded improved results. In this implementation of the NBI the participating municipalities were divided into the same categories as in the 2004/2005 implementation, but the local municipalities were further sub-categorised according to size. These four sub-categories were: secondary cities, large towns, small towns (mainly urban) and small towns (mainly rural).

In this round 70 municipalities were identified, of which 48 responded – a 13% increase in the proportion of respondents responding in the first round. In this round more indicators and context data were also requested from the participating municipalities – despite the finding of the 2004/2005 NBI report (DWAF, SALGA and WRC, 2005) that many of the participating municipalities lacked the capacity to complete the original questionnaire.

The water service authorities (WSA's) participating in this round of the NBI accounted for approximately 86% of the entire population of South Africa. By way of comparison, the highly successful Canadian municipal benchmarking initiative covered 60% of the population.

Although the number of municipalities which participated increased in the 2005/6 initiative, it was still burdened by unreliable and insufficient information. Only three participating WSA's provided data which (on average) was considered reliable or better and, in total, 69% of the required information was returned by the WSA's. It was deduced that there remained a lack of capacity within the individual municipalities but that the situation had improved. When comparing the progress of the NBI with that of the Canadian initiative, it was deduced progress was being made.

However, owing to the large number of estimated data in the responses, it was felt that caution should be exercised in drawing deductions from the data, and the aim of the initiative was changed to look for overall patterns and trends in the data sets. The results of each section, along with the level of confidence were reported for each key performance indicator. From this information best-in-class performers were identified for each category of municipality. The best performers were identified using an overall points score for five performance areas: access to basic services, drinking water quality, effluent quality, asset management and financial performance. The best performers (by category) were:

- Metropolitan: Johannesburg
- District: Ugu
- Local: Mangaung.

At this stage of the benchmarking process, experience from the Canadian initiative indicated that all the effort should be channelled into "data collection" and "refinement", and that "internal review" only be implemented in the subsequent roll out of the initiative (DWAF, SALGA and WRC, 2007: 9).

At this stage of the initiative the type of benchmarking being undertaken would be the "comparison of performance statistics" (Ammons *et al.* 2001, and discussed in Section 1.2.1).

2006/2007 Initiative

The report of the 2006/7 initiative is the *National Water Services Benchmarking Initiative: Report on 2006/2007 Performance* (DWAF, SALGA and WRC, 2008). The third and last round of the NBI was (as expected) secured the highest level of participation and best quality of data. The benchmarking categories of municipal water service authorities identified were: metropolitan, district, secondary cities, large towns, small towns and rural areas, with the last three areas falling under the broader heading of "local" municipalities.

There were 67 municipalities out of a total of 84 that participated in the third round of the NBI, an 11% increase in the proportion of participating municipalities over the second round.

In addition, 3% more data was collected from each municipality and the data provided was assessed to be more reliable – the municipalities were learning by doing.

As the second rounds before it had, the third round of benchmarking required additional performance indicators and data from the participating municipalities. It found that there was institutional commitment to the programme, but queried whether the initiative was sustainable on the grounds that high staff turnover within the municipalities quickly eroded the learning gained from the year before and the total dependence of the initiative on central government funding.

At this juncture in the Canadian initiative an internal review was implemented. This review was left to be addressed DWAF, SALGA and WRC (2008: 4). Notable conclusions drawn out of the third round of the NBI were that public access to water and sanitation services remained a substantial challenge, drinking water quality and the treatment of waste water needed to be addressed and there was an overall concern about performance in the areas financial performance and asset management. The report also identified accountability and the role of human agency as gaps in knowledge which needed to be addressed.

This round of benchmarking saw the use of a more advanced method (which was moderated by data confidence/reliability) to determine the best performing WSA's in each of the five categories. The method also allowed for the overall ranking of each of the WSA's within each category. This allowed for municipalities to identify their areas of weakness and strengths and to identify the best-in-class performers in their peer group. The best performers per category were:

- Metropolitan: Johannesburg
- Districts: Ilembe
- Secondary Cities: Drakenstein
- Large Towns: Knysna
- Small Towns: Gamagara.

The report backs to the municipalities provided them with insight into how to improve their functionality – one of the primary objectives for engaging in performance benchmarking. The report concluded that a number of additional performance measures should be included in the next round of benchmarking.

What was learnt?

The addition of performance measures round of benchmarking did not happen, but during 2010 SALGA and the WRC conducted a critical review of the NBI (Moraka *et al.* 2011). On the positive side, it was deduced that the participating municipalities considered the exercise to have much merit – that there was indeed useful insight that could be gained by comparing performances with each other and through the assessment process the municipalities where able to identify shortcomings in their data management systems. The NBI also made the municipalities aware of performance measures which were previously not considered by them in this light and encouraged them to develop the necessary systems to generate performance indicators. On the negative side, several municipalities reported feeling that the initiative merely imposed data gathering obligations on them with no real return benefit. The feedback from the co-ordinators of the initiative was also criticised as lacking guidance. The feedback provided consisted of a document of statistics. There was no direction provided as to how the poorly performing participants could improve their performance.

The expectation of participants in benchmarking initiatives is typically higher than what the initiative can achieve (Ammons *et al.* 2001). In the face of these high expectations, even noteworthy achievements may be overlooked or dismissed. Relating to benchmarking exercises undertaken in other countries, Taylor (2009) found that participants in a performance measuring exercise will often only want to use performance indicators for the purpose of meeting the reporting requirements – they do not really want to go the expected extra mile of committing the effort to improving their overall performance. This reluctance may be because there are other constraints imposed on them (funding and personnel).

The lessons learnt through the implementation of the NBI were (Moraka *et al.*, 2011):

- the data management within municipalities needs to be strengthened,
- the initiative should provide practical and tangible benefits to the municipalities by focusing
- the benchmarking initiative on how to improve performance, and
- data collection should use a "less-is-more" approach and reporting should be captured online with hands on support provided to municipalities.

The lessons learnt from the implementation of the NBI were intended to be used to re-launch a national benchmarking initiative for water services by the WRC and SALGA under the title of the Municipal Benchmarking Initiative (MBI) – a project is discussed further in Section 1.3.4.

1.3.3 Blue and Green Drop Certification: 2008-Present

The provision of adequate water and wastewater services is imperative to both economic and social development in South Africa. Following the international trend of using specialised agencies to regulate the provision of water services, the Department of Water Affairs (DWA) introduced a Regulation Strategy for the South African Water Sector. The purpose of this regulation was to protect consumers of water services by clarifying the requirements and responsibilities of water service authorities (WSA's). On 11 September 2008 an Incentive-based Regulation was introduced to the water sector. This regulation brought about a Blue Drop Certification Programme to assess drinking water quality and a Green Drop Programme to assess the management of wastewater (DWA, 2011a).

These programmes aim to evaluate municipal drinking water and wastewater treatment systems and rank or certify the individual performance of each municipality. They fall into the "targets as benchmarks" method of benchmarking as outlined in Section 1.1. Specific targets of minimum standards must be met and the benchmarking process is intended to assist the poor performers to reach these standards. The programmes are also based on an incentive based regulation programme uniquely developed for the South African Context DWA (2011a). Incentive based regulation is defined as "*the conscious use of rewards as well as penalties to encourage performance excellence and continuous improvement, based on an innovative performance rating system*" DWA (2011a, pg. 1-3).

The results of the annual assessments are made publicly available to increase accountability and allow the public access to credible information.

Blue Drop Certification

The Blue Drop certification programme has gained increasing publicity in South Africa as more and more of its results are made public. Over time the requirements to obtain a Blue Drop status have become more stringent. All municipalities are coverage by the programme, so enabling the DWA to accurately identify strengths and gaps in the area of water service provision (The Water Wheel, 2011b).

The participation of municipalities is not voluntary. Municipalities which are WSA's are compelled to provide all the required information and perform a proper analysis of the quality of water services and performance (DWA, 2011a). The assessment process involved in the Blue Drop programme is a thorough investigation into the ability of municipalities to provide drinking water at a sufficient standard. Each municipality is assessed according to five key performance areas (KPA's), namely:

- water safety planning,
- drinking water quality process management and control,
- drinking water quality compliance,
- management, accountability and local regulation and
- asset management.

The performance of each municipality in each of the areas is assessed and scored, when combined a total score out of 100 is given to the municipality. For a municipality to be awarded a Blue Drop Certification a total score of 95% or higher must be achieved (DWAF, 2009). Each of these areas is assessed in order to determine the level of each of the key performance indicators (KPI's) in Table 1. For a detailed explanation on how each of the KPI's are assessed the interested reader is referred to *The Blue Drop Handbook* (DWA, 2011a).

Table 1.1: Key Performance Areas and Indicators considered in Blue Drop Certification

Key Performance Area	**Key Performance Indicator**	**KPA %**
Water Safety Planning	Water Safety Planning Process Risk Assessment and Review of Control Measures Risk-Based Monitoring Programmes Credibility and Submission of Drinking Water Quality Data Incident Management	30-35
Drinking Water Quality Process Management & Control	Compliance with Regulation – Works Classification Compliance with Regulation – Process Controller Registration Availability of WTP logbook	10-15
Drinking Water Quality Compliance	Compliance per Determinand (according to Monitoring Programme) Risk Assessment Defined Health Index Operational Efficiency Index	25-30
Management, Accountability & Local Regulation	Management Commitment Publication of Performance Service Level Agreements/Performance Agreements	10-15
Asset Management	Annual Process Audit Asset Register Availability & Competence of Maintenance Team Operations and Maintenance Manual Maintenance and Operations Budget and Expenditure Design Capacity vs Operational Capacity	15

Source: DWA, 2011a: 2-2

The Blue Drop assessment is solely focused on the performance of a municipality with respect to the quality of the potable water service they render to the community they are mandated to serve. It encourages municipalities to improve the quality of their product. The mechanism by which the scheme works is point incentives. The Blue Drop programme awards bonuses and issues penalties to municipalities according to their performance.

Bonuses are awarded in various KPA's in the form of additional points on a Blue Drop assessment. These bonuses are given awarded to municipalities who are top performers and those that perform tasks over and above the requirements of the Blue Drop Programme. Similarly, penalties are issued for those that fail to meet the requirements, falsify records or fail to report incidents. The penalties take the form of a reduction in the score of a Blue Drop assessment. The methodology of issuing bonuses and penalties is described in *The Blue Drop Handbook* (DWA, 2011a).

As one can see by the areas assessed in Table 1 the performance assessment is thorough. It is intended that obtaining a Blue Drop Certification be a prestigious and sought after status for municipalities – something to boast about (to show that a good job is being done). In 2010 38 potable water systems were awarded a Blue Drop Certification while in 2011 66 systems were (The Water Wheel, 2011b) – suggesting municipalities have embraced the programme and have been striving for improved certification performance.

Green Drop Certification

Introduced alongside the Blue Drop Programme in 2008, the Green Drop Certification Programme assesses the ability of a WSA to adequately collect, treat and discharge wastewater (DWAF, 2009). There are more than 800 (in fact 821) wastewater treatment facilities in South Africa (The Water Wheel, 2011a). These treatment facilities and the corresponding infrastructure are essential for the provision of adequate sanitation services to the South African public.

Like the Blue Drop programme, the Green Drop certification programme uses an incentive based regulation strategy, but makes participation is mandatory. The process (for both programmes) involves the regular assessment of wastewater treatment facilities to ensure effective quality management and compliance with regulations. The goal of these initiatives is to encourage individual municipalities to strive towards improved performance. These assessments are conducted according to Key Performance Areas (KPA's) and the publication of the results provides feedback to the municipalities so that they might assess their own performance and enhance their accountability. Each assessment is intended to aids in the setting of targets and criteria to guide improvement in the service rendered in the following cycle (DWA, 2011b).

The Green Drop certification programme assesses the ability of each wastewater treatment facility to adequately collect, treat and discharge wastewater. This ability is assessed according to the following eleven KPA's:

- operational staff,
- wastewater quality monitoring,
- wastewater sample analysis,
- submission of wastewater quality results,
- wastewater quality compliance,
- management of wastewater quality failures,
- storm-water management,
- by-laws,
- wastewater treatment capacity,
- publication of wastewater quality performance and
- wastewater asset management.

The performance of each wastewater treatment facility in each of these KPA's is assessed and rated resulting in a total performance score out of 100 for each facility. For a facility to be awarded a Green Drop Certification a total weighted score of 90% or higher must be achieved (DWAF, 2009). Each of these areas is assessed determining the level of each of the key performance indicators (KPI's) in Table 1.1.

The weights associated with each KPA are not included in the table as they are variables that are intended to be changed with each implementation of the programme. These changes serve to 'raise the bar' making the absolute requirement to achieve Green Drop certification more difficult with each assessment cycle. An asterisk (*) in Table 2 indicates the number of opportunities which are available for the issuing of bonuses/penalties in each KPA. For a detailed explanation on how each of the KPI's are assessed and the procedure for issuing bonuses and penalties, the interested reader is referred to *The Green Drop Handbook* (DWA, 2011b).

Table 1.2: Key Performance Areas and Indicators considered in Green Drop Certification

Key Performance Area	**Key Performance Indicator**	**Bonus**	**Penal-ty**
Operational Staff	Compliance with Regulation (Works Classification) Compliance with Regulation (Process Controller/Supervisor Registration) Availability and Competence of Maintenance Team Operation and Maintenance Manual	**	
Wastewater Quality Monitoring	Operational Monitoring Compliance Monitoring		
Wastewater Sample Analysis	Proof and name of laboratory used Credibility, Traceability and Submission of Wastewater Quality Data Use of Monitoring results to Amend/Improve Process Controlling	*	
Submission Of Wastewater Quality Results	WWQ Data Submission to DWA		*
Wastewater Quality Compliance	Effluent Quality Limits Known and Available Compliance with Effluent Quality Categories Wastewater Risk Assessment Plan in Place and Annually Reviewed	**	**
Management Of Wastewater Quality Failures	Wastewater Incident Management Quality Protocol Evidence of Implementation of Protocol		*

Storm-Water Management	Storm-water Management Plan Water Demand Management Plan	
By-Laws	By-Laws By-Laws Enforcement Management Performance and Commitment	*
Wastewater Treatment Capacity	Documented Design Capacity (Hydraulic and Organic) and Documented Daily Receiving Flows Medium to Long Term Plans (Capacity for Treatment System, Effluent Quality Compliance) Medium to Long Term Plan (Capacity for Collecting System)	
Publication Of Wastewater Quality Performance	Annual Publication of Wastewater Management Performance Public Publication of Performance	
Wastewater Asset Management	Annual Audit Report Asset Register Operation and Maintenance Budget Operation and Maintenance Expenditure Maintenance Records	*

Source: DWA, 2011b: 27-29

The inclusion of municipalities in the programme has rapidly increased since it was implemented. In 2009 there were 449 municipal treatment works that participated in the programme, while by 2011 all 821 municipal wastewater treatment works were included in the programme (The Water Wheel, 2011a). However, only 40 (5%) Green Drop certifications were issued in 2011, 19 in the Western Cape; suggesting the improvement of wastewater treatment facilities to be an area of great current concern.

1.3.4 Municipal Benchmarking Initiative (MBI): 2012-Present

Following the lessons learnt from the initial National Benchmarking Initiatives (2004-2007) for Water services, SALGA and WRC launched an new and improved benchmarking initiative in 2011 (Moraka *et al.*, 2011), known as the Municipal Benchmarking Initiative (MBI). This new initiative aims to take advantage of the maturing regulation tools currently used in the South Africa water sector. In order to encourage buy in by the municipalities it is marketed as "benchmarking for municipalities, by municipalities, to the benefit of municipalities" with the declared aim being performance improvement of water service delivery (Moraka *et al.*, 2012).

It is motivated on two grounds:

- that the South African local government system is under strain to meet the necessary government mandates regarding the provision of safe and effective water services (Moraka *et al.* 2011) – as reflected in the Green Drop certification results, and to a lesser extent, the Blue Drop certification results
- that benchmarking along the lines explored through the NBI, when done effectively, is a key tool to the optimisation of service quality and provision at the municipal level.

A key requirement of this revitalised imitative is the ability of the municipalities to perform introspective performance benchmarking, as distinct from regulatory assessment of the Blue and Green Drop type (Moraka *et al.*, 2011). It has four main aims:

- use comparative performance benchmarking to facilitate improved efficiency and effectiveness of water service provision,
- enhancing the performance measurement, monitoring and management,
- establishing communities of good practice within (and between) municipalities, and
- building relationships of mutual trust and respect between municipalities (Moraka *et al.*, 2012).

The initiative entails two types of benchmarking. The first is called "metric benchmarking" (Moraka *et al.* 2011) – and is essentially a 'comparison of performance statistics' benchmarking approach (Ammons *et al.* (2001) see Section 1.1). This type of benchmarking allows for reflection on a municipalities performance when compared with previous performances, or the performance of others. This form of benchmarking is good for the ranking of municipalities according to their performance. However, it does not offer support in the way of providing reasons as to why the performances differ (Moraka *et al.*, 2011). It is very much in the tradition of the blue and green drop certification approaches.

Where it is different is that it explicitly incorporates a second type of benchmarking called "process benchmarking" (Moraka *et al.* 2011). This component of the initiative incorporates the actual process involved in performing a specific task or function. It fits into the corporate style benchmarking approach (Ammons *et al.* 2001, see also Section 1.1). It is an approach that aims to allow poorly performing municipalities to learn from top performers through a detailed investigation into the processes used by the top performer. The lessons learnt are then adapted and used by the poor performers to improve their performance.

1.4 The international use of efficiency analysis

Using various methods efficiency analysis has been implemented in many regions of the world, including both first and third world countries. In Africa, Singh et al. (2011) reported that efficiency studies of water service provision using data envelopment analysis were implemented in Uganda and Ghana in 2007 and 1995 respectively. Singh et al. (2011) implemented a similar technique to evaluate the efficiency of water utilities in the urban areas of Northern India. Much research has been done into the efficiency evaluation of water services providers in the UK by Thanassoulis (2000a) and Thanassoulis (2000b), where the use and application of efficiency measurement techniques common.

1.5 A case for incorporating efficiency into South African benchmarking

The main attention of public sector benchmarking in South Africa is focused on benchmarking/evaluating the provision of water services by municipalities. The initiatives which have been undertaken seek to establish a system whereby the performance of a municipality can be used to rank each municipality in comparison with others.

Benchmarking of this kind is good practice for many reasons. The comparison and publication of performance measurement encourages municipalities to perform well so as to not be ranked near the bottom and to meet predetermined quality standards (as with the Blue and Green Drop programmes). This practice of collecting performance data within an institution enhances the skill level of the staff, and when used effectively benchmarking can be harnessed to enhanced performance in specific areas through "process benchmarking".

These positives are very important, but there still appear to be some gaps that have remained unfilled in the South African adoption of benchmarking. The current benchmarking approaches allow one to identify the top municipal performers, but they do not provide any indication of the cost effectiveness of the water services provided by the municipalities to their clients/customers/users.

Suppose two municipalities both achieve a high relative level of service provision, but one provides this service at half the (relative) cost. The service provider which supplied the water at a lower cost is more efficient, but the current benchmarking initiatives do not incorporate this crucial information. This is very important information that is being omitted in the benchmarking exercises. It is not only the relative quality of provision that is a benchmark of water service provision but also the level of efficiency by which this water service is provided.

Inevitable as more municipal funding resources are added to water service provision the quality of the provision will improve, but this may be at a cost elsewhere within the municipality – either in municipal functionality or on the residents and businesses falling under that municipality. At the very least it is important that a cost minimisation component be incorporated. Ultimately though, if a benchmarking initiative is to achieve its social objectives, the quality performance assessments must be combined with various efficiency performance assessments, so that a holistic benchmarking measure is generated with respect to water service provision.

A limited but still useful measure of the efficiency of an organisation in production can be measured by how well various inputs, e.g. number of staff, operating costs, are converted into outputs, e.g. volume of water delivered, and volume of water treated. This measure is a cost effectiveness measure (a supply side measure of efficiency). An efficient organisation will minimise their input cost to achieve a given output.

Incorporating this vital piece of information into the evaluating municipal water service performance still needs to be done. It follows that the cost effectiveness of municipalities needs to be benchmarked along with performance. Fortunately, this will not be difficult because this type of "efficiency" benchmarking is not uncommon. Cost effectiveness efficiency benchmarking has been being implemented in the UK, Peru, Uganda, Ghana, Brazil, India, Japan, Palestine, Canada and the USA, inter alia (Singh *et al.* 2011: 88). It is clearly the next logical step to incorporate it into the benchmarking initiative currently being implemented to improve South African municipal water service delivery.

The typical tools used to measure cost effectiveness efficiency are stochastic frontier analysis (SFA) and/or data envelopment analysis (DEA).

1.6 Conclusion

The South African municipal water services providing sector is well suited to performance benchmarking. The uniqueness and welfare counter arguments against performance benchmarking are weak for this sector. There are a large number of separate water service providers and their inputs and outputs are similar.

As could be expected, there is evidence that South African municipal water service delivery is being increasingly monitored, both through regulation enforced monitoring initiatives like Blue and Green Drop certification, and through other active benchmarking projects, such as the Municipal Benchmarking Initiative. That these monitoring exercises serve to highlight areas in which water service provision can be improved is beyond question. They almost certainly are also yielding positive spinoffs in terms of quality of service delivered by, at least, some municipalities.

But good as they are, the current South African municipal water service benchmarking initiatives leave some important performance information gaps, one of which is with respect to the efficiency of this provision. It is a shortcoming that is relatively easy to address. The remainder of the report is structured in the following way. Chapters 2 and 3 respectively, will outline the parametric and non-parametric methodologies for generating efficiency performance measures and benchmarks. Chapter 4 will discuss the sample design and composition. Chapters 5 and 6 respectively will apply the parametric and non-parametric methodologies to generate efficiency performance measures and benchmarks for a sample of metropolitan, district and local municipalities. Chapter 7 will concludes with a policy perspective.

CHAPTER 2

THEORY AND METHODOLOGY OF PARAMETRIC FRONTIER ESTIMATION FOR THE PURPOSE OF GENERATING MEASURES/INDICES OF EFFICIENCY IN WATER SERVICE PROVISION

2.1 Introduction

Many government firms and institutions produce goods or services (output) that are very important, if not essential, and use large quantities of public resources in this production. One such set of institutions that fit into this category are the water service providing institutions. Being an essential service and one that uses large quantities of public resources in production, it follows that the monitoring of their productive efficiency is a socially responsible thing to do.

It is not enough for a water service providing institutions to produce enough output to satisfy the demand for water services. This production also needs to be accomplished by a minimal use of raw materials, labour and other available resources (inputs). A producer's capacity to maximise outputs (services) while minimising their use of inputs (resources) is an indicator of the *efficiency* of the firm (Bogetoft and Otto, 2010). This literature survey will describe the basic idea of econometric efficiency measurement and review two of the most prominent parametric methodologies used for this purpose – stochastic frontier analysis (SFA) and corrected ordinary least squares (COLS).

2.2 Theory of Econometric Efficiency Measurement

A producer is an entity responsible for taking a set of inputs and converting them into valuable outputs. The efficiency of production is a measure of how successful a producer is in generating a maximum level of outputs for a given set of inputs (Farrell, 1957) or as the success of a producer at limiting the amount of inputs used to achieve a desired level of output.

Suppose that there exists N producers and consider a single producer j in $\{1,2,\ldots,n\}$. This producer (and all others) uses k inputs $\boldsymbol{x}_j = \left(x_1^j, x_2^j, \ldots, x_k^j\right) \in \Re_+^k$ to produce a set of n outputs $\boldsymbol{y}_n = \left(y_1^j, y_2^j, \ldots, y_n^j\right) \in \Re_+^n$ (Bogetoft and Otto, 2010). The input requirements set (Greene, 2008) can now be defined as

$$R(\boldsymbol{y}) = \{\boldsymbol{x} \mid (\boldsymbol{x}, \boldsymbol{y}) \text{ is producible}\}, \tag{2.1}$$

and is a set which contains all $(\boldsymbol{x}, \boldsymbol{y})$ such that the vector of outputs, $\boldsymbol{y}$, can be produced from the vector of inputs $\boldsymbol{x}$. In essence this is a broad inclusive set of all feasible combinations of inputs and outputs. Within this set it is desired to identify the $(\boldsymbol{x}, \boldsymbol{y})$ combinations which are considered to be efficient. This set of input-output combinations is known as the efficient production function and which is defined by Farrell (1957, pg. 254) as "the output that a perfectly efficient firm (producer) could obtain from any given combination of inputs". As such, the production function can be defined as (Greene, 2008)

$$T(\boldsymbol{y}) = \{\boldsymbol{x} \mid \boldsymbol{x} \in R(\boldsymbol{y}) \text{ and } \gamma\boldsymbol{x} \notin R(\boldsymbol{y}) \, \forall \, \gamma \in [0,1)\}. \tag{2.2}$$

This function (an isoquant) defines the boundary of the input requirements set and states that any equiproportional reduction in input $\boldsymbol{x}$ (i.e. $\gamma\boldsymbol{x}$) will result in the output $\boldsymbol{y}$ no longer being producible from that input (Ruggiero, 2011). As such the input-output pair $(\boldsymbol{x}, \boldsymbol{y})$ can be considered as technically efficient if $\boldsymbol{x} \in T(\boldsymbol{y})$.

To measure the efficiency of a producer, the most widely used approach is that of Farrell efficiency (Bogetoft and Otto, 2010) (also known as the Debreu-Farrell measure of efficiency) which was suggested by Debreu (1951) and Farrell (1957). The Debreu-Farrell measure of technical efficiency can be defined in terms of input reduction as (Greene, 2008)

$$E(\boldsymbol{y}, \boldsymbol{x}) = \min_{\delta}\{\delta \mid \delta\boldsymbol{x} \in R(\boldsymbol{y})\}. \tag{2.3}$$

The technical efficiency is thus given as the maximum amount that an input vector $\boldsymbol{x}$ can be reduced (equiproportionally) while still being able to produce the output vector $\boldsymbol{y}$. Clearly if $E(\boldsymbol{y}, \boldsymbol{x}) = 1$ then the producer j is considered to be technically efficient. As such, the production function $T(\boldsymbol{y})$ can be redefined according to Knox Lovell (1993) as

$$T(\boldsymbol{y}) = \{\boldsymbol{x} \mid E(\boldsymbol{y},\boldsymbol{x}) = 1\}. \tag{2.4}$$

To illustrate this idea of technical efficiency Farrell (1957) and Greene (2008) considered the example of a producer *j* who uses two inputs $\boldsymbol{x}_j = (x_1^j, x_2^j)$ to produce unit output $\boldsymbol{y}_j = 1$. This scenario is graphically represented in figure 1. The isoquant *AB* in the figure depicts the values of $\boldsymbol{x}$ such that $\boldsymbol{x} \in T(\boldsymbol{y})$ where $\boldsymbol{y} = 1$, and is a representation of the production function. That is, *AB* represents all the combinations of the inputs X_1 and X_2 that a perfectly efficient producer would use to produce unit output (Farrell, 1957).

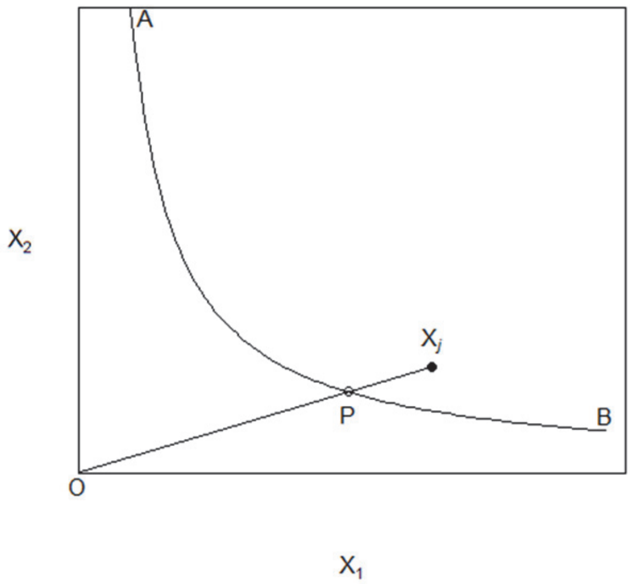

Figure 2.1: Graphical representation of technical efficiency (Farrell, 1957; Greene, 2008)

If the inputs (x_1^j, x_2^j) were each reduced by the ratio $\frac{\|OP\|}{\|OX_j\|}$ this would result in the point X_j being moved to coincide with point *P* and fall on the efficient isoquant *AB*. (Note: the operator $\|\cdot\|$ is a length operator. The term $\|OP\|$ denotes the length of the (straight) line segment between point *O* and point *P* on the graph). The input vector $\boldsymbol{x}_j = (x_1^j, x_2^j)$ can be reduced while still maintaining unit output. As such the point X_j (which represents the input vector $\boldsymbol{x}_j$) is inefficient. Farrell (1957) proposed that the technical efficiency of the producer j could be defined as the ratio $\eta = \frac{\|OP\|}{\|OX_j\|}$ where $0 < \eta \leq 1$. An efficient producer would be able to produce unit output using an input of only $\eta \mathbf{x}_j$. Using this measure for efficiency Farrell (1957) explains that perfectly efficient producers will achieve an efficiency rating of 1, while this value will decrease as the producer moves further away from the efficient isoquant. This approach only considers technical efficiency, without taking the costs of the inputs into account. Figure 2 builds on figure 1 by adding an isocost line, the gradient of which represents the price ratio between input X_1 and X_2 (Greene, 2008). This line is given by *CD* in figure 2.

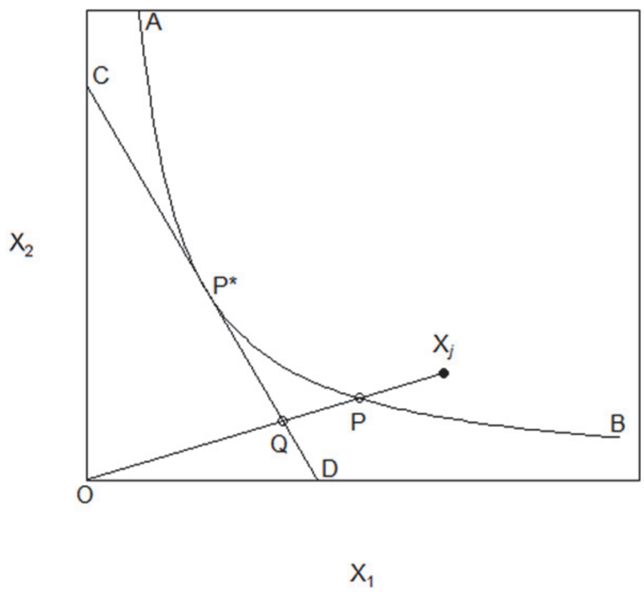

Figure 2.2: Graphical representation of allocative and overall efficiency
(Farrell, 1957; Greene, 2008)

Taking this into account, Farrell (1957) explains that the point P^* and not *P* provides the optimal combination of inputs X_1 and X_2. This is evident as the point P^* lies on the efficient isoquant *AB* and is thus fully technically efficient. However, this is achieved at a fraction $\frac{\|OQ\|}{\|OP\|}$ of the cost of production at point *P*. The price efficiency or allocative efficiency (denoted ρ) of point *P* is thus defined as $\rho = \frac{\|OQ\|}{\|OP\|}$ (Farrell, 1957). As before, a producer who is perfectly efficient in this regard will have an allocative efficiency measure of 1.

Combining the two efficiency measures (technical and allocative) will result in a measure of overall efficiency. Farrell (1957) proposed that this overall efficiency (denoted ω) be calculated as the product of the individual measures, i.e.

$$\omega = \eta \cdot \rho = \frac{\|OP\|}{\|OX_j\|} \cdot \frac{\|OQ\|}{\|OP\|} = \frac{\|OQ\|}{\|OX_j\|} \tag{2.5}$$

Greene (2008) defined the overall inefficiency of a producer as $1-\omega$. Similarly, technical and allocative inefficiency of a producer as $1-\eta$ and $\eta-\omega$ respectively. Using these definitions of efficiency (and inefficiency) producers can be ranked and compared, according to the relevant efficiency measure (technical, allocative, overall).

The approach to efficiency analysis described above is frontier efficiency analysis and is the focus of the remainder of this report.

2.3 Parametric Frontier Estimation

Considering the method of efficiency measurement outlined in Section 2.2 it is important to determine the functional form of the production function (now denoted $f(\cdot)$). This function is taken to be quasi-concave, smooth, continuous and continuously differentiable (Greene, 2008). These assumptions are necessary for the manipulations and derivations that follow. Furthermore Greene (2008) points out that the producers in this case are assumed to be "price takers" and as such all input prices are considered to be exogenous. For the following discussions we consider producers of a single scalar output, *y*, using *k* inputs $x = (x_1, x_2, ..., x_k)$. The production function can now be expressed as (Ganley and Cubbin, 1992)

$$y_i \leq f(\boldsymbol{x}_i, \boldsymbol{\beta}) \tag{2.6}$$

where y_i is the observed output of the producer *i* when utilising the inputs $\boldsymbol{x}_i$. The vector of coefficients, $\boldsymbol{\beta}$, are the parameters which are used to define the function $f(\cdot)$ (Ganley and Cubbin, 1992). Recall that the production function is defined as the maximum amount of output which is achievable from the given set of inputs (Johnston, 1960), and as such the output of the producer *i* reaches its maximum at the production function. (The production function $f(\cdot)$ is thus interpreted as a frontier (Ganley and Cubbin, 1992).) So, assuming that inefficiency is possible among producers, the output y_i obtained by producer *i* using the inputs $\boldsymbol{x}_i$, can only be less than (inefficient) or equal to (efficient) the output predicted by the production function $f(\boldsymbol{x}_i,\boldsymbol{\beta})$. Using an output based Debreu-Farrell measure of efficiency, the technical efficiency for producer *i* can be defined as (Greene, 2008)

$$E(y_i,\boldsymbol{x}_i)=\frac{y_i}{f(\boldsymbol{x}_i,\boldsymbol{\beta})} \tag{2.7}$$

where $0<E(y_i,\boldsymbol{x}_i)\leq 1$. Letting $\xi_i = E(y_i,\boldsymbol{x}_i)$ the model for production can be written as

$$y_i = f(\boldsymbol{x}_i,\boldsymbol{\beta})\cdot \xi_i \tag{2.8}$$

The two most popular forms of the production function (used in econometric efficiency estimation) are the Cobb-Douglas and the more general translog production functions (Greene, 2008). A simple form of the Cobb-Douglas production function for one output and two inputs is (Aigner and Chu, 1968)

$$f(\boldsymbol{x}_i,\boldsymbol{\beta})= Ax_1^{\beta_1}x_2^{\beta_2} \tag{2.9}$$

where x_1 and x_2 are the inputs and A,β_1,β_2 are the model parameters. This model can easily be generalised to include *k* inputs. The translog production function for one output and *k* inputs is defined by Greene (2008) as

$$\ln f(\boldsymbol{x}_i,\boldsymbol{\beta})=\alpha+\sum_{i=1}^{k}\beta_i \ln x_i +\tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}\gamma_{ij}\ln x_i \ln x_j. \tag{2.10}$$

The translog function expands the Cobb-Douglas function by including interaction terms between the inputs, as given by the $\tfrac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{k}\gamma_{ij}\ln x_i \ln x_j$ term in equation 2.10. Using these production functions it is clear that the functions are linear in the logs of the variables, and as such equation 2.8 can be expressed as (Greene, 2008)

$$\ln y_i = \ln f(\boldsymbol{x}_i,\boldsymbol{\beta})+\ln \xi_i \tag{2.11}$$

where $\ln \xi_i \leq 0$ (since $0<\xi_i \leq 1$). Equation 2.11 is often reformulated as

$$\ln y_i = \ln f(\boldsymbol{x}_i, \boldsymbol{\beta}) - u_i \tag{2.12}$$

where $u_i = -\ln \xi_i \geq 0$ provides a measure of technical inefficiency (Greene, 2008). The technical efficiency of producer *i* can thus be expressed as

$$\xi_i = e^{-u_i} \tag{2.13}$$

The next step is to develop an estimate for the value of u_i.

We digress from the measurement and estimation of production efficiency and turn our attention to cost efficiency. The formulation of the cost function is similar to that of the production function discussed above, however the frontier is now a minimum rather than a maximum value. This is intuitive as the effort to maximise production is akin to that of minimising the costs of production. Ganley and Cubbin (1992) express the cost function for producer *i* as

$$c_i = g(\boldsymbol{z}_i, \boldsymbol{\varphi}) \tag{2.14}$$

where z_i are the factors which determine costs (such as prices, quantities, etc.), $\boldsymbol{\varphi}$ is a vector of parameters and c_i is the observed cost for producer *i*. The function $g(\cdot)$ has a similar interpretation to the production function discussed above. Whereas the production function $f(\cdot)$ represents the maximum production, the function $g(\cdot)$ represents the minimum possible cost achievable from the given inputs. As such, the function $g(\cdot)$ represents a minimisation frontier. The ratio of the observed to minimum possible costs (c_i and $g(\boldsymbol{z}_i, \boldsymbol{\varphi})$ respectively) given by

$$\psi_i = \frac{g(\boldsymbol{z}_i, \boldsymbol{\varphi})}{c_i} \tag{2.15}$$

is a measure of the cost efficiency of producer *i* (Ganley and Cubbin, 1992). Similar to equation 2.8 this relationship can be expressed as

$$c_i = g(\boldsymbol{z}_i, \boldsymbol{\varphi}) \cdot \psi_i \tag{2.16}$$

where $0 < \psi_i \leq 1$ is the cost efficiency of producer *i*. There are a number of complications to deal with when estimating cost efficiency, these are discussed in Greene (2008, pg. 110-111). For now the discussion of cost frontiers will be left at this elementary point and revisited in later sections.

2.4 Statistical Model

This section covers the estimation of the efficiency of various producers in relation to a specified production function (frontier). For illustrative purposes, an extension of the Cobb-Douglas production (given in equation 2.9) is used. This production function is used owing to its popularity in the related literature. The approach to efficiency estimation described below can be used in an identical manner for other specifications of the production function.

The production function used expands the formulation in equation 2.9 to include *k* inputs, that

$$f(\boldsymbol{x}_i, \boldsymbol{\beta}) = A\prod_{j=1}^{k} x_{ij}^{\beta_j}. \tag{2.17}$$

This function is estimated using the observations of input and output for each of a set *N* producers. Suppose the observed data consist of output values y_i where *i* = 1,2,...,*N*, and input values x_{ij} where *i* = 1,2,...,*N* and *j* = 1,2,...,*k* for each producer. The model can be expressed according to equation 2.8 as

$$y_i = A\prod_{j=1}^{k} x_{ij}^{\beta_j} \xi_i \tag{2.18}$$

where $0 < \xi_i \leq 1$ is a random disturbance used to determine technical efficiency. Equation 2.18 can be simplified by taking logs on both sides which results in

$$\ln y_i = \ln A + \sum_{j=1}^{k} \beta_j \ln x_{ij} + \varepsilon_i \tag{2.19}$$

where $\varepsilon_i = \ln \xi_i$. There are two schools of thought regarding the composition of the random disturbance term ε_i. The first is that the disturbance term ε_i only comprises of a measure of inefficiency and disregards any measurement error (Ruggiero, 1999). This the deterministic frontier approach as the stochastic term of the model consists entirely of the inefficiency measure, i.e. $\varepsilon_i = -u_i$. The other school of thought is that the disturbance term ε_i comprises of both a statistical noise component, v_i, and an inefficiency component, u_i (Jondrow *et al.*, 1982). This approach is named the stochastic frontier model, or the composed error model, and the error term is constructed as $\varepsilon_i = v_i - u_i$ where the components are independent of each other (Jondrow *et al.*, 1982). Estimation using these models is discussed in further detail in the sections that follow.

2.4.1 The Deterministic Frontier Model

The deterministic frontier model for the Cobb-Douglas production function given in Section 2.2 is given as

$$\ln y_i = \alpha + \sum_{j=1}^{k} \beta_j \ln x_{ij} - u_i \qquad (2.20)$$

where $\alpha = \ln A$. All deviations from the frontier are attributed to the inefficiency of the producer, u_i.

In order to determine the measures of inefficiency the production function, represented by $\alpha + \sum_{j=1}^{k} \beta_j \ln x_{ij}$, in equation 2.20, must first be estimated. The parameters, β_j, of the production function can be consistently estimated using ordinary least squares (OLS) (Greene, 2008). However, the estimate of the intercept term, α, is biased (Ruggiero, 1999). Fitting a linear regression model to the model (equation 2.20) will result in the linear fit illustrated in figure 2.3.

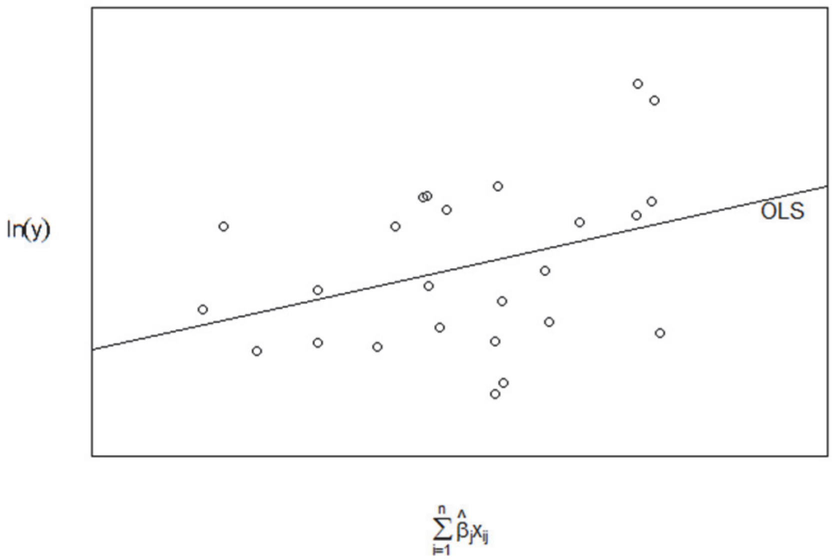

Figure 2.3: Production function estimation using OLS (uncorrected)

Since the model (represented by the OLS line in figure 2.3) does not represent a "frontier" it needs to be corrected. The slope parameters are estimated consistently and (thus) do not need to be corrected. The biased estimate of the intercept parameter is corrected by shifting the OLS line upwards so that all the residuals save for one are negative (Ruggiero, 1999). The non-negative residual is zero, and is the observation on which the production function is "hung" (Greene, 2008). This places the frontier above the observed values of output, thus predicting the maximum achievable production value given the inputs. As such, the frontier constructed using this method satisfies the conditions required of a production function. This process is illustrated in figure 2.4, and is called corrected ordinary least squares (COLS). The COLS line is a consistent estimate of the production function/frontier (Greene, 2008).

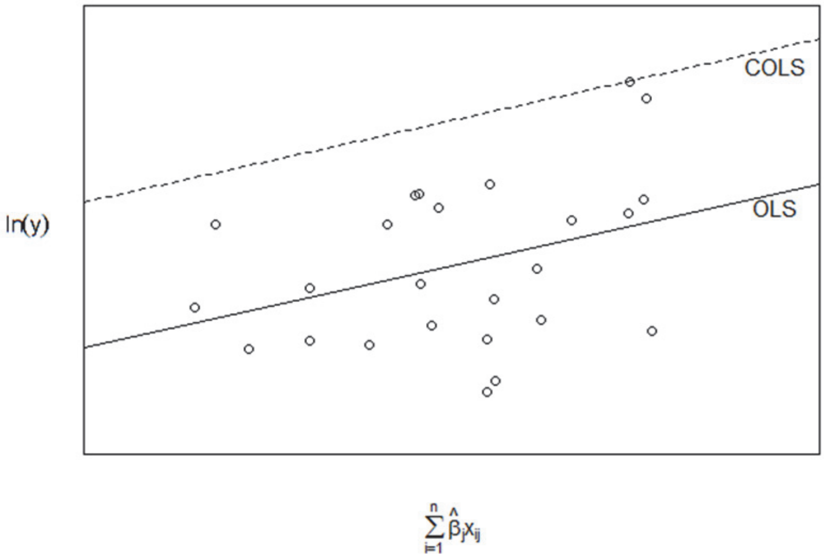

Figure 2.4: Production function estimation using OLS (corrected)

Mathematically, suppose that OLS is used to estimate the production function in equation 2.20. The OLS estimates of the model parameters are given as $\hat{\alpha}$ and $\hat{\beta}_j$ (j = 1,2,…,k). Using the COLS method described above, the estimated intercept term, $\hat{\alpha}$, is adjusted so that all residuals are non-positive. That is

$$\hat{\alpha}_{COLS} = \hat{\alpha} + \max_i u_i \tag{2.21}$$

and the adjusted residual

$$\hat{u}_i = u_i - \max_i u_i \tag{2.22}$$

provides an estimate of the technical inefficiency of producer i (Ruggiero, 1999). The formulation of an efficiency measure using the result of equation 2.22 is covered in section 2.5.

Estimation of a cost frontier is also possible using the COLS method (Greene, 2008). In this case, the dependent variable would be cost (or a function thereof) and the correction to the OLS line would be a downward shift. This allows the cost frontier to lie below all the observed costs, thus providing a minimum value as the theory (discussed previously) requires.

The main drawback of the deterministic frontier model is its treatment of the so called "error" term, $-u_i$. The deterministic specification embeds all measurement error and possible stochastic variation in the dependent variable (e.g. output, cost) into this economic efficiency measure (Greene, 2008). As such, there is a conceptual flaw in the formulation of this model. It does not allow for the possibility of statistical noise or measurement error (Ruggiero, 1999), which is an inevitable part of estimation. Owing to this, applications of this deterministic approach to efficiency estimation are fairly uncommon (Murillo-Zamorano (2004), Greene (2008)).

The stochastic frontier approach overcomes the drawbacks of the deterministic model by allowing for the possibility of statistical noise and measurement error in the model (Murillo-Zamorano, 2004). For this reason, the model used in this case is the stochastic frontier model.

2.4.2 The Stochastic Frontier Model

The stochastic frontier model addresses the shortcomings of the deterministic model by including the possibility of statistical noise as a source of deviation between observed values and the frontier. This model was proposed both by Meeusen and van Den Broeck (1977) and Aigner *et al.* (1977) to account for variations in output measurements of a producer which occurred outside the control of the producer. Using the deterministic approach, any external and uncontrollable factors which effect the output of a producer, e.g. bad weather, strikes, measurement errors, would be measured as inefficiency (Battese, 1991; Greene, 2008). Furthermore, misspecification of the production function, could result in significantly decreased efficiency measures (Greene, 2008) calculated for the producers. The stochastic frontier model for the Cobb-Douglas production function can be represented as (Murillo-Zamorano, 2004)

$$\ln y_i = \alpha + \sum_{j=1}^{k} \beta_j \ln x_{ij} + v_i - u_i. \tag{2.23}$$

The stochastic frontier model differs from the deterministic model in its treatment of the error term. The stochastic model used what is termed a "composed error" formulation, where the term $v_i - u_i$ is a composite error term (Murillo-Zamorano, 2004). It is so named because the overall error in the model is composed of two differing and independent sources for error, v_i and u_i. The error component v_i represents the random statistical noise associated with factors such as bad weather or measurement errors, which are beyond the control of the producer (Battese, 1991). The error component u_i, represents the technical efficiency of the producer.

The statistical noise components (v_i) of the error term are assumed to be independently and identically distributed variables from a zero mean, symmetric distribution with a constant variance σ_u^2. The Gaussian (normal) distribution $N(0, \sigma_u^2)$, as proposed by Aigner *et al.* (1977) and Meeusen and van Den Broeck (1977), is the most commonly assumed distribution for v_i.

The technical efficiency components ($u_i \geq 0$) of the error term are distributed independently of v_i and are assumed to have a one sided or skewed distribution. The most commonly assumed distributions for the u_i variables are the half normal and exponential distributions (Murillo-Zamorano, 2004). For the purpose of this report we will assume a half normal distribution for the u_i terms owing to its prominent use in the related literature (Ruggiero, 1999; Murillo-Zamorano, 2004).

The basic structure of a stochastic frontier can be explained by considering the case of producers using a single input (*x*) to produce a single output (*y*) (Battese, 1991); a scenario depicted in figure 2.5. The line *AB* in the figure represents the deterministic production function $f(x, \boldsymbol{\beta})$. The point Y_i represents the observed output of producer *i* using X_i of input *x*. Similarly, the point Y_j represents the observed output of producer *j* using X_j of input *x*. Both the producers are producing output values below the deterministic frontier *AB*. The stochastic frontier output for producer *j*, Z_j, is higher than that of the deterministic frontier, as this corresponds to the productive activity under favourable external conditions ($v_j > 0$) (Battese, 1991). Similarly, the frontier output for producer *i*, Z_i, corresponds to productive activity under unfavourable external conditions ($v_i < 0$). It is, of course, possible for the observed output (*Y*) and the frontier output (*Z*) to fall above the corresponding deterministic frontier value (Battese, 1991). This scenario would correspond to a highly efficient producer operating under favourable external conditions.

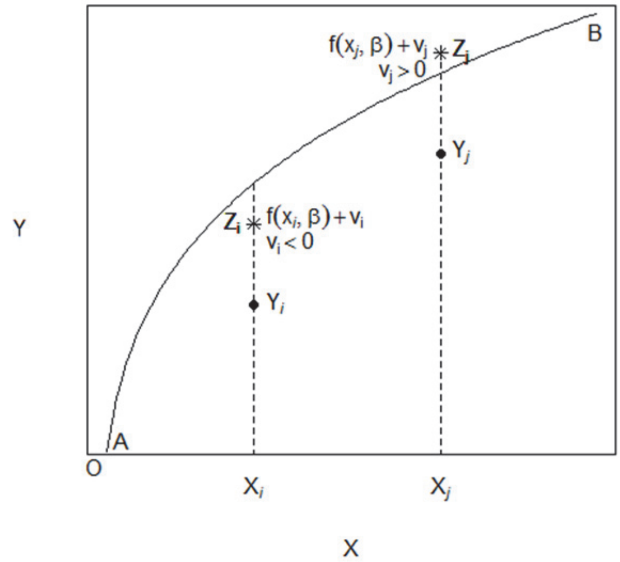

Figure 2.5: The stochastic production frontier (Battese, 1991)

The stochastic frontier (for the Cobb-Douglas production function considered) is:

$$\alpha + \sum_{j=1}^{k} \beta_j \ln x_{ij} + v_i \tag{2.24}$$

with the value of u_i remaining as the measure of inefficiency for producer i (Greene, 2008). As such, the values of $\ln y_i$ (the natural logarithm of the observed production) are bounded above by the stochastic value given in equation 2.24 (Battese, 1991).

The driving objective of economic frontier model estimation is the estimation of the inefficiency measure u_i (Greene, 2008), which is only possible once the distribution of each error component is specified (Ruggiero, 1999). As mentioned above, this report will assumes the “normal-half-normal” specification of the composite error term.

The first step towards estimating the efficiency of producers is to determine the parameter estimates for α and β_j (for j = 1,2,...,k), σ_u and σ_v (Greene, 2008). Maximum likelihood estimation (MLE) and COLS (method of moments) estimation techniques have been introduced for this purpose.

The maximum likelihood estimation technique for the "normal-half-normal" case begins by obtaining the density function for the composed error term $\varepsilon_i = v_i - u_i = \ln y_i - \alpha - \boldsymbol{\beta} \ln \boldsymbol{x}_i$. This density function is given as (Aigner *et al.*, 1977; Greene, 2008)

$$f_\varepsilon(\varepsilon_i) = \frac{2}{\sigma\sqrt{2\pi}}\phi\left(\frac{\varepsilon_i}{\sigma}\right)\left[1 - \Phi\left(\frac{\varepsilon_i \lambda}{\sigma}\right)\right] \tag{2.25}$$

where $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \frac{\sigma_u}{\sigma_v}$ and $\phi(\cdot)$ and $\Phi(\cdot)$ represent the density and distribution functions for the standard normal distribution respectively. The next step is to determine the log-likelihood function for the density function given in equation 2.25. The log-likelihood function is easily determined from the density function and is given as (Greene, 2008)

$$\ln L(\alpha, \boldsymbol{\beta}, \sigma, \lambda) = K - n \ln \sigma + \sum_{i=1}^{n}\left[\ln \Phi\left(\frac{-\varepsilon_i \lambda}{\sigma}\right) - \frac{1}{2\sigma^2}\left(\varepsilon_i^2\right)\right] \tag{2.26}$$

where K is a known constant value. This likelihood function is maximised with respect to each of the variables in order to determine estimates for the parameter values. This estimation is done iteratively using statistical software packages. The values of the estimates are then used to determine the estimates of the technical efficiencies of the various producers.

Stochastic Cost Frontiers

The dual to the stochastic production frontier is the stochastic cost frontier. While the production frontiers provide a maximum level of output, the cost frontiers provide a minimum value of input. The dual cost function for the Cobb-Douglas production function is given by Schmidt and Lovell (1979) as

$$\ln C_i = K + \frac{1}{r}\ln Y_i + \sum_{j=1}^{k}\frac{\beta_j}{r}\ln p_j - \frac{1}{r}(v_i - u_i) \tag{2.27}$$

where $r=\sum_{j=1}^{k}\beta_j$ represents the returns to scale, p_j indicate the prices of the *k* inputs, and Y_i and C_i indicate the output and costs of producer *i* respectively. This specification implies that a producer will only perform above the cost frontier if it performs below the production frontier (Murillo-Zamorano, 2004). This approach models technical efficiency only and does not take allocative efficiency into account. The value $\frac{u}{r}$ can be interpreted as a measure of "the extra cost of producing under the production frontier" (Schmidt and Lovell, 1979, pg. 348). In the public sector, outputs can be thought of as exogenous as the output is produced to meet demand. The cost frontier is constructed assuming that the output Y_i is an exogenous variable. Assuming that the cost function is linearly homogeneous in input prices equation 3.27 can be reformulated as (Schmidt and Lovell, 1979):

$$\ln\left(\frac{C_i}{p_k}\right)=K+\frac{1}{r}\ln Y_i+\sum_{j=1}^{k-1}\frac{\beta_j}{r}\ln\left(\frac{p_j}{p_k}\right)-\frac{1}{r}(v_i-u_i) \tag{2.28}$$

Equation 2.28 may be used to construct estimates for the values of σ_u and σ_v. The maximum likelihood estimators are determined using the following log-likelihood function

$$\ln L(K,\boldsymbol{\beta},\sigma,\lambda)=M-n\ln\sigma+\sum_{i=1}^{n}\left[\ln\Phi\left(\frac{\varepsilon_i\lambda}{\sigma}\right)-\frac{1}{2\sigma^2}\left(\varepsilon_i^2\right)\right] \tag{2.29}$$

where *M* is a constant (Schmidt and Lovell, 1979). Other formulations of the cost function have also been used. For example, Greene (2008) used the following function to model costs in U.S. electric power generating firms:

$$\ln\left(\frac{C_i}{p_k}\right)=\alpha_1+\alpha_2\ln Y_i+\alpha_3\left(\frac{1}{2}\ln^2 Y_i\right)+\sum_{j=1}^{k-1}\beta_j\ln\left(\frac{p_j}{p_k}\right)+\varepsilon_i \tag{2.30}$$

A very popular specification of the cost function is the translog formulation (given for production frontier in equation 2.10). A translog cost function representation is given in Christensen and Greene (1976) as

$$\ln C = \alpha_0 + \alpha_Y \ln Y + \gamma_{YY}\left(\frac{1}{2}\ln^2 Y\right) + \sum_i \alpha_i \ln p_i$$
$$+\frac{1}{2}\sum_i\sum_j \gamma_{ij} \ln p_i \ln p_j + \sum_i \gamma_{Yi} \ln Y \ln p_i \tag{2.31}$$

where $\gamma_{ij} = \gamma_{ji}$. A cost function must be homogeneous of degree one for it to correspond to a well-behaved production function (Christensen and Greene, 1976). In order to satisfy this requirement, the following constraints are imposed on the parameters

$$\sum_i \alpha_i = 1, \text{and} \tag{2.32}$$

$$\sum_i \gamma_{Yi} = \sum_i \gamma_{ij} = \sum_j \gamma_{ij} = \sum_i\sum_j \gamma_{ij} = 0. \tag{2.33}$$

2.5 Calculation of Efficiency Measures/Indices

There are two formulations of the parametric frontier; a deterministic and a stochastic method. The nature of each of these formulations was discussed in the previous section. What remains is to use these approaches to calculate specific estimates of technical efficiency for a set of producers given their corresponding input and output values. This section provides the efficiency estimates resulting from each of the frontier approaches discussed above.

For the deterministic frontiers, the efficiency measure results from a transformation of the estimate of u_i given in equation 2.22. The transformation used is given in equation 2.13 and provides an estimate of the Farrell measure of the technical efficiency of producer *i* (Ruggiero, 1999)

$$\hat{\xi}_i = e^{-\hat{u}_i}. \tag{2.34}$$

These estimates of efficiency can now be used to rank and assess the performance of the producers. The producers are ranked according to their measure of efficiency according to the value calculated in equation 2.34 above with higher values of $\hat{\xi}_i$ indicating a higher efficiency ranking.

The method of determining efficiency measures when using the stochastic framework is more complex than its deterministic counterpart. Assuming a normal-half-normal distribution of the composite error term, Jondrow *et al.* (1982) showed that the conditional distribution of the inefficiency term *u* given the composite error term *ε* is that of a normal distribution, $N\left(\frac{-\sigma_u^2\varepsilon}{\sigma_u^2+\sigma_v^2}, \frac{\sigma_u^2\sigma_v^2}{\sigma_u^2+\sigma_v^2}\right)$, truncated at zero. The Jondrow *et al.* (1982) paper was a ground-breaking at the time it was written as it provided a means to estimate the level of technical inefficiency for each producer in the sample, which was not possible prior to this paper. This measure of technical inefficiency for a producer *i* is estimated as (Jondrow *et al.*, 1982)

$$E[u_i \mid \varepsilon_i] = \frac{\sigma\lambda}{1+\lambda^2}\left[\frac{\phi\left(\frac{\varepsilon_i\lambda}{\sigma}\right)}{1-\Phi\left(\frac{\varepsilon_i\lambda}{\sigma}\right)} - \left(\frac{\varepsilon_i\lambda}{\sigma}\right)\right] \tag{2.35}$$

where $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \frac{\sigma_u}{\sigma_v}$, $\varepsilon_i = v_i - u_i$ (the composed error term) and $\phi(\cdot)$ and $\Phi(\cdot)$ represent the density and distribution functions for the standard normal distribution respectively. Equation 2.35 defines the point estimator for the level of technical inefficiency (u_i) for producer *i* given the composite error term ε_i. Using this as the measure of technical inefficiency implies that the measure of technical *efficiency* is $1 - E[u_i \mid \varepsilon_i]$, which is not suggested for the when the logarithm of production is used in the formulation of the model (Battese and Coelli, 1988). In this case it is preferable to use e^{-u_i} as an estimate of the technical efficiency of a producer. To correct for this, Battese and Coelli (1988) proposed the following measure of technical efficiency for the normal-half-normal model (as given in Greene (2008))

$$E\left[e^{-u_i} \mid \varepsilon_i\right] = \frac{\Phi\left(\frac{u_i^*}{\sigma^*} - \sigma^*\right)}{\Phi\left(\frac{u_i^*}{\sigma^*}\right)} e^{\left\{-u_i^* + \frac{1}{2}(\sigma^*)^2\right\}} \tag{2.36}$$

where $\sigma^* = \frac{\sigma_u\sigma_v}{\sigma}$ and $u_i^* = \frac{-\sigma_u^2\varepsilon_i\lambda}{\sigma^2}$. The $E\left[e^{-u_i} \mid \varepsilon_i\right]$ measure of technical efficiency is recommended for values of u_i which are not close to zero (Murillo-Zamorano, 2004).

In the stochastic model, producers can be ranked according to the value of their *in*efficiency estimate, $E[u_i \mid \varepsilon_i]$ (Jondrow *et al.*, 1982), or efficiency estimates, $E[e^{-u_i} \mid \varepsilon_i]$ (Battese and Coelli, 1988).

With respect to the cost frontier, efficiency estimates are found in a similar manner to the production frontier. Cost efficiencies are determined simply be changing the sign of the ε_i value (Greene, 2008).

The box below summarizes the process followed in implementing the SFA and COLS techniques for a given municipality.

Implementation of the SFA and COLS techniques (for a given municipality 'X')

1. The data are collected for all variables and all municipalities (including X) which are taking part in the efficiency evaluation.
2. A functional form for the frontier is selected (usually Cobb-Douglas or translog) and the parameters which require estimation are determined.
3. The parameters of the model are estimated:

   SFA: by fitting the function using a maximum likelihood technique where the residuals (errors) are considered to be made up of an error (random) component, v_i, and an efficiency component, u_i. The procedure is outlined in section 2.4.2.

   COLS: by fitting the function selected in (2) using ordinary least squares (OLS) and shifting the fitted model by changing the intercept value as outlined in section 2.4.1 and given in equation 2.21.

4. The model function defined in (2) with the parameters fitted in (3) now forms the frontier for the efficiency evaluation.
5. According to the position of the municipality X in relation to the frontier fitted to the data as in (4) above the municipality is allocated an efficiency score:

SFA: the efficiency of municipality X is calculated given the composed error (residual) using equation 2.35 or 2.36. Note in this cases the normal-half-normal model is being utilized for the composition of the error term.

COLS: the efficiency of municipality X is calculated using the error (residual) of the fitted OLS model. The efficiency is calculated using the adjusted residual given in equation 2.22. If the SFA model is unable to determine inefficiency in the residuals, the COLS approach will be used to determine efficiencies.

2.6 Critique of Indices

The parametric frontier approach has come under some scrutiny, especially by advocates of data envelopment analysis (DEA). The main criticism of the approach is the requirement of an *a priori* specification of the functional form of the production function. In the parametric case a production function (the Cobb-Douglas or translog) is imposed on the data, and any misspecification of this model can dramatically affect the results of an efficiency analysis (Ruggiero, 1999). There are also potential problems of bias in the stochastic frontier regression approach (Cooper and Tone, 1997).

Criticisms of the deterministic approach are evident from its formulation. Ruggiero (1999) argues that the distributions for the error components (v_i and u_i) under the stochastic frontier approach are chosen purely for mathematical convenience, and not on the observed data values. Ondrich and Ruggiero (2001) expand on this argument to show that, while advocates of the stochastic frontier approach believe that the measure given in equation 2.35 is an absolute measure of technical inefficiency, this can be true only if the users of this approach have full confidence in their chosen distributions for the error components. They show that, if distributions for these error terms are empirically indistinguishable from the original distributions are used, inefficiency measures can change quite substantially.

A particularly alarming demonstration of potential error is presented in Ruggiero (1999). The author simulates a stochastic production scenario with a normal-half-normal composition of the error components. Using the simulated values and the estimation approach discussed above, it is shown that the correctly specified stochastic frontier model does not perform as well in identifying the true efficiencies of the simulated producers as the conceptually flawed deterministic. Ruggiero (1999, p. 558) suggests that, although the conceptual problems exist in the formulation of the deterministic model, the model should not be abandoned unless the alternative stochastic methods are indeed able to overcome the problems introduced by the deterministic formulation (which they may well not be able to).

2.7 Variable Selection

There have been a number of investigations into the efficiency of water service providers using a parametric frontier approach. The predominant frontier used is the cost frontier (water service providers wish to minimise costs), although one could just as easily use a production frontier with the output being the volume of water delivered. Nevertheless, the literature covered in this area focuses specifically on a cost minimisation approach (stochastic cost frontiers).

Cubbin and Tzanidakis (1998) compared the results of a COLS based parametric approach to a non-parametric (DEA) approach for ranking the efficiencies of water service providers in England and Wales. The approach used was the basic regression approach of dropping independent (explanatory) variables until the most parsimonious model is found. The dependent variable used was the natural logarithm of operating expenditure. The dependent and independent variables used were:

Dependent	log(Operating expenditure)
Independent	log(Volume of water delivered); log(length of mains); proportion of distribution to non- households

This approach stands in contrast to the methodology of cost frontiers discussed in section 2.2.2 as there is no use of input prices in the model. This study was conducted using data from the England and Wales water industry.

In a study conducted in the United Kingdom, Ashton (2000) used a translog cost function to determine cost efficiencies in the water and sewerage industry. The level of output used in the study was proxied by the number of households connected to the water system. The input prices (inputs) used were the price of labour, the price of consumables and the price of other costs (service charges, etc.), and the cost variable was defined as the operating costs. Ashton (2000) found that there was a large degree of dispersion in the efficiency results obtained.

Bottasso and Conti (2003) also conducted an assessment of the cost efficiency of water service providers in the UK. The cost variable used in this study was the logarithm of the operating expenditure, and the price variables were for “labour” and “other inputs”. Other variables used in the study were the volume of water distributed and the amount of capital stock. Technical variables for the length of the mains, the average pumping head and the proportion of river sources, as well as a dummy variable indicating the supply of sewerage services were also included in the translog model used. All the variables were also recorded and modelled over the 1995-2001 time period. Bottasso and Conti (2003) estimated that the average cost inefficiency for water service providers steadily decreased over the time period observed.

In a study of cost efficiencies in the Italian water sector Fraquelli and Moiso (2005) also used a translog cost function. The cost variable used was the total costs incurred, and the output variable was the volume of water delivered. The price variables included in the study were the price of labour, the price of electricity and the price of materials, service and capital. Additional variables for the length of the network and the level of water losses were also included. The authors also included a time trend component in order to account for technological advancement.

Filippini *et al.* (2008) estimated cost efficiencies in the Slovenian water sector. In this study Filippini *et al.* (2008) used a translog cost model specification. The cost variable was the total costs incurred, and the output variable was the volume of water delivered. The price variables used in this 2008 study were the price of labour, the price of electricity and the price capital. Extra technical variables for the number of customers served and the size of the service area were also included. Dummy variables indicating the level of water loss, water treatment requirements, the use of surface water and the use of underground water were added to the model. A time variable was also included to capture the progress in technology. In this model all prices and the total cost variable were divided by the price of capital variable.

2.8 Conclusion

Analysis of cost efficiencies dominates the literature relating to water service provision. Operating expenditure and the total costs incurred are popular choices for the cost variable, while the total volume of water delivered is a popular choice for the output variable. The inputs popularly selected are “labour” and “other consumables”. The price of each is required in order to estimate the cost. Additional variables for length of water network add sophistication to the models if information is available on them for all the producers. The convenience of using a parametric approach is that, once a model is fitted to the data, using OLE or MLE, the contribution of each variable to the model is quantified, and those variables which do not contribute can be dropped in favour of a more parsimonious model.

The analysis of the data can be performed in the open-source statistical software *R* (R Core Team, 2013) using the "benchmarking" package (Bogetoft and Otto, 2013) as well as the "frontier" package Coelli and Henningsen (2012). These packages perform analysis using the normal-half-normal composition of the error term as discussed above.

The main criticism of the parametric approach is the requirement of an *a priori* specification of the functional form of the production function. Typically, a production function (the Cobb-Douglas or translog) is imposed on the data, and any misspecification of this model can dramatically affect the results of an efficiency analysis.

CHAPTER 3

THEORY AND METHODOLOGY OF DATA ENVELOPMENT ANALYSIS FRONTIER ESTIMATION FOR THE PURPOSE OF GENERATING MEASURES/INDICES OF EFFICIENCY IN WATER SERVICE PROVISION

3.1 Introduction

The parametric approach discussed in Chapter 2 requires the knowledge of the functional form of production, prior to the analysis. This drawback can be overcome through the use of so called "non-parametric" methods. The most prominent form of non-parametric efficiency analysis is based on linear programming and is called data envelopment analysis (DEA) (Greene, 2008).

The foundation and approach to what was eventually termed DEA was introduced by Farrell (1957). The basic concept of the non-parametric approach is to use the data itself to construct the production frontier, rather than imposing a specific functional form to the frontier. In so doing, the method of DEA constructs a "piecewise-linear, quasi-convex hull around the data points" (Greene, 2008, p. 112). This hull represents the production frontier against which measures of relative efficiency can be determined. This procedure is depicted for two inputs (X_1 and X_2) and a single output in the figure 3.1.

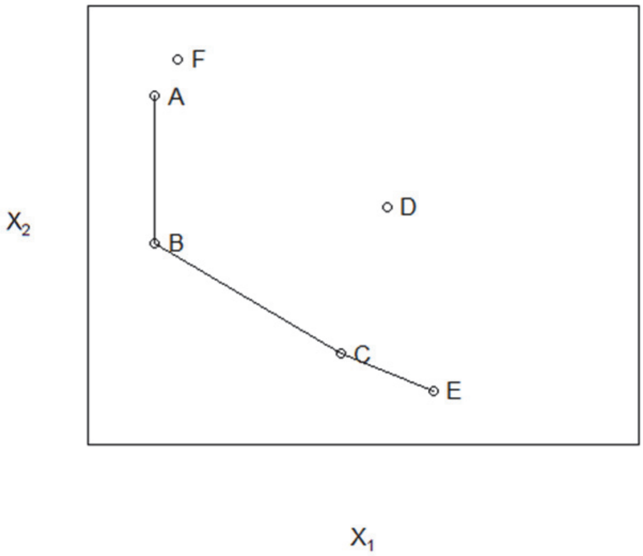

Figure 3.1: Graphical representation of the DEA technique

The DEA "hull" is represented by the line segments through *ABCE*, indicating these as the efficient producers. The producers *D* and *F* in the figure are considered to be inefficient. Chapter 3 will outline and discuss the theory of DEA.

3.2 Theory of the Model

The great benefit of using DEA is its ability to handle situations with both multiple inputs and multiple outputs. Suppose that a group of *n* producers each use the same *l* inputs, X_i where $i = 1,2,\ldots,l$, to produce the same *m* outputs Y_j where $j = 1,2,\ldots,m$. For the sake of consistency with the related literature, producers are now referred to as decision making units (DMUs).

DEA can be thought of as a combination of fractional and linear programming techniques (Ganley and Cubbin, 1992). The fractional program represents the conceptual idea of DEA, while the linear program is used for the efficiency calculations. The conceptual framework (fractional program) is discussed first, and this leads into the discussion of the related linear program.

An intuitive and commonly used measure of efficiency is the ratio of output to input, that is Cooper *et al.* (2006):

$$\frac{\text{Output}}{\text{Input}}, \tag{3.1}$$

The Ouput:Input ratio gives an indication of the amount of output that is generated per unit of input. This forms the basis of the Charnes-Cooper-Rhodes (CCR) model, proposed by Charnes *et al.* (1978) and is discussed in detail in what follows.

Suppose one were to combine all inputs into a single 'virtual' input as

$$\sum_{t=1}^{l} a_t x_t = a_1 x_1 + a_2 x_2 + \cdots + a_l x_l \tag{3.2}$$

where a_t represent weights (as yet not determined) given to each of the inputs, and all outputs are combined into a single 'virtual' output

$$\sum_{s=1}^{m} b_s y_s = b_1 y_1 + b_2 y_2 + \cdots + b_m y_m \tag{3.3}$$

where b_s represent weights (as yet not determined) given to each of the outputs (Cooper *et al.*, 2006). The elementary efficiency ratio given in equation 3.1 can thus be represented as

$$\frac{\sum_{s=1}^{m} b_s y_s}{\sum_{t=1}^{l} a_t x_t}. \tag{3.4}$$

The method of DEA then maximises the efficiency ratio given in equation 3.4 for each DMU k (k = 1,2,…,n) according to the following fraction programming model(FP) (Cooper *et al.*, 2006):

$$FP: \qquad \max_{b_s, a_t} \left(\frac{\sum_{s=1}^{m} b_s y_{sk}}{\sum_{t=1}^{l} a_t x_{tk}} \right) \tag{3.5}$$

subject to $$0 \le \frac{\sum_{s=1}^{m} b_s y_{sk}}{\sum_{t=1}^{l} a_t x_{tk}} \le 1 \left(i = 1,2,\ldots,n\right) \tag{3.6}$$

$$b_s, a_t \ge 0 \text{ for all } s, t \tag{3.7}$$

where y_{sk} and x_{tk} represent output *s* and input *t* of DMU *k* respectively. The constraints ensure that the maximum value (efficiency score) for DMU *k* is in line with a frontier interpretation of performance (Ganley and Cubbin, 1992). A score of unity will imply that a DMU represents "best practice", while values of less than one indicate that the DMU is relatively inefficient. The efficiency scores are based on observed performances of the DMUs and not on an imposed production function.

A problem with the above formulation of the fractional program FP is that there could be an infinite number of solutions (Murillo-Zamorano, 2004). Consider the case where $\hat{a}_t$ and $\hat{b}_s$ represent the solutions to FP. It then follows that $\kappa\hat{a}_t$ and $\kappa\hat{b}_s$ are also solutions for any non-zero value of κ. This drawback can be overcome by converting the above fractional program into a linear program (*LP*). This can be completed in one of two ways; an output maximisation formulation $\left(LP_k^1\right)$ and an input minimisation formulation $\left(LP_k^2\right)$, each of which will result in the maximisation of equation 3.5. This also allows for a researcher to identify the requirements of their analysis and select the more appropriate method. The first linear program is formulated by adding a unity constraint on the 'virtual' input (equation 3.10). This results in the following linear program (Ganley and Cubbin, 1992)

LP_k^1 : $$\max_{b_s, a_t} \gamma_1 = \sum_{s=1}^{m} b_s y_{sk} \tag{3.8}$$

subject to $$\frac{\sum_{s=1}^{m} b_s y_{sk}}{\sum_{t=1}^{l} a_t x_{tk}} \le 1 \left(i = 1,2,\ldots,n\right) \tag{3.9}$$

$$\sum_{t=1}^{l} a_t x_{tk} = 1 \tag{3.10}$$

$$b_s, a_t \geq 0 \text{ for all } s, t \tag{3.11}$$

for each DMU k (k = 1,2,...,n). In the second case of the linear program formulation the 'virtual' output (equation 3.14) is constrained to unity. This formulation is given as (Ganley and Cubbin, 1992)

$$LP_k^2: \qquad \min_{b_s, a_t} \gamma_2 = \sum_{t=1}^{l} a_t x_{tk} \tag{3.12}$$

subject to

$$\frac{\sum_{s=1}^{m} b_s y_{sk}}{\sum_{t=1}^{l} a_t x_{tk}} \leq 1 \left(i = 1,2,\ldots,n\right) \tag{3.13}$$

$$\sum_{s=1}^{m} b_s y_{sk} = 1 \tag{3.14}$$

$$b_s, a_t \geq 0 \text{ for all } s, t \tag{3.15}$$

for each DMU k (k = 1,2,...,n). In order to calculate relative efficiency measures for each DMU, dual formulations (*DP*) to the above linear programmes are required. Cooper *et al.* (2006) represent the dual for LP_k^1 as

$$DP_k^2: \qquad \min_{\theta, \lambda_i} \theta \tag{3.16}$$

subject to

$$\theta x_{tk} - \sum_{i=1}^{n} \lambda_i x_{ti} \geq 0 \left(t = 1,2,\ldots,l\right) \tag{3.17}$$

$$y_{sk} \leq \sum_{i=1}^{n} \lambda_i y_{si} \left(s = 1,2,\ldots,m\right) \tag{3.18}$$

$$\lambda_i \geq 0 \text{ for all } i \tag{3.19}$$

for each DMU k (k = 1,2,...,n). The optimal value for $\theta = \theta^*$ can be thought of as the measure of Farrell efficiency for DMU k, i.e. $1 - \theta^*$ represents the maximum radial reduction of all the inputs which will still ensure at least the given outputs are producible for DMU k.

This represents an "input minimization" formulation. Owing to the inequalities given in equations 3.17 and 3.18 one can define input excesses for DMU k as (Cooper *et al.*, 2006)

$$s_t^- = \theta x_{tk} - \sum_{i=1}^{n} \lambda_i x_{ti} \ (t = 1,2,\ldots,l) \tag{3.20}$$

and output shortfalls as

$$s_s^+ = \sum_{i=1}^{n} \lambda_i y_{si} - y_{sk} \ (s = 1,2,\ldots,m) \tag{3.21}$$

These variables are known as "slack" variables, and represent excess use of inputs and shortfall values of output. The concept of slack is illustrated in figure 3.1. In the figure, DMU *A* and DMU *B* both produce the same quantity of output, but DMU *B* does so using less of input X_2 (while using the same amount of input X_1). This indicates that DMU *A* uses an excess amount of input X_1 (when compared to DMU *B*) and thus there exists a slack in that variable.

Solving the CCR model for a DMU k is now given as a two phase linear programming problem (Cooper *et al.*, 2006). The first phase consists of determining the optimal value of $\theta = \theta^*$ using the linear program given in DP_k^1 above. The second phase is to maximise the sum of the slack variables using the optimal value, θ^*. It is achieved through the following linear program, representing the second phase of the procedure (Cooper *et al.*, 2006)

SP_k :

$$\max_{\lambda_i, s_t^-, s_s^+} \left(\sum_{t=1}^{l} s_t^- + \sum_{s=1}^{m} s_s^+ \right) \tag{3.22}$$

subject to

$$s_t^- = \theta x_{tk} - \sum_{i=1}^{n} \lambda_i x_{ti} \ (t = 1,2,\ldots,l) \tag{3.23}$$

$$s_s^+ = \sum_{i=1}^{n} \lambda_i y_{si} - y_{sk} \ (s = 1,2,\ldots,m) \tag{3.24}$$

$$\lambda_i, s_t^-, s_s^+ \geq 0 \text{ for all } i, s, t \tag{3.25}$$

These two phases are unified into a single model and the complete formulation for the CCR model for input minimisation as (Ganley and Cubbin, 1992)

$$CCR_k^{In}: \qquad \min_{\lambda_i} \theta - \varepsilon\left(\sum_{t=1}^{l} s_t^- + \sum_{s=1}^{m} s_s^+\right) \tag{3.26}$$

subject to

$$\theta x_{tk} - s_t^- = \sum_{i=1}^{n} \lambda_i x_{ti} \; (t = 1,2,\ldots,l) \tag{3.27}$$

$$y_{sk} + s_s^+ = \sum_{i=1}^{n} \lambda_i y_{si} \; (s = 1,2,\ldots,m) \tag{3.28}$$

$$\lambda_i, s_t^-, s_s^+ \geq 0 \text{ for all } i,s,t \tag{3.29}$$

where ε is a "non-Archimedean infinitesimal", i.e. is smaller than any positive number (Cooper *et al.*, 2006).

The CCR model for output maximisation is similarly formulated as (Ganley and Cubbin, 1992)

$$CCR_k^{Out}: \qquad \max_{\lambda_i} \psi + \varepsilon\left(\sum_{t=1}^{l} s_t^- + \sum_{s=1}^{m} s_s^+\right) \tag{3.30}$$

subject to

$$x_{tk} - s_t^- = \sum_{i=1}^{n} \lambda_i x_{ti} \; (t = 1,2,\ldots,l) \tag{3.31}$$

$$\psi y_{sk} + s_s^+ = \sum_{i=1}^{n} \lambda_i y_{si} \; (s = 1,2,\ldots,m) \tag{3.32}$$

$$\lambda_i, s_t^-, s_s^+ \geq 0 \text{ for all } i,s,t \tag{3.33}$$

A DMU is only considered to be CCR efficient if $\theta^* = 1$ (or $\psi^* = 1$) *and* all slack variables equal zero, i.e. $s_t^-, s_s^+ = 0$ for all *s, t* (Cooper *et al.*, 2006).

For illustrative purposes our attention is turned back to the scenario depicted in figure 3.1 duplicated below as figure 3.2 with a ray drawn from the origin *O* to the inefficient DMU *D*.

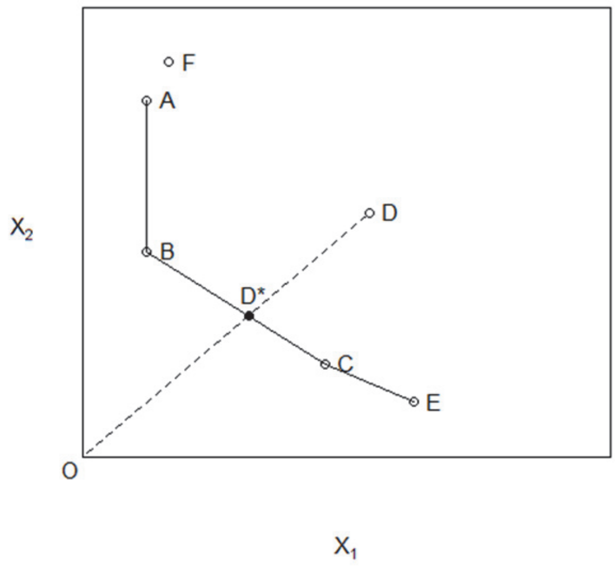

Figure 3.2: Illustration of DEA for two inputs

The DEA measure of efficiency for DMU *D* using the input minimisation formulation of the CCR model, can be calculated from the figure as $\theta_D^* = \frac{\|OD^*\|}{\|OD\|}$. If DMU *D* reduced its use of input X_1 and input X_2 by the values $\theta_D^* X_1$ and $\theta_D^* X_2$ respectively, while maintaining the same output, DMU would be regarded as efficient. From the figure it can be seen that the efficient position for DMU *D* is at point D^*, which falls on the line segment between DMU *B* and DMU *C*. The point D^* is a linear combination of the inputs used by the efficient DMUs *B* and *C*. This will be indicated in the DEA solution, as the coefficients λ_i for these efficient DMUs will be non-zero. For the scenario depicted in figure 3.2, the DEA solution for DMU *D* will have $\lambda_B, \lambda_C \geq 0$ and $\lambda_j = 0$ for $j \neq B, C$. This is a useful advantage for the use of DEA in benchmarking, as it indicates the efficient DMUs (in this case *B* and *C*), known as a "peer group" from which the inefficient DMU *D* can learn.

The DEA methodology considered up to this point is based on the assumption of constant returns to scale (CRS). The nature of CRS is depicted in figure 3.3.

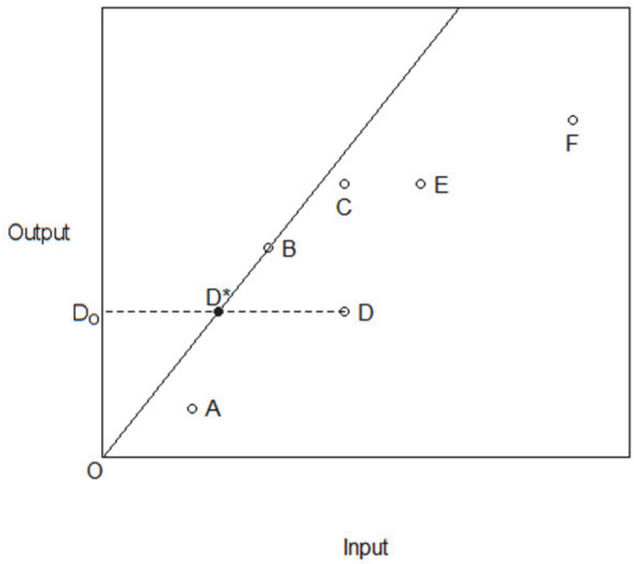

Figure 3.3: Constant returns to scale (CRS)

The most (and only) efficient DMU in figure 3.3 is DMU *B* which lies on the constant returns frontier, line *OB* (in the figure). Each DMU will thus only achieve efficient production if they fall on the line through *OB*. The relative efficiency of each DMU (θ^*) (when considering input minimisation) is measured as a ratio of the optimal input (for the given output) and the observed input. In the case of DMU *D* in the figure, this results in $\theta^* = \frac{\|D_0 D^*\|}{\|D_0 D\|}$.

The assumption of CRS is that all participating DMUs are performing at an optimal scale (Murillo-Zamorano, 2004). Owing to many circumstances, however, this assumption is not realistic and can yield misleading results. Specifically, scale inefficiency in the CRS technology contributes to the technical inefficiency measurement (Murillo-Zamorano, 2004).

The drawback of the above approach was addressed by Banker *et al.* (1984) where a varying returns to scale (VRS) technology was used. The concept of VRS is depicted in figure 3.4.

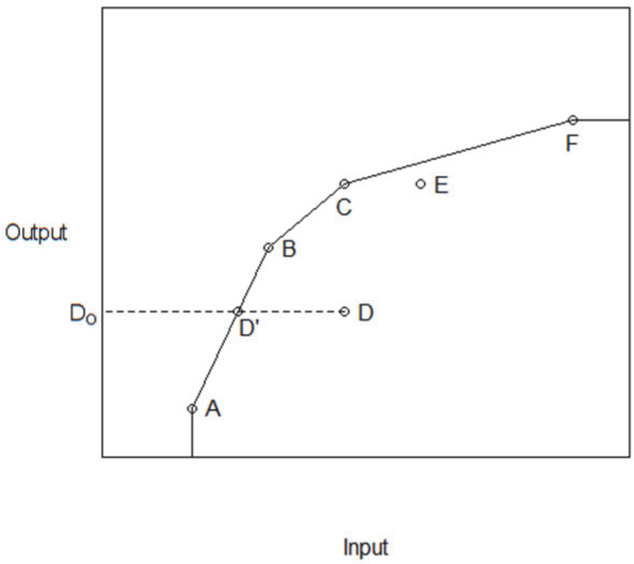

Figure 3.4: Varying returns to scale (VRS)

The VRS frontier in figure 3.4 allow the returns to scale to vary from line segment to line segment (Ganley and Cubbin, 1992). The efficient DMUs are now *A*, *B*, *C* and *F*. The addition of VRS to the model is achieved by adding a unity constraint to the sum of the λ_i values. This DEA model is called the Banker-Charnes-Cooper (BCC) model (after Banker *et al.* (1984)) and is formulated for input minimisation as (Banker *et al.*, 1984)

$$BCC_k^{In}: \qquad \min_{\lambda_i} \vartheta - \varepsilon\left(\sum_{t=1}^{l} s_t^- + \sum_{s=1}^{m} s_s^+\right) \tag{3.34}$$

subject to

$$\vartheta x_{tk} - s_t^- = \sum_{i=1}^{n} \lambda_i x_{ti} \; (t = 1,2,\ldots,l) \tag{3.35}$$

$$y_{sk} + s_s^+ = \sum_{i=1}^{n} \lambda_i y_{si} \; (s = 1,2,\ldots,m) \tag{3.36}$$

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{3.37}$$

$$\lambda_i, s_t^-, s_s^+ \geq 0 \text{ for all } i, s, t \tag{3.38}$$

and for output maximisation as

$$BCC_k^{Out}: \qquad \max_{\lambda_i} \varphi - \varepsilon\left(\sum_{t=1}^{l} s_t^- + \sum_{s=1}^{m} s_s^+\right) \tag{3.39}$$

subject to

$$x_{tk} - s_t^- = \sum_{i=1}^{n} \lambda_i x_{ti} \; (t = 1,2,\ldots,l) \tag{3.40}$$

$$\varphi y_{sk} + s_s^+ = \sum_{i=1}^{n} \lambda_i y_{si} \; (s = 1,2,\ldots,m) \tag{3.41}$$

$$\sum_{i=1}^{n} \lambda_i = 1 \tag{3.42}$$

$$\lambda_i, s_t^-, s_s^+ \geq 0 \text{ for all } i, s, t \tag{3.43}$$

The calculated efficiency measure for DMU D in figure 3.4 using the above BCC DEA model results in an efficiency estimate of $\vartheta^* = \frac{\|D_0 D'\|}{\|D_0 D\|}$.

It is important to consider the measure of scale inefficiency which causes the discrepancy in efficiency measures between the CCR (CRS) and BCC (VRS) models. (Murillo-Zamorano, 2004) explain that "total technical efficiency" for a DMU k denoted T_k^* is expressed as the product of "pure technical efficiency" (E_k^*) and "scale efficiency" (S_k^*), i.e.

$$T_k^* = E_k^* \times S_k^* \tag{3.44}$$

The value for T_k^* is found as the solution to the CCR (CRS) model for DMU k, which is effected by scale inefficiencies. The solution to the BCC (VRS) model provides the value for E_k^*, the "pure" technical efficiency which is unaffected by the scale efficiency. The scale efficiency for DMU k is thus calculated as

$$S_k^* = \frac{T_k^*}{E_k^*} = \frac{\theta_k^*}{\vartheta_k^*} \tag{3.45}$$

Both the CRS and VRS models are illustrated in figure 3.5

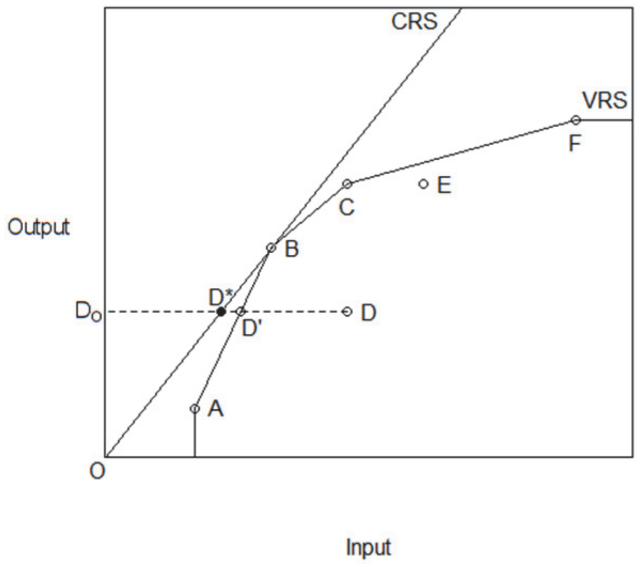

Figure 3.5: The relationship between technical and scale efficiency

The CCR the BCC efficiency for a DMU D are given as $\theta^* = \frac{\|D_0D^*\|}{\|D_0D\|}$ and $\vartheta^* = \frac{\|D_0D'\|}{\|D_0D\|}$ respectively. Using equation 91, the scale efficiency for DMU D is given as $S_k^* = \frac{\|D_0D^*\|}{\|D_0D'\|}$. This illustration was conducted under the input minimisation orientation for the DEA models. The results follow similarly for the output maximisation orientation.

3.3 Calculation of DEA Efficiency Measures/Indices

The method of DEA allows for an intuitive ranking of the DMUs. The rankings according to the DEA methodology are completed by using the value of θ_k^* (or ϑ_k^*) calculated for each DMU k. DMUs with the highest θ_k^* (or ϑ_k^*) values will achieve the highest ranking, and those with the lowest values will achieve the worst ranking. This ranking is indicated for the input minimisation orientation of the DEA models, but follows identically for the output maximisation orientation. It is common for two or more DMUs to achieve a fully efficient ranking (particularly when VRS is used). In this case the effected DMUs will simply assume the same ranking. The box below summarizes the implementation of the DEA technique for a given municipality.

Implementation of the DEA technique for a given municipality 'X'

1. The data are collected for all variables and all municipalities (including X) which are taking part in the efficiency evaluation.
2. A convex hull – which forms the frontier – is then fitted around all the data points using the techniques described in Section 3.2 (either using a CCR (CRS) or BCC (VRS) model) and illustrated in Figure 3.1. (This is typically performed using a computer program)
3. According to the position of the municipality X in relation to the frontier fitted to the data as in (2) above the municipality is allocated an efficiency score.

   a) if municipality X is positioned on the frontier (as illustrated in Figure 3.2 as points A,B,C and E) the municipality A will achieve an efficiency score of 1 (100%)

   b) if municipality X is positioned within the boundaries of the frontier (as illustrated in Figured 3.2 as points D and F) the municipality will achieve and efficiency rating in relative to its distance from the fitted frontier. Higher efficiencies (close to 1) will be achieved municipality X if its position is close to the frontier and lower values (close to 0) with be achieved as it moves further from the frontier.

3.4 Critique of Indices

The main critique of the results of a DEA analysis is that non-parametric frontier techniques are deterministic by nature. As such they share the drawbacks of the deterministic frontier techniques discussed in section 3.4, most notably the fact that all deviations from the determined frontier are attributed to the inefficiency of the DMU. Statistical noise and any measurement error is thus not considered as part of the model (Greene, 2008). This has resulted in some literature on the subject describing the methods as non-statistical (Murillo-Zamorano, 2004).

There have, however, been some advances in the field to address these shortcomings. Most notably, perhaps, is the use of bootstrapping of DEA estimates (Greene, 2008) to construct confidence intervals for DEA measures of efficiency. The drivers behind this approach are Simar and Wilson who published a number of papers on this topic at the turn of the twenty-first century (see Simar and Wilson (1998) and Simar and Wilson (2000)). These approaches require significant computing power to run. Murillo-Zamorano (2004) concede that although using the bootstrapping techniques to overcome the statistical inference problem of these non-parametric techniques, other concerning issues remain. The most prominent of these is the sensitivity of non-parametric methods to outliers in the data. But once again, there are ways of managing this problem (Cazals *et al.*, 2002).

3.5 Variable Selection

Various studies have been conducted in the efficiency of water service provision in many countries. Singh *et al.* (2011, p. 88) provide a table which provides a good summary of the studies that used of DEA in the benchmarking and efficiency analysis of the water sector. This report focuses on a number of such studies to develop knowledge of the nature of the variables which have been used and the variables which should be used in DEA efficiency analysis in the water sector.

The first study considered is that of Aida *et al.* (1998) who used (a range adjusted measure) of efficiency using DEA in two regions in Japan. The number of DMUs included in the study exceeded one hundred and the inputs and outputs used are given below:

Inputs	Number of employees; operating expenses (before depreciation); net plant; equipment value, population size; length of pipes,
Outputs	Operating revenues; the amount of water billed net of leakage.

Aida *et al.* (1998) used a (range adjusted) BCC model with an input minimization orientation.

Cubbin and Tzanidakis (1998) used a CCR model with an input minimization orientation in the England and Wales water industry. The inputs and output used were

Inputs	Operating expenditure
Outputs	Volume of water delivered; length of water main

An additional non-controllable variable of "proportion of distribution to non-households" was also included in the study.

An input minimization oriented CCR model was also used in UK water sector by Thanassoulis (2000b), where the inputs and output used were

Inputs	Operating expenditure
Outputs	Volume of water delivered; length of water main; number of connections served

Thanassoulis (2000b) used the correlations between a number of variables to help reduce the number of outputs included in the model. The rationale behind this was that if large correlations are found between two output variables, only one of these should be considered in the DEA model.

In a study of the relative efficiency of water service provider in the Brazilian water and sewage sector, Tupper and Resende (2004) considered the following inputs and outputs in a BCC model using and output maximization orientation.

Inputs	Labour expenses; operational costs (e.g. materials); other operational costs (e.g. interest on debts)
Outputs	Water produced; treated sewage; population served (treated sewage); population served (water)

This study was conducted using data for 5 years (1996-2000) and considered 20 state utilities.

An input minimisation oriented CCR model was used to determine the efficiencies of Spanish municipalities by García-Valiñasa and Muñiza (2007).

Inputs	Operational costs
Outputs	Volume of water delivered; length of water main; population supplied with water

An interesting addition to the DEA model was a non-controllable variable for rainfall. This is advantageous as it gives an indication of the effect of natural circumstances on the efficiency results. The municipalities were assessed over a period of 16 years (1985-2000) and the evolution of each according to their DEA efficiency score was observed.

The last study considered for this report is that conducted in Peru by Berg and Lin (2008). This study used both the CCR and the BCC models with an input minimisation orientation. The inputs and outputs used are given below.

Inputs	Operating costs, number of employees, number of water connections
Outputs	Volume of water billed; number of customers, coverage of service; continuity of service.

The purpose of including the outputs "coverage of service" and "continuity of service" was to bring the study in line with an already implemented government benchmarking system. Berg and Lin (2008) used the volume of water billed as opposed to the volume of water delivered as solution to the serious water loss problem (unaccounted for water) in the area. Lastly, the number of water connections input variable was included as an indication of the capital of the assessed DMUs.

The majority of the DEA methodologies used in the water sector use operating costs as the input variable in an input minimization oriented model. The input orientation used in the majority of the studies is the most advisable method to use for water service providers as these utilities are required to meet a demand in the community (making the output exogenous) (Berg and Lin, 2008). As such, the ability to limit the level of operating expenditure (input), while still meeting the demand for water provision, is what will determine the efficiency level of the provider.

Thanassoulis (2000a) provided a review of the use of DEA in the regulation of water service providers. The author's suggestions for input and output variables were based on factors initially identified by the UK Office of Water Services (OFWAT) and are given in Table 3.1.

Table 3.1: Possible input and output variables for the use in a DEA model for water service providers (Thanassoulis, 2000a, pg. 8)

Potential Inputs	*Potential Outputs*
Operating Expenditure	Number of connections served Length of mains (dispersion) Water delivered to clients (metered and non-metered) Measured amount of water delivered Estimated remainder of water delivered Expenditure incurred for repairs (pipe bursts)

Using this as a basis, the number of outputs considered in the analysis can be reduced by checking the correlations between the output variables (Thanassoulis, 2000a). As mentioned before, if two variables have a high correlation coefficient, only one of the two need be considered in the DEA model. This approach was discussed and used by (Thanassoulis, 2000b) and (Thanassoulis, 2000a).

The DEA measures can be calculated using the open-source statistical software *R* (R Core Team, 2013) using the "benchmarking" package (Bogetoft and Otto, 2013).

3.6 Conclusion

The most prominent form of non-parametric efficiency analysis is one based on linear programming and is called data envelopment analysis (DEA) (Greene, 2008). It overcomes the weakness of the parametric approach (discussed in Chapter 2) which requires knowledge of the functional form of production, prior to the analysis.

Chapter 3 has shown how the DEA method uses the data itself to construct the production frontier, a "piecewise-linear, quasi-convex hull around the data points" (Greene, 2008, p. 112). This hull represents the production frontier against which measures of relative efficiency can be determined. While expenditure on water is typically used as an output in the parametric modelling (SFA), it is typically used as an input in the model and related to outputs such as volume of water served, number of connections and length of pipes making up the distribution network.

CHAPTER 4

SAMPLE DESIGN AND COLLECTION OF DATA

4.1 Sample Frame

South Africa is split into 283 municipal regions, each classified as local (B), district (C) and metropolitan (A) municipalities. The B and C municipalities are further sub-categorised into category B1, B2, B3, B4, C1 and C2 municipalities. The characteristics of each of these sub-categories are given in Table 4.1.

Table 4.1: Categorisation of South African Municipalities

Classification	**Description**	**Characteristics**
A	Metropolitan	Large urban complex, > 1 million population
B1	Local	Large budgets, secondary cities
B2	Local	Large town
B3	Local	Small towns, significant urban population
B4	Local	Mainly rural, one or two small towns
C1	District	Not water service authorities (WSA)
C2	District	Water service authorities (WSA)

(Source: CoGTA (2009))

District municipalities (category C) are comprised of a number of smaller local municipalities (category B). Each district municipality is sub-categorised according to whether it is a water service authority (WSA) or not.

Since district municipalities are made up of a collection of local municipalities it does not make sense to include both district municipalities and local municipalities into a single DEA (or SFA) evaluation, owing to the fact that there could then be an overlap of data, because as the local municipality data is incorporated into the district municipality data. As such efficiency evaluations should focus on comparing municipalities within each categorisation (A, B or C). However, if it is desired to evaluate metropolitan municipalities along with local/district municipalities, these evaluations should combine the municipalities in one (or both) of the following ways:

- metropolitan and local municipalities, or
- metropolitan and district municipalities.

Using the above combinations will avoid any overlap in the regions evaluated both by geographical boundary and by data contents. For this reason it was decided to undertake two DEAs – one incorporating only local municipalities and a second incorporating both municipal and district municipalities.

4.2 Sample selection constraints

It is desired that all the South African municipalities be included in an efficiency evaluation, but this is only possible if all of the required data is available for each municipality considered. Many municipalities are thus excluded from efficiency evaluations on the basis of insufficient data.

Data was available from the municipalities described in Table 4.2.

Table 4.2: Sample Selection by data availability

MUNICIPALITY	Category	WSA	Population	Households
NELSON MANDELA BAY METROPOLITAN MUNICIPALITY	A	Yes	1056350	287595
CACADU DISTRICT MUNICIPALITY	C1	Yes	395931	107214
CAMDEBOO LOCAL MUNICIPALITY	B3	Yes	45912	11213
MAKANA LOCAL MUNICIPALITY	B2	Yes	76841	19281
NDLAMBE LOCAL MUNICIPALITY	B3	Yes	58571	17251
BAVIAANS LOCAL MUNICIPALITY	B3	Yes	15550	4136
KOUGA LOCAL MUNICIPALITY	B3	Yes	73587	21064

MUNICIPALITY	Category	WSA	Population	Households
AMATHOLE DISTRICT MUNICIPALITY	C2	Yes	1617769	424989
MBHASHE LOCAL MUNICIPALITY	B4	No	256885	55780
MNQUMA LOCAL MUNICIPALITY	B4	No	267443	65324
GREAT KEI LOCAL MUNICIPALITY	B3	No	44142	11826
AMAHLATHI LOCAL MUNICIPALITY	B3	No	130782	24055
BUFFALO CITY LOCAL MUNICIPALITY	B1	Yes	694362	198368
NGQUSHWA LOCAL MUNICIPALITY	B4	No	77273	20786
NKONKOBE LOCAL MUNICIPALITY	B3	No	121478	31816
O.R. TAMBO DISTRICT MUNICIPALITY	C2	Yes	1576165	334138
MBIZANA LOCAL MUNICIPALITY	B4	No	229872	44841
NTABANKULU LOCAL MUNICIPALITY	B4	No	128619	26659
NGQUZA HILL LOCAL MUNICIPALITY	B4	No	239593	49485
PORT ST JOHNS LOCAL MUNICIPALITY	B4	No	136487	28077
NYANDENI LOCAL MUNICIPALITY	B4	No	257212	53458
MHLONTLO LOCAL MUNICIPALITY	B4	No	187948	42176
KING SABATA DALINDYEBO LOCAL MUNICIPALITY	B2	No	396434	89442
XHARIEP DISTRICT MUNICIPALITY	C1	No	143155	43308
LETSEMENG LOCAL MUNICIPALITY	B3	Yes	46591	13667
Dr RUTH SEGOMOTSI MOMPATI DISTRICT MUNICIPALITY	C1	No	781471	232570
MANGAUNG LOCAL MUNICIPALITY	B1	Yes	695083	209141
MANTSOPA LOCAL MUNICIPALITY	B3	Yes	56365	14725
LEJWELEPUTSWA DISTRICT MUNICIPALITY	C1	No	671443	197676
TSWELOPELE LOCAL MUNICIPALITY	B3	Yes	53892	13092
MATJHABENG LOCAL MUNICIPALITY	B1	Yes	407800	126175
THABO MOFUTSANYANE DISTRICT MUNICIPALITY	C1	No	767614	203327
SETSOTO LOCAL MUNICIPALITY	B3	Yes	134668	37579
DIHLABENG LOCAL MUNICIPALITY	B2	Yes	142846	38424
FEZILE DABI DISTRICT MUNICIPALITY	C1	No	484565	133227
MOQHAKA LOCAL MUNICIPALITY	B2	Yes	176473	45806
NGWATHE LOCAL MUNICIPALITY	B3	Yes	120546	34205
METSIMAHOLO LOCAL MUNICIPALITY	B2	Yes	124806	36462
MAFUBE LOCAL MUNICIPALITY	B3	Yes	62740	16754
CITY OF JOHANNESBURG METROPOLITAN MUNICIPALITY	A	Yes	3913967	1282624
CITY OF TSHWANE METROPOLITAN	A	Yes	2353332	699686

MUNICIPALITY	Category	WSA	Population	Households
MUNICIPALITY				
EKURHULENI METROPOLITAN MUNICIPALITY	A	Yes	3015386	950844
SEDIBENG DISTRICT MUNICIPALITY	C1	No	934542	277999
EMFULENI LOCAL MUNICIPALITY	B1	Yes	773532	230691
MIDVAAL LOCAL MUNICIPALITY	B2	Yes	77142	24626
LESEDI LOCAL MUNICIPALITY	B3	Yes	83868	22682
METSWEDING DISTRICT MUNICIPALITY	C1	No	219160	62834
NOKENG TSA TAEMANE LOCAL MUNICIPALITY	B2	Yes	66070	18775
KUNGWINI LOCAL MUNICIPALITY	B2	Yes	153089	44058
WEST RAND DISTRICT MUNICIPALITY	C1	Yes	876100	256839
MOGALE CITY LOCAL MUNICIPALITY	B1	Yes	355038	107458
RANDFONTEIN LOCAL MUNICIPALITY	B2	Yes	157288	46317
WESTONARIA LOCAL MUNICIPALITY	B2	Yes	124207	35745
MERAFONG CITY LOCAL MUNICIPALITY	B2	Yes	239567	67318
ETHEKWINI METROPOLITAN MUNICIPALITY	A	Yes	3509256	912458
UGU DISTRICT MUNICIPALITY	C2	Yes	735583	165319
UMDONI LOCAL MUNICIPALITY	B2	No	67779	17457
UMZUMBE LOCAL MUNICIPALITY	B4	No	198911	41221
EZINQOLENI LOCAL MUNICIPALITY	B4	No	55911	11506
HIBISCUS COAST LOCAL MUNICIPALITY	B2	No	232818	57647
UMGUNGUNDHLOVU DISTRICT MUNICIPALITY	C2	Yes	1003785	246429
UTHUKELA DISTRICT MUNICIPALITY	C2	Yes	700721	151179
EMNAMBITHI-LADYSMITH LOCAL MUNICIPALITY	B2	No	247220	58168
INDAKA LOCAL MUNICIPALITY	B4	No	118206	23322
UMTSHEZI LOCAL MUNICIPALITY	B3	No	67183	15435
NTAMBANANA LOCAL MUNICIPALITY	B4	No	143359	29100
IMBABAZANE LOCAL MUNICIPALITY	B4	No	124753	25154
UTHUNGULU DISTRICT MUNICIPALITY	C2	Yes	954382	193538
UMHLATHUZE LOCAL MUNICIPALITY	B1	Yes	345776	82000
CAPRICORN DISTRICT MUNICIPALITY	C2	Yes	1144529	281501
WATERBERG DISTRICT MUNICIPALITY	C1	No	648401	163267
MOGALAKWENA LOCAL MUNICIPALITY	B2	Yes	294655	70482
GERT SIBANDE DISTRICT MUNICIPALITY	C1	No	1006160	247760

MUNICIPALITY	Category	WSA	Population	Households
MSUKALIGWA LOCAL MUNICIPALITY	B2	Yes	142673	35618
GOVAN MBEKI LOCAL MUNICIPALITY	B1	Yes	237957	69537
NKANGALA DISTRICT MUNICIPALITY	C1	No	1137416	288304
STEVE TSHWETE LOCAL MUNICIPALITY	B1	Yes	161199	42927
DR JS MOROKA LOCAL MUNICIPALITY	B4	Yes	270568	62522
EHLANZENI DISTRICT MUNICIPALITY	C1	No	1521148	362740
MBOMBELA LOCAL MUNICIPALITY	B1	Yes	527056	130306
BOJANALA PLATINUM DISTRICT MUNICIPALITY	C1	No	1234265	355032
RUSTENBURG LOCAL MUNICIPALITY	B1	Yes	433701	134239
NGAKA MODIRI MOLEMA DISTRICT MUNICIPALITY	C2	Yes	793580	196655
MAFIKENG LOCAL MUNICIPALITY	B2	No	278351	72823
RAMOTSHERE MOILOA LOCAL MUNICIPALITY	B3	No	139599	34087
Dr KENNETH KAUNDA DISTRICT MUNICIPALITY	C1	No	674040	181206
TLOKWE LOCAL MUNICIPALITY	B1	Yes	143201	37530
NAMAKWA DISTRICT MUNICIPALITY	C1	Yes	119660	31199
RICHTERSVELD LOCAL MUNICIPALITY	B3	Yes	10155	2744
NAMA KHOI LOCAL MUNICIPALITY	B3	Yes	47296	12105
HANTAM LOCAL MUNICIPALITY	B3	Yes	27792	6940
KAROO HOOGLAND LOCAL MUNICIPALITY	B3	Yes	10753	3158
KHAI-MA LOCAL MUNICIPALITY	B3	Yes	12397	3145
PIXLEY KA SEME DISTRICT MUNICIPALITY	C1	Yes	167758	43693
THEMBELIHLE LOCAL MUNICIPALITY	B3	Yes	14863	3815
SIYATHEMBA LOCAL MUNICIPALITY	B3	Yes	18046	4434
FRANCES BAARD DISTRICT MUNICIPALITY	C1	Yes	342883	92485
SOL PLAATJIE LOCAL MUNICIPALITY	B1	Yes	211828	55426
JOHN TAOLO GAETSEWE DISTRICT MUNICIPALITY	C1	Yes	194365	46985
GA-SEGONYANA LOCAL MUNICIPALITY	B3	Yes	72367	18541
WEST COAST DISTRICT MUNICIPALITY	C1	Yes	340379	92785
CEDERBERG LOCAL MUNICIPALITY	B3	Yes	47098	13018
CAPE WINELANDS DISTRICT MUNICIPALITY	C1	Yes	751436	186918
DRAKENSTEIN LOCAL MUNICIPALITY	B1	Yes	232469	55745
EDEN DISTRICT MUNICIPALITY	C1	Yes	533178	147479

MUNICIPALITY	Category	WSA	Population	Households
CENTRAL KAROO DISTRICT MUNICIPALITY	C1	Yes	63831	16614

This reduced sample frame of 106 municipalities, 37 of which were not WSAs, served a total population of 48 810 248, making up 13 127 306 households and serving 11 102 003 water connections. From this list of municipalities the ones that qualified for inclusion in this preliminary investigation were those listed in Table 4.3.

Table 4.3: List of municipalities used in the preliminary DEA water services efficiency evaluation.

Metropolitan	**District**	**Local**	
Nelson Mandela Bay	O.R. Tambo	Camdeboo	Moqhaka
eThekwini	Ugu	Makana	Ngwathe
City Of Johannesburg	Umgungundhlovu	Ndlambe	Metsimaholo
City Of Tshwane	Uthukela	Baviaans	Mafube
Ekurhuleni	Uthungulu	Buffalo City	Emfuleni
	Capricorn	Letsemeng	Midvaal
	West Coast	Mangaung	Lesedi
	Central Karoo	Mantsopa	Nokeng Tsa Taemane
		Tswelopele	Kungwini
		Matjhabeng	Mogale City
		Merafong City	Nama Khoi
		Umhlathuze	Hantam
		Mogalakwena	Karoo Hoogland
		Msukaligwa	Khai-Ma
		Govan Mbeki	Cederberg
		Steve Tshwete	Drakenstein
		Dr JS Moroka	Dihlabeng

Metropolitan	District	Local	
		Mbombela	Westonaria
		Rustenburg	Richtersveld
		Mafikeng	Ramotshere Moiloa
		Setsoto	Randfontein

The municipalities listed in Table 4.3 were judged to satisfy the data requirements for the application of DEA.

4.3 Collection and Description of the data

4.3.1 Introduction

The ability to produce useful results goes hand in hand with the ability to collect valid and complete data. That being said, collecting data which is correct, complete and consistent is a common problem faced by those investigating the South African public services sector. This problem is particularly evident at a municipality level in South Africa. Recent benchmarking (NBI and MBI) and quality control (Blue Drop, Green Drop) endeavours conducted by local government departments have brought to light the importance of consistent and complete data records. These initiatives have helped develop an awareness of the importance of data collection and record keeping within many municipalities. Even with these improvements, gathering valid and complete data, particularly with regard to water services, continues to be a challenge.

Data regarding water usage and provision, as well as data regarding the financial position of South African Municipalities, is required in order to perform comparative efficiency evaluations. The sections that follow describe the data required, as well as the process undertaken to gather the data. Shortcomings of the data and the collection process are also discussed.

4.3.2 Data Requirements

The data required to perform an efficiency analysis must encompass information regarding both the inputs and the outputs of the firms under investigation. When it comes to efficiency analysis within the water sector, a large amount of research has been conducted and the input and output variables used in these studies are similar. Typically, the input variables account

for the resources used by the water service provider, while the output variables provide information on the extent of the infrastructure, as well as the volume of water delivered.
The data required for an efficiency study using data envelopment analysis (DEA) can use any number of inputs as well as outputs. The ability to include more than one input variable distinguishes DEA from regression based models (namely stochastic frontier analysis (SFA)) as these models are limited to a single input variable. The number of variables used in either of these models is constrained by the number of utilities being assessed. The number of utilities being assessed in a DEA evaluation must be at least three times larger than the total number of variables (input and output) used. For SFA studies, the number of utilities investigated is also constrained by the number of variables included in the model (single input, multiple output). In this case, the recommended number of utilities is determined by the guidelines set out for standard regression modeling. A study of related literature (discussed previously) indicates that there are common choices for input and output variables. These common variables are listed in Table 4.4.

Table 4.4: Common variables used in efficiency evaluations of water service providers

Operating expenditure (Cost)
Volume of water delivered
Volume of water billed
Volume of water treated
Size of population served
Number of connections served
Length of mains/network
Price of labour, materials, consumables
Expenditure on repairs
Proportion of water delivered to non-households
Number of Employees

Although the variables listed in Table 4.4 can be used in both DEA and SFA, the price variables are typically used in the SFA approach. Other variables which are common to the SFA approach in the literature are given in Table 4.5.

Table 4.5: Variables included in SFA studies of water service providers

Supply of sewerage services
Amount of capital stock
Proportion of river sources
Use of surface water
Use of underground water

Frequently studies using the stochastic frontier approach also include a variable for time. This variable is used to account for advances in the efficiency and the capabilities of the utilities over a period of time (usually a few years).

A variable which can play an important role is the topography of the area in which the utility is situated. This may affect the costs of pumping and storage of water and also the ability to conduct repairs.

4.3.3 Data

The data required are not available from a single source in South Africa. As such, multiple data sources will be used in order to gather a sufficient amount of data to perform an efficiency analysis. All the required data must come through the municipalities and the record keeping in some of these institutions is not of a high standard. An example of the challenges faced by researchers and authorities attempting to collect data from municipalities is given in the reports of the National Water Services Benchmarking Initiative which ran from 2004-2007. The report of the 2005 implementation of this initiative stated that most of the data supplied by the municipalities was an estimate or guess, with only 1% of the responses coming from audited data (DWAF, SALGA and WRC, 2005).

More recently a study performed by the Water Research Commission (WRC), the data collected from the municipalities was only marginally more reliable. This study by Mckenzie et al. (2012) investigated the state of non-revenue water in South Africa in 2012. This study gathered data for many of the variables used in this study and the reliability of the data collected is given in Table 4.6.

Table 4.6: Reliability of record keeping of South African Municipalities

Reliability of record keeping	**Percent of total**
Good records	19
Gaps in and questions about records	21
Erratic Worthless Records	28
Poor Records	17
No Records	15
Total	100

Source: McKenzie et al. (2012)

The study by McKenzie et al. (2012) provided much insight into the process by which data of this kind is collected and sourced in a South African setting and explained many of the challenges still faced by researchers when gathering data from municipalities.

The following data sources were selected owing to the ease with which the information can be obtained as well as their prominence in related literature/studies:

- StatsSA/National Treasury;
- Previous Studies and
- Blue Drop Reports,

and a single primary source:

- collection of data directly from the municipalities.

4.3.4 StatsSA

Each year a financial (Report P9114) and non-financial (Report P9115) census of the municipalities is conducted and the reports are compiled by StatsSA. These reports are published, and made available to the public on their website in September of each year with a

delay of 365 days. The data for these audits are collected for municipal years which run from 1 July to 30 June. The data currently available on the StatsSA website (www.statssa.gov.za, August 2013) only covers the period 2005-2011.
These reports provide data for many of the variables mentioned above and as such are the main source of data for this study. The data available in these documents, for variables which are of particular importance to this study, are given below.

- P9114 – Financial Census of Municipalities
    - Expenditures on employees, repairs, depreciation, etc.
    - Value of fixed assets (property, etc.)

- P9115 – Non-financial Census of Municipalities
    - Number of employees (Permanent, Temporary and Vacant)
    - Number of units served with water (Inside yard, < 200m and > 200m)
    - Number of units receiving water services
    - Number of units receiving free water services
    - Number of units receiving sewerage and sanitation services
    - Number of units receiving free sewerage and sanitation services

The data for all of the above data (excluding fixed assets) is provided for both water and sanitation services.

It is interesting to note that the 2005 and 2006 non-financial census of municipalities report (P9115) included values of water supply and water loss (in kl) for each of the municipalities. These variables were excluded in all subsequent publications of the non-financial audits.
Additional data on the financial activities of the municipalities can be found on the website for the National Treasury. This data can be used for verification purposes.

The data provided by StatsSA in the P9114 and P9115 reports include a number of discrepancies which require attention. Examples of these include the fact that some municipalities have been renamed between 2005 and 2011. It is thus necessary to update these names in each report. Missing data is also prevalent in these reports. This results in the omission of these municipalities from the consideration in this study. Other examples of discrepancies in the data include contradictions in employee numbers and employee costs. Added to this the municipalities of Manguang and Buffalo City were classified as "local" municipalities when they are in fact “metropolitan” municipalities. It is thus required that this data be cleaned thoroughly before it is used in an efficiency evaluation.

The first (preliminary) evaluation of the efficiencies of municipal water service provision (using DEA) was undertaken using data from the 2010 municipal year.

4.3.5 Previous Studies

The WRC report TT522/12, The State of Non-Revenue Water in South Africa by Mckenzie et al. (2012), provided detailed information on the water usage and infrastructure of each municipality for the municipal years from 2005 to 2010. The authors had overall concerns about the validity of the data, as described in Section 4.3.3. That said, the report is considered to be the most comprehensive of its kind presently available in South Africa (Mckenzie et al., 2012). The data which is available in this survey is:

- System input volume
- Non-revenue water
- Connections (metered and unmetered)
- Length of mains
- Population served
- Households served
- Billed and unbilled consumption (metered and unmetered)

4.3.6 Municipalities & DWA National Information System

At this stage the data for the years from 2005 to 2010 have been collected and documented. Additional data will have to be sourced from each of the municipalities individually through direct contact and survey questionnaires. The main drawback of this approach is that one would have to rely on the capability of the employees and their willingness to co-operate.

The Department of Water Affairs (DWA) hosts a website for the National Information System (NIS). This is a possible avenue for data on additional variables (for example, the number of water pumps) should these variables be required.

4.4 Description of the data collected

Data were collected for the year 2010 relating to the variables identified in Tables 4.4 and 4.5 above. These data are shown in Appendix C.

4.5 Selection of variables

Considering the data available to the researchers through StatsSA's financial and non-financial audits of the municipalities (Section 4.3.4) as well as the WRC report TT522/12 (Section 4.3.5) a number of variables are available for inclusion into an efficiency analysis relating to the year 2010. To determine the variables which would be most appropriate, the guidelines set out by Thannassoulis (2000a) and given in Table 3.1 as well as a review of related literature were used.

The most prominently used input variable in the related literature, as well as the only suggestion given by Thanassoulis (2000a), was the variable of "Operating Expenditure". This variable was included into the analysis as the total expenditure on both water and sanitation for each municipality. The other input variable included was number of employees. Although this variable was not indicated in the guidelines of Thannassoulis (2000a), it was regularly used in the related literature and similar studies.

The output variables chosen for the analysis were total system volume, number of connections and length of mains. These outputs were all suggested as potential output by Thanassoulis (2000a) and also were prominently used in similar studies reportedly undertaken elsewhere. Furthermore, these variables provide an indication of the size of the municipality, the population served and the infrastructure respectively. These are all important factors to take into account in the provision of water services.

More variables would have been chosen for inclusion into the model, but for gaps in the data. Any municipality which had missing data value for any one of the above variables had to be excluded from the study. This restriction limited what variables could be included to the ones recommended by Thannasouliss (2000a) and which feature prominently in the related literature and in similar studies.

4.6 Conclusion on the future availability and credibility of the data

In the future the data should still be able to be gathered from the Financial and Non-Financial Census of Municipalities (StatsSA Reports P9114 and P9115 respectively). However, the data on the water supply (system input volume, etc.) reported in the WRC TT522/12 Report (McKenzie et al, 2012) is not expected to be available for years from 2010 onwards, as this was a once off initiative. This type of data will have to be gathered through other means, such as direct correspondence with the municipalities and the DWA.

Currently, much of the data reported by municipalities is based on poor record keeping. Less than one fifth of municipalities keep god records of their water inputs and outputs. It would greatly add to the credibility of the SFA and the DEA if the data reported were subject to some form of audit or verification, not only with respect to inputs but also with respect to outputs.

The data has many gaps. Filling the data gaps will allow the methodology to deliver a much improved, less controversial and more encompassing product and it will make it more amenable for implementation over a multi-year period. Experience in working with municipalities indicates that this task will not be an easy one.

CHAPTER 5

ESTIMATING THE STOCHASTIC FRONTIER AND CALCULATING COLS AND SFA EFFICIENCY INDICES

5.1 Methodology

Estimating an efficiency frontier by means of regression analysis falls under the heading of stochastic frontier analysis. Stochastic frontier analysis (SFA) differs from data envelopment analysis (DEA) in the method by which the frontier is estimated and constructed. DEA uses nonparametric techniques to wrap the frontier around the data points, whereas SFA imposes a functional form on the shape of the frontier and uses the data to estimate the parameters of this predefined function (regression analysis). Typically a Cobb-Douglas or translog function is used to estimate the production frontier. The Cobb-Douglas function is used in this study. The stochastic production frontier model is defined as (Greene, 2008):

$$\ln y_i = \alpha + \underline{\beta}' \underline{x}_i + v_i - u_i \tag{5.1}$$

where y_i and $\underline{x}_i$ represent the output and inputs of firm i respectively. The residual of the model is split into two components; a strictly positive inefficiency component, $u_i > 0$, and an error component, v_i which is assumed to be symmetric (typically normally distributed). In this case the function $\ln y_i = \alpha_i + \underline{\beta}' \underline{x}_i + v_i$ represents the stochastic frontier (Greene, 2008), which models best practice.

The deterministic alternative to the above approach is to consider the following frontier model:

$$\ln y_i = \alpha + \underline{\beta}' \underline{x}_i - u_i \tag{5.2}$$

where $\ln y_i = \alpha_i + \underline{\beta}' \underline{x}_i$ represents the deterministic frontier (best practice) and this is found using the method of corrected ordinary least squares (COLS). COLS, in summary, is the method by which the production function is fitted to the data using ordinary least squares, and this fitted model is shifted upwards to until all but one of the residuals is negative (the remaining residual is zero). The resulting adjusted residuals represent the inefficiency of each firm. This method is not often used as it disregards the stochastic nature of statistical models (which typically include a symmetric error term). Ruggiero (1999) suggests that the stochastic methods do not necessarily overcome the problems introduced by this deterministic formulation, and that, until they do, the deterministic model should not be abandoned.

The cost function formulation of the models given in equations 5.1 and 5.2, are needed for implementation in the water services scenario presented in this study. The cost function representation of in the Cobb-Douglass case is given by

$$\ln\left(\frac{C_i}{P_{ir}}\right) = \delta + \sum_{j=1}^{m} \beta_j \ln(y_{ij}) + \sum_{\substack{j=1 \\ j \neq r}}^{n} \gamma_j \ln\left(\frac{P_{ij}}{P_{ir}}\right) + \varepsilon_i \tag{5.3}$$

where $\varepsilon_i = v_i + u_i$ for the stochastic case and $\varepsilon_i = u_i$ for the deterministic case. P_{ij} in this case represent the prices for each of the input variables and P_{ir} can be considered to be a reference price which is used to ensure that the model is linearly homogeneous in input prices. C_i is the total cost (or expenditure) and Y_{ij} represents the output variables. The Cobb-Douglas representation is chosen owing to the lack of sufficient data to fit a more elaborate model (i.e. one with more variables).

5.2 Applying the Methodology

The cost function estimated in in this study is

$$\ln\left(\frac{C_i}{P_{ir}}\right) = \delta + \beta_1 \ln(Y_i) + \beta_2 \ln(L_i) + \gamma \ln\left(\frac{P_{il}}{P_{ir}}\right) + \varepsilon_i \tag{5.4}$$

where C_i is the total expenditure water and sanitation for municipality i. The expenditure value used in this study was calculated as the aggregate expenditure of the ’’Waste Water Management” and ’’Water” divisions for each municipality (as gathered from the StatsSA 2010 Financial Audit of Municipalities). Y_1 is the total system volume input for municipality i, L_i is the length of the mains for municipality i, P_{il} is the price of labour for municipality i and P_{ir} is the price of other operating expenditure (e.g. repairs, etc.) for municipality i. The price of labour is calculated by dividing the cost of labour by the number of employees (permanent or part time) for each municipality. The price of other operating expenditure is calculated as the cost of other expenditure divided by the number of water connections served by the municipality. The formulation of the composed error term $\varepsilon_i = v_i + u_i$ will be that of a $v_i \sim N\left(0, \sigma_v^2\right)$ and $u_i \sim N_+\left(0, \sigma_u^2\right)$, which is termed the "normal-half-normal" formulation of the error term. The parameters $\sigma_u^2 = \sigma_v^2 + \sigma_u^2$ and $\lambda = \sqrt{\frac{\sigma_u^2}{\sigma_v^2}}$ are required to be estimated from the residuals of the model. These parameters are estimated using the method of maximum likelihood. Bogetoft and Otto (2010) explain that values of λ close to zero (i.e. σ_u^2 close to zero) will result in the ordinary least squares model. As the value of λ gets larger, this will indicate that the majority of the residual is due to the efficiency differences (Bogetoft and Otto, 2010).

The stochastic frontier approach will be used unless the results dictate this model is not suitable. The stochastic approach can be considered unsuitable if the distribution of the residuals does not exhibit a skewness which can be attributed to inefficiency (i.e. $\lambda \approx 0$). This makes estimation of efficiencies difficult in the stochastic frontier approach. In this case, the deterministic frontier approach will be used to determine efficiencies. Efficiencies in this case are considered to be "cost efficiencies" and could thus comprise of both allocative and technical efficiency components (Greene, 2008).

Of the municipalities which were included in the DEA evaluation, Capricorn district municipality and Umgungundhlovu district municipality were excluded from the stochastic frontier analysis. This was necessary as these municipalities reported having a number of employees, while at the same time reported a zero for employee costs. As such, the price variable for labour, P_{il}, could not be determined for these municipalities.

5.3 Results

The efficiency estimates using the procedure described above are given in Table 5.3 for district and metropolitan municipalities and in Table 5.5 for local municipalities.

5.3.1 District and Metropolitan Municipalities

The SFA model was fitted to the same data as was employed in the DEA evaluation, using the benchmarking package (Bogetoft and Otto, 2013) in R. The results of this fit are given in Table 5.1.

Table 5.1: Model fitted to district and metropolitan municipalities using SFA.

Parameter	Estimate	Std Error	*p*-value
δ	-5.5535	0.85089	0.000
β_1	1.0607	0.2143	0.001
β_2	-0.6084	0.2854	0.070
γ	0.0148	0.04393	0.746
λ	-1.8607		
σ^2	0.12594		

The fit has a negative estimate for the λ parameter, indicating one of two things: misspecification of the model or no discernible inefficiency in the models residuals. The former is a possibility since there are very few data points to fit the model (only 13), but the latter is more likely the cause. The latter typically occurs when the residuals are close to symmetric, making the ability to distinguish between error (v_i) and inefficiency (u_i) difficult.

Owing to this drawback it was decided that a deterministic frontier would be used to establish efficiency measures for the district and metropolitan municipalities. The deterministic frontier was determined using the method of corrected ordinary least squares (COLS) methodology. The COLS model parameters are given in Table 5.2.

Table 5.2: Model fitted to district and metropolitan municipalities using COLS.

Parameter	Estimate
δ	5.695879
β_1	1.155587
β_2	-0.726603
γ	0.009842

In this model the relative efficiencies of each of the municipalities are represented by the residuals of the COLS model (u_i). The efficiencies are calculated according to the following formula (Greene, 2008):

$$eff_i = \exp\{\min(e_i) - e_i\} \quad (5.5)$$

where e_i are the residuals of the OLS model. The efficiency estimates so calculated for the municipalities are presented in Deliverable 6.

Calculation of the efficiency ratings – SFA model

The efficiency estimates so calculated for the municipalities are presented in given in Table 5.3.

Table 5.3: Efficiencies of district and metropolitan municipalities using COLS.

Municipality		**eff**
Nelson Mandela Bay Metro Municipality	EC	0.5281
Buffalo City Metro Municipality	EC	0.3759
O.R. Tambo District Municipality	EC	0.5469
Mangaung Metro Municipality	FS	0.6657
City Of Johannesburg Metro Municipality	GT	0.7980
City Of Tshwane Metro Municipality	GT	0.6219
Ekurhuleni Metro Municipality	GT	0.6326
eThekwini Metro Municipality	KZN	0.6094
Ugu District Municipality	KZN	1.0000
Uthukela District Municipality	KZN	0.5258
Uthungulu District Municipality	KZN	0.9172
West Coast District Municipality	WC	0.5086
Central Karoo District Municipality	WC	0.6777

The COLS derived relative efficiency estimates are listed in Table 5.3 and displayed in Figure 5.1. They show that the Ugu District municipality in KwaZulu-Natal to be the most (fully) efficient. In the COLS model, the frontier will rest on this observation. The next best district municipality is the Uthungulu district municipality, which achieved an efficiency estimate of 0.9152. This municipality is also situated in KwaZulu-Natal.

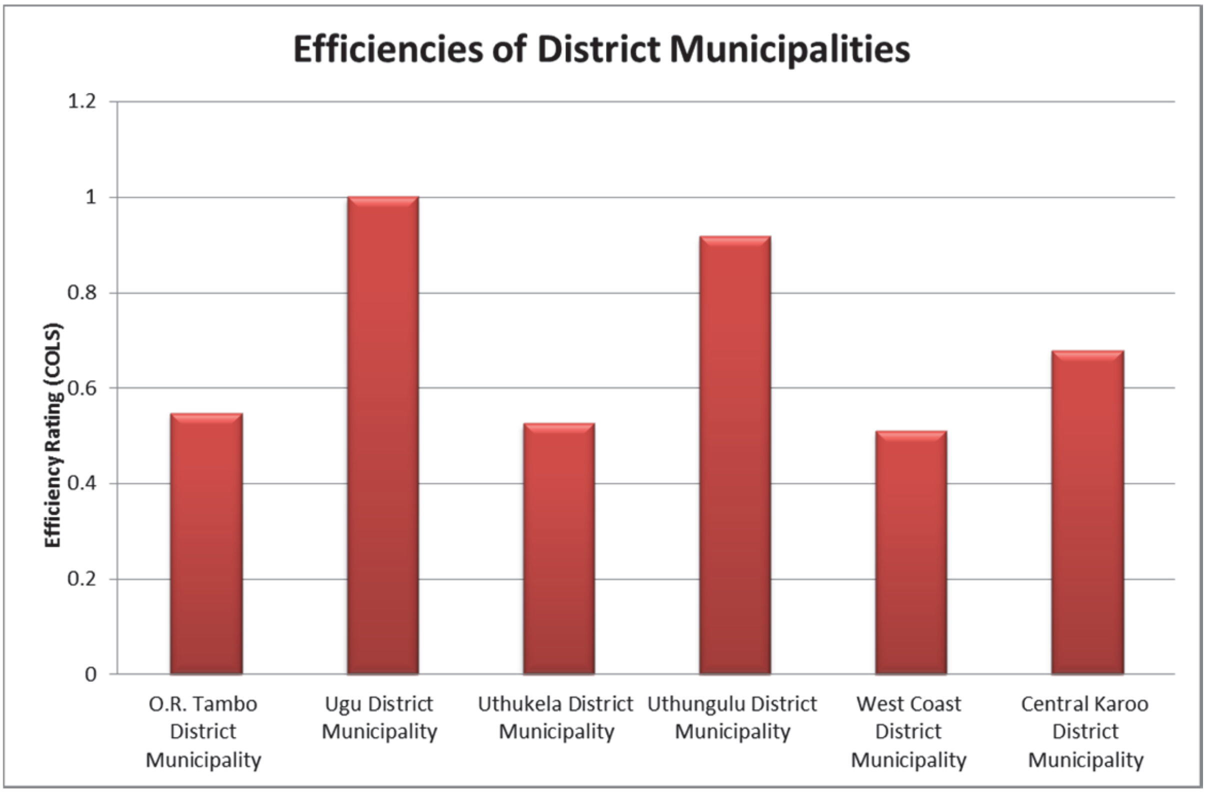

Figure 5.1: Graph comparing the efficiencies of the district municipalities. (Efficiencies calculated using COLS method)

The estimated efficiencies of the metropolitan municipalities are shown in Figure 5.2. The best performing metropolitan municipality was the city of Johannesburg metropolitan municipality (0.7980) followed by the Ekurhuleni metropolitan municipality (0.6326), both of which are situated in Gauteng province. The worst performing municipality in terms of efficiency estimates, in this case, was the Buffalo City metropolitan municipality (0.3759).

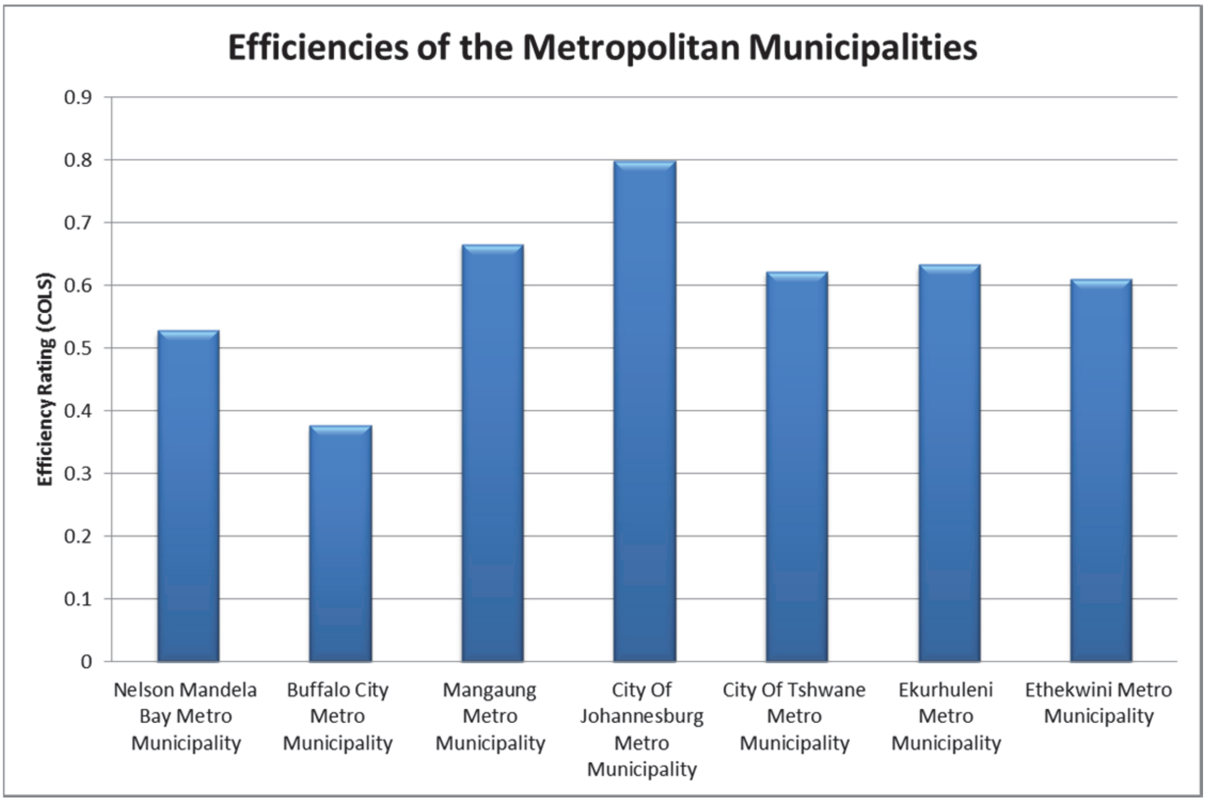

Figure 5.2: Graph comparing the efficiencies of the metropolitan municipalities. (Efficiencies calculated using COLS method)

The best performing province according to the average efficiencies of the included municipalities is KwaZulu-Natal, with an average efficiency estimate of 0.7631. The worst performing municipality evaluated in the same manner is the Eastern Cape with an average efficiency estimate of 0.4836. The average efficiency for all the district and metropolitan municipalities is 0.6468, which indicates that there is a large scope for improvement in the efficiency levels of the district and metropolitan municipalities in South Africa.

Figure 5.1 provides a histogram of the efficiency estimates of the district and metropolitan municipalities using COLS. It is clear that the two most efficient municipalities (Ugu and Uthungulu) are separated from the remaining municipalities by a considerable distance. This could indicate that these municipalities are outliers. The presence of these observations creates a symmetric shape to the efficiency estimates.

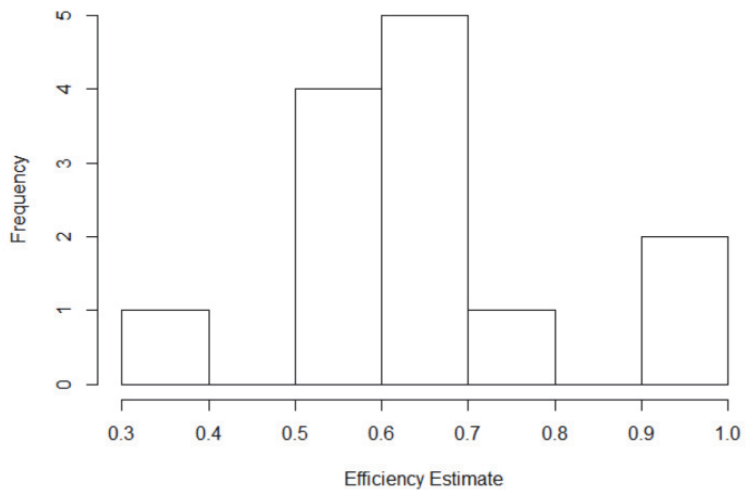

Figure 5.3: Histogram of district and metropolitan municipality efficiency estimates using COLS.

For the purpose of relative efficiency determination with an SFA model, symmetry is not a virtue. The SFA requires skewness in the residuals in order to identify inefficiency. This could be the reason that the SFA failed to determine efficiency estimates for this data.

5.3.2 Local Municipalities

The SFA model was also fitted to the local municipality data using the benchmarking package (Bogetoft and Otto, 2013) in R. The results of this fit are given in Table 5.4.

Table 5.4: SFA model fitted to local municipality data.

Parameter	Estimate	Std Error	*p*-value
δ	-3.529	0.0156	0.000
β_1	0.1744	0.0489	0.001
β_2	0.7371	0.0369	0.000
γ	0.1216	0.0433	0.008
λ	3159.9763		
σ^2	0.29966		

The fit has a very large estimate for the λ parameter, indicating that the u_i term dominates the v_i term in their composition of the error term ε_i, i.e. that the majority of the error term occurs as a result of differences in efficiency between the local municipalities. The efficiency estimates calculated using the SFA model for local municipalities is given in Table 5.5. The local municipalities with the highest estimated efficiency using the SFA model are the Camdeboo (Eastern Cape), Kungwini (Gauteng) and Khai-Ma (Northern Cape) local municipalities, each with an estimated efficiency of 0.9995, that is, very close to fully efficient. These are followed closely by the Midvaal local municipality in Gauteng, which achieved an efficiency estimate of 0.9709. The worst performing local municipality was the Msukaligwa local municipality in Mpumalanga, which achieved an efficiency estimate of 0.1648 (indicating a very low level of efficiency).

Table 5.5: Efficiencies of local municipalities using SFA.

Municipality		**eff**
Camdeboo Local Municipality	EC	0.9995
Makana Local Municipality	EC	0.5832
Ndlambe Local Municipality	EC	0.6058
Baviaans Local Municipality	EC	0.6976
Letsemeng Local Municipality	FS	0.7592
Mantsopa Local Municipality	FS	0.4532
Tswelopele Local Municipality	FS	0.7866
Matjhabeng Local Municipality	FS	0.7829
Setsoto Local Municipality	FS	0.6810
Moqhaka Local Municipality	FS	0.7325
Ngwathe Local Municipality	FS	0.7221

Municipality		eff
Metsimaholo Local Municipality	FS	0.9280
Mafube Local Municipality	FS	0.5839
Emfuleni Local Municipality	GT	0.7539
Midvaal Local Municipality	GT	0.9709
Lesedi Local Municipality	GT	0.5717
Nokeng Tsa Taemane Local Municipality	GT	0.8506
Kungwini Local Municipality	GT	0.9995
Mogale City Local Municipality	GT	0.6031
Randfontein Local Municipality	GT	0.8264
Westonaria Local Municipality	GT	0.8284
Merafong City Local Municipality	GT	0.7799
Umhlathuze Local Municipality	KZN	0.4332
Mogalakwena Local Municipality	LIM	0.4053
Msukaligwa Local Municipality	MP	0.1648
Govan Mbeki Local Municipality	MP	0.8027
Steve Tshwete Local Municipality	MP	0.5651
Dr JS Moroka Local Municipality	MP	0.2025
Mbombela Local Municipality	MP	0.6141
Rustenburg Local Municipality	NW	0.8484

Municipality		eff
Richtersveld Local Municipality	NC	0.8867
Nama Khoi Local Municipality	NC	0.8399
Hantam Local Municipality	NC	0.9131
Karoo Hoogland Local Municipality	NC	0.8124
Khai-Ma Local Municipality	NC	0.9995
Cederberg Local Municipality	WC	0.8338
Drakenstein Local Municipality	WC	0.7889

The province which achieved the highest average efficiency rating among the municipalities considered in this study was the Northern Cape province. The Northern Cape achieved a 0.8903 average efficiency estimate, followed closely by the Western Cape and Gauteng, with average efficiency estimates of 0.8114 and 0.7983 respectively. The Eastern Cape achieved the second worst average rating with an average efficiency estimate of 0.7215, which is in fact not a score that gives cause for concern. The worst performing province, by a substantial distance, was Mpumalanga which achieved an average efficiency rating of 0.4698, indicating that, as a province, its local municipalities are performing well below what could reasonably be classed as efficient. It must be noted, however, that the provinces of KwaZulu-Natal, Limpopo and North West were excluded from the analysis of average efficiency as each only had a single municipality included in the evaluation.

The average efficiency for all the local municipalities included in this assessment was 0.7192, which could be thought of as being a satisfactory level.

Figure 5.4 shows the efficiencies of local municipalities that are categorised B1, which indicates a municipality with a large budget or a secondary city. The average efficiency rating for this category of local municipality was calculated to be 0.6880. From Figure 5.2 it can be seen that most of the municipalities perform in a region close to this average; exceptions being the Rustenburg local municipality (0.8484) which performed the best for this category and the Umhlathuze local municipality (0.4332) which performed well below the average.

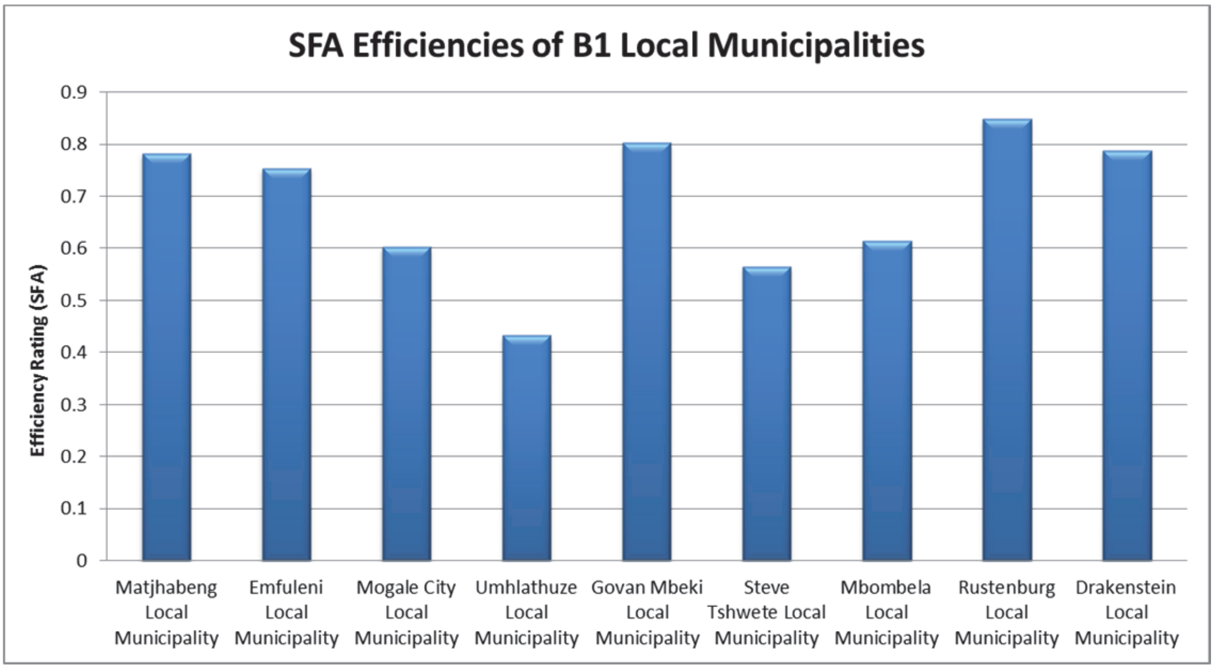

Figure 5.4: Graph comparing the efficiencies of Category B1 Local Municipalities. (Efficiencies calculated using SFA method)

Figure 5.5 shows the efficiencies of the category B2 local municipalities. Municipalities in this category are described as municipalities with large towns. The majority of the municipalities falling in this category performed well in terms of efficiency ratings. The Kungwini (0.9995), Midvaal (0.9709) and Metsimaholo (0.9280) local municipalities performed particularly well, achieving efficiency ratings above 0.9. The average efficiency rating for this category of local municipalities was 0.7336, which is higher than that of the B1 municipalities. Most of the municipalities performed above this average. The two worst performing municipalities in this category were the Mogolakwena (0.4053) and Msukaligwa (0.1648) local municipalities, both of which performed significantly below the average for this category.

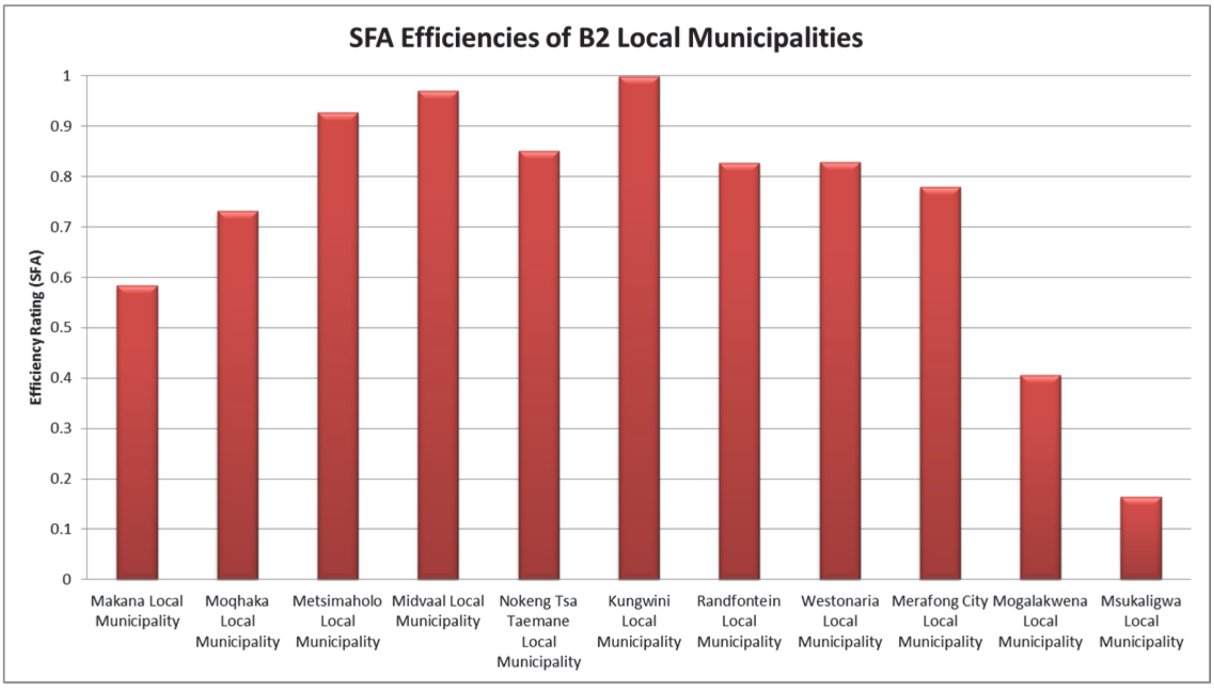

Figure 5.5: Graph comparing the efficiencies of Category B2 Local Municipalities. (Efficiencies calculated using SFA method)

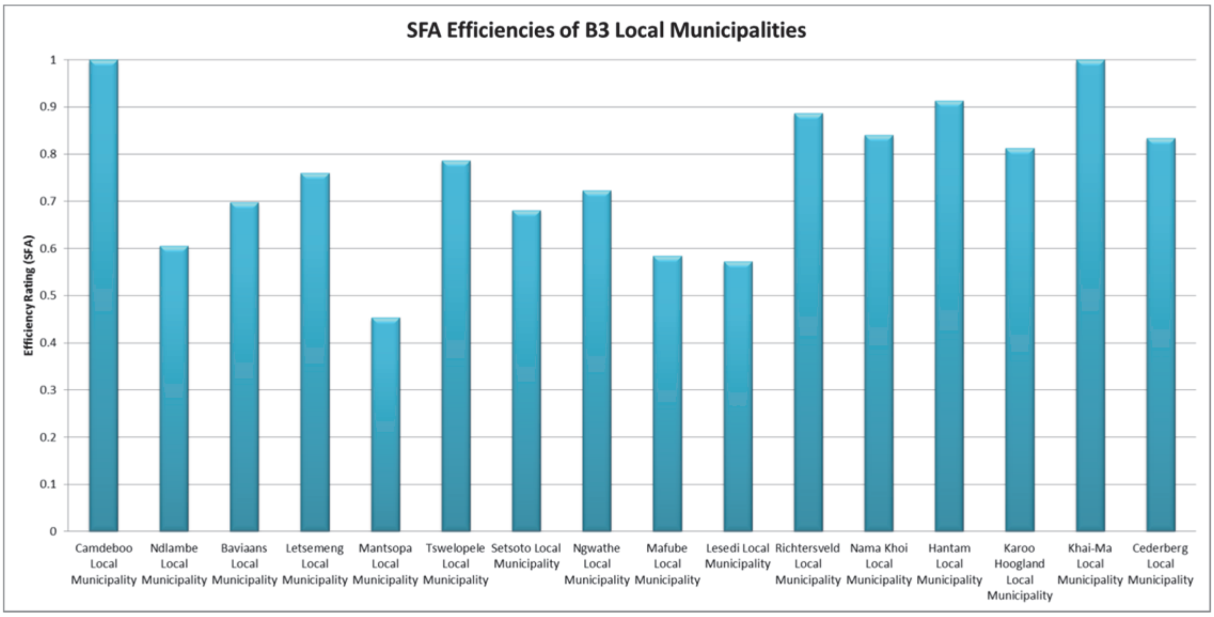

Figure 5.6: Graph comparing the efficiencies of Category B1 Local Municipalities. (Efficiencies calculated using SFA method)

The graph provided in Figure 5.6 shows the calculated efficiency ratings of the category B3 local municipalities. Municipalities in this category are small towns with a significant urban population. The average efficiency rating for this category of municipalities was calculated to be 0.7591. The two stand-out municipalities were the Camdeboo (0.9995) and the Khai-Ma (0.9995) local municipalities which were performing at almost full efficiency. The remaining municipalities performed close to the average value, the exception being the Mantsopa local municipality which achieved the lowest efficiency rating for this category (0.4532).

There was only a single local municipality in the sample that fell into the category of B4, which are mainly rural municipalities with one or two small towns. The Dr JS Moroka local municipality in this category achieved an efficiency rating of 0.2025, which was the second lowest of all local municipalities in the evaluation.

5.4 Conclusion on the SFA Efficiency Ratings

The performance of district and metropolitan municipalities in terms of efficiency indicate that there is room for improvement. There were only 13 municipalities included in the assessment for efficiency. A larger data set could have rectified the problem of an absence of skewness of the residuals. The histogram presented in Figure 5.1 indicates two possible outliers on the right which, when removed, could result in the skewness required of the residuals. However, removal of these observations would further compound the problem of insufficient data. A similar case could be for removal of the single observation on the far left of the histogram as also an outlier. Removing all these (possible) outliers would leave the same problem of symmetric residuals, i.e. no efficiency difference indication present in the residuals.

The use of the deterministic approach to estimate efficiency is not an ideal approach. However, when the stochastic approach failed, it was the only parametric alternative method remaining that could be considered. The local municipalities were shown to be more efficient, on average, than the district and metropolitan municipalities and the stochastic frontier did not encounter any symmetry problems in the residuals during estimation (as was the case for the district and metropolitan municipalities).

Although the district and metropolitan municipalities may exhibit similar quantities in terms of the data considered in this study, their operating environments are substantially different. This difference can make comparison in an efficiency evaluation misleading. For this reason, the municipalities should be grouped according to operating environment (urban and rural) in order to develop efficiency estimates which are more relatively reflective of what they aim to do.

CHAPTER 6

ESTIMATION OF THE DATA ENVELOPLMENT ANALYSIS FRONTIER AND EFFICIENCY INDICES

6.1 Variable Choice and Data

There have been a number of studies which have investigated the efficiency of water service providers. They have been conducted in both developed and developing nations. These studies have been used to determine which variables should be used as possible input and outputs in the DEA model. The variables used in this evaluation were chosen owing to the prominence of their use in the related literature.

Two input variables that have achieved prominence in the relevant literature are total operating expenditure and total number of employees. Aida et al. (1998), Cubbin and Tzanidakis (1998), Thanassoulis (2000b), Tupper and Resende (2004), García-Valñasa and Muñiza (2007), Berg and Lin (2008) and Singh *et al.* (2011) all included a measure of the operating costs or expenditure into their DEA models. The total number of employees was used as an input in Aida *et al.* (1998), Berg and Lin (2008) and Singh *et al.* (2011) with other studies (i.e. Tupper and Resende (2004)) including labour expenses as an input.

The output variables used in this study were selected in a similar manner. A measure of the total volume of water delivered (or billed) is an output which is included in the vast majority of DEA studies aimed at identifying the relative efficiency of water service providers. All of the studies mentioned above include a measure of input volume as an output. The length of the main water network is a measure of the infrastructure available to the municipality. It has been used as an input in Aida *et al*. (1998) and an output in Cubbin and Tzanidakis (1998), Thanassoulis (2000b) and García-Valñasa and Muñiza (2007). The last output considered was the number of connections served by the water service authority. This measure was used as an output in the DEA study of Thanassoulis (2000b) but as an input in Berg and Lin (2008). The outputs selected for use in this study were influenced by the above analyses. They were: system input volume (in kl), length of mains (in km) and number of connections.

Adding further justification to the above selection of variables, Thanassoulis (2000a) provided a review of the use of DEA in the regulation of water service providers. The author's suggestions for input and output variables were based on factors initially identified by the UK Office of Water Services (OFWAT) and are given in Table 2.1.

This study added an additional input variable to those given in Table 2.1; viz. number of employees, and omitted the output variables: metered water delivered to clients and expenditure on repairs. There are constraints on the number of variables which can be used in a DEA evaluation (Dyson *et al*., 2001). If m and n indicate the number of input and output variables receptively, and k represents the number of municipalities included in the efficiency assessment, the number of variables used should be constrained to:

$$2 \times m \times n \leq k. \tag{6.1}$$

As two inputs and three outputs were considered for inclusion in this study, a minimum number of 2 × 2 × 3 = 12 municipalities are required to run a DEA evaluation. The inclusion of more municipalities would improve the validity of the assessment.

The data on inputs and outputs were collected from the StatsSA website and a 2012 water research commission (WRC) report on the state of non-revenue water in South Africa (McKenzie *et al*., 2012). More specifically, the data regarding the volume of water input into the system and the nature of the infrastructure (length of mains and number of connections) were gathered from McKenzie *et al*. (2012). The data for the two inputs, total expenditure and number of employees, were gathered from the StatsSA website. The operating expenditure data was collected from the StatsSA document ''P9114 – Financial Census of Municipalities" (StatsSA, 2011a) and the employee numbers were collected from the StatsSA document ''P9115 – Non-Financial Census of Municipalities" (StatsSA, 2011b). The expenditure value used in this study was calculated as the aggregate expenditure of the ''Waste Water Management" and ''Water" divisions for each municipality. The number of employees used was calculated in the same way. This aggregated figure was used since some prominent municipalities, e.g. Nelson Mandela Metropolitan and City of Johannesburg Metropolitan, do not distinguish between employees in these two divisions and only report a single value.

For each municipality i in 1, …, k the following data vector is collected

$$\underline{X}_i = \left[Y_{i1} \quad Y_{i2} \quad X_{i1} \quad X_{i2} \quad X_{i3} \right] \tag{6.2}$$

where Y_1 and Y_2 respectively represent the input variables; expenditure and number of employees, and X_1, X_2 and X_3 respectively represent the output variables; system volume input, length of mains and number of connections, respectively. Efficiency scores using the statistical software *R* (R Core Team, 2013) and package *Benchmarking* (Bogetoft and Otto, 2013) were calculated using these data. The data used in this study were for the municipal year ending 30 June 2010; this being the most recent relevant data available.

6.2 Model

Owing to the fact that the volume of water input into the system (as well as other outputs) is exogenous and, as such, an efficient municipality is one that achieves these levels of output whilst using a minimum amount of input, the input minimisation DEA approach to determine efficiency scores is the most popular. An appealing sophistication to this model is to allow for varying returns to scale (VRS), as proposed by Banker *et al*. (1984). The input orientated Banker-Charnes-Cooper (BCC) DEA linear model is discussed and defined in section 2.2 equations 2.34 to 2.38.

The DEA analysis is run using the dea() function in the benchmarking package in *R* (Bogetoft and Otto, 2013). The efficiency value ϑ^* (which represents represents the Farrell efficiency of each municipality) which results is confined to the interval $(0,1]$. Values of ϑ^* close to 1 indicate efficient WSA's, whilst values of ϑ^* far from 1 indicate WSA's with poor efficiency.

6.3 Statistical Inference

Simar and Wilson (2011) emphasise that, although it is common practice to speak of efficiency calculations or values when conducting a DEA, it is more appropriate to speak of efficiency *estimates*. This correction in terminology is required since the production set and efficient frontier constructed from a sample of firms is itself an *estimate* of the true production set and efficient frontier. As such, it follows that efficiency calculations made using this estimated frontier will themselves be estimates of the true values. Being estimates, these values are subject to variations and uncertainty which open the door to statistical inference. The inference used in this report is restricted to the estimations of bias and confidence intervals of the efficiency values, but could be expanded to include tests for returns to scale (increasing, decreasing or constant), amongst others.

The production set is defined as

$$T \equiv \{(x, y) | x \text{ can produce } y\} \tag{6.3}$$

which includes all input-output combinations (x, y) which are possible. The boundary of the production set provides the production frontier against which efficiency is measured. The boundary of a production set can be described in terms of input and output. This report considers an input-oriented DEA model, so the boundary is defined in terms of input minimisation:

$$X^{\Omega}(y) = \{x | (x, y) \in T, (\theta x, y) \notin T \ \forall \ 0 < \theta < 1\}. \tag{6.4}$$

For an input value x to fall on the boundary (i.e. to be efficient) it must be the minimum possible input quantity possible to produce the output y. By implication, any radial contraction in the input θx $(0 < \theta < 1)$ will result in the input-output pair $(\theta x, y)$ being infeasible. The varying returns to scale assumption used in the DEA model estimates the production set by ''wrapping" a convex hull around the data points. This estimator, denoted $\hat{T}$, exhibits the following property (Simar and Wilson, 2011):

$$\hat{T} \subseteq T. \tag{6.5}$$

If the true efficiency and DEA estimated efficiency of a firm are denoted as ϑ and $\hat{\vartheta}$ $(= \vartheta^*)$ respectively, from $\hat{T} \subseteq T$. it follows that $\hat{\vartheta} \geq \vartheta$ and hence DEA estimated efficiencies are upwardly biased (Bogetoft and Otto, 2010), that is, the DEA estimates of efficiencies are typically greater than their corresponding true values. For this reason it is sensible to construct interval estimates for the DEA efficiency values which account for the bias as well as the stochastic nature of statistical estimates. This adjustment may be achieved using a bootstrapping methodology.

Bootstrapping is a nonparametric sub-sampling technique which allows for statistical inference through re-sampling from an observed data set. The bootstrapping procedure involves the creation of B samples through resampling, with replacement, from the originally observed data $\underline{x}$ (Efron and Tibshirani, 1993). These bootstrap samples, denoted $\underline{x}_b^*$

for $b = 1, \ldots, B$, are used to estimate the underlying distribution of the statistic (the efficiencies) in question. The statistic for each of the B bootstrap samples is estimated and used to construct an estimate of the sampling distribution of the statistic. The quartiles of this sampling distribution can then be used to estimate confidence intervals for the statistic of interest which, in this case, is the efficiency estimate of the water service providers[1]. The construction of confidence intervals of DEA efficiency estimates is discussed below.

Since DEA efficiency estimates are upwardly biased, this bias must be corrected for when constructing confidence intervals. The bias of the DEA efficiency estimator may be defined as

$$bias\left(\hat{\vartheta}\right) = E\left[\hat{\vartheta}\right] - \vartheta \qquad (6.6)$$

and can be estimated using the bootstrap samples as

$$bias\left(\hat{\vartheta}\right) = \frac{1}{B}\sum_{b=1}^{B} \vartheta_b^* - \hat{\vartheta} = \overline{\vartheta}^* - \hat{\vartheta} \qquad (6.7)$$

where ϑ_b^* is the efficiency estimate of the firm using the b^{th} bootstrap replication and $\overline{\vartheta}^*$ is the average of the bootstrapped estimates of ϑ. The bias corrected estimate for efficiency is calculated as (Bogetoft and Otto, 2010):

$$\hat{\vartheta}_{bc} = 2\hat{\vartheta} - \overline{\vartheta}^{**}. \qquad (6.8)$$

A method for constructing confidence intervals, whilst simultaneously adjusting for bias was introduced in Simar and Wilson (1999). In order to construct the $(1-\alpha)\times 100\%$ confidence intervals of bias-corrected efficiency measures, Simar and Wilson (2008) reparameterise the efficiency scores using the distance function (φ) proposed by Shepard (1970); this being the inverse of the DEA estimated efficiency:

$$\varphi = \frac{1}{\vartheta}. \qquad (6.9)$$

[1] The construction of bootstrap confidence intervals for efficiency estimates involves a far more detailed approach than the simplistic approach outlined here. A comprehensive account of the process of bootstrapping for DEA efficiency estimates is given in Simar and Wilson (2008).

It can easily be seen that $\varphi \geq 1$. The estimates for these parameters are denoted by $\hat{\varphi}$ and $\hat{\vartheta}$. It follows that the estimate for the Shepard efficiency is $\hat{\varphi} = \frac{1}{\hat{\vartheta}}$. This reparameterisation corrects for instances in the bootstrapping process which result in negative confidence limits for the efficiency intervals. The confidence intervals can be derived from the distribution of $\hat{\varphi} - \varphi$. This distribution is unknown and is estimated using the empirical distribution of $\varphi_b^* - \hat{\varphi}$ for $b = 1, \ldots, B$, where $\varphi_b^* = \frac{1}{\vartheta_b^*}$ is the Shepard efficiency calculated for the b^{th} bootstrap replication. Using the empirical distribution, the values of the quantiles $\hat{q}_{\frac{\alpha}{2}}$ and $\hat{q}_{1-\frac{\alpha}{2}}$ are found such that

$$P\left(\hat{q}_{\frac{\alpha}{2}} \leq \varphi_b^* - \hat{\varphi} \leq \hat{q}_{1-\frac{\alpha}{2}}\right) = 1 - \alpha\,. \tag{6.10}$$

Using the empirical distribution of $\varphi_b^* - \hat{\varphi}$ to estimate the distribution of $\hat{\varphi} - \varphi$ it follows that (Simar and Wilson, 2008)

$$P\left(\hat{q}_{\frac{\alpha}{2}} \leq \hat{\varphi} - \varphi \leq \hat{q}_{1-\frac{\alpha}{2}}\right) \approx 1 - \alpha \tag{6.11}$$

Solving equation 6.11 for φ,

$$P\left(\hat{q}_{\frac{\alpha}{2}} \leq \hat{\varphi} - \varphi \leq \hat{q}_{1-\frac{\alpha}{2}}\right) \approx 1 - \alpha$$

$$\Rightarrow P\left(\hat{\varphi} - \hat{q}_{1-\frac{\alpha}{2}} \leq \varphi \leq \hat{\varphi} - \hat{q}_{\frac{\alpha}{2}}\right) \approx 1 - \alpha\,. \tag{6.12}$$

In terms of the true Farrell efficiency, ϑ, it can be shown from the above that (Simar and Wilson, 2008)

$$P\left(\frac{1}{\hat{\varphi} - \hat{q}_{\frac{\alpha}{2}}} \leq \vartheta \leq \frac{1}{\hat{\varphi} - \hat{q}_{1-\frac{\alpha}{2}}}\right) \approx 1 - \alpha \tag{6.13}$$

and the estimated bias-corrected confidence interval for ϑ is

$$\left[\frac{1}{\hat{\varphi} - \hat{q}_{\frac{\alpha}{2}}}; \frac{1}{\hat{\varphi} - \hat{q}_{1-\frac{\alpha}{2}}}\right] \tag{6.14}$$

Correcting for bias introduces a substantial amount of noise to the estimate (Efron and Tibshirani, 1993) and it is suggested in Simar and Wilson (2008) that bias-correction be avoided unless

$$\frac{\left|bias\left(\hat{\vartheta}\right)\right|}{\hat{\sigma}_b} > \frac{1}{\sqrt{3}} = 0.577 \qquad (6.15)$$

where $\hat{\sigma}_b^2$ is the variance of the bootstrap estimates of $\hat{\vartheta}$, i.e. $\hat{\sigma}_b^2 = \frac{1}{B}\sum_{b=1}^{B}\left(\vartheta_b^*\right)^2 - \left(\frac{1}{B}\sum_{b=1}^{B}\vartheta_b^*\right)^2$.

6.4 Results

The validity of DEA is only as good as the validity of the data utilised. This data has been described in the previous section. Using this data the results of the DEA for district and metropolitan municipalities are shown in Table 6.1 and for local municipalities in Table 6.2. Since district municipalities are made up of local municipalities geographically it is not meaningful to compare these two types of municipalities.

The results given are the original DEA efficiency estimates E^{VRS} (which is equivalent to $\hat{\vartheta}$) and the bias corrected efficiency estimates E_{bc}. The bias correction in this case was done for the Shepard efficiency estimate and then converted to the Farrell efficiency estimate. Bias correction was done in this way to match the method of bias correction implemented in the construction of the confidence interval estimates. The lower and upper limits for the 95% bootstrap estimated confidence intervals for the efficiency estimates are given as $CL_{0.025}^{bc}$ and $CL_{0.975}^{bc}$ respectively. The ratio of the bias to the standard error is listed in the last column. The bias corrected value is preferred to the original estimate if this ratio is large (Simar and Wilson, 2008). If this ratio is larger than 0.577, the bias corrected estimate can be used without too much concern about introducing additional noise to the estimate. As can be seen in the tables, this ratio is substantially larger than 0.577 in all cases. As such the bias corrected estimate is preferred in all cases.

6.4.1 District and Metropolitan Municipalities

The efficiencies given in table compare the efficiencies district and metropolitan municipalities which are WSA's in South Africa (for which there was sufficient data available). The original efficiency estimates indicate that the Gauteng based metropolitan municipalities (Johannesburg, Pretoria, Ekurhuleni) perform perfectly efficiently.

Table 6.1: Efficiencies of district and metropolitan municipalities using DEA, 95% bias corrected confidence intervals are provided for each efficiency estimate.

Municipality		E^{VRS}	E_{bc}	$CL^{bc}_{0.025}$	$CL^{bc}_{0.097}$	$\lvert bias \rvert / \hat{\sigma}_b$
Nelson Mandela Bay Metro Municipality	EC	0.569	0.499	0.429	0.563	1.616
Buffalo City Metro Municipality	EC	0.481	0.421	0.363	0.476	1.738
O.R. Tambo District Municipality	EC	1.000	0.812	0.652	0.991	1.484
Mangaung Metro Municipality	FS	0.700	0.626	0.545	0.693	1.583
City Of Johannesburg Metro Municipality	GT	1.000	0.805	0.618	0.988	1.436
City Of Tshwane Metro Municipality	GT	1.000	0.821	0.669	0.990	1.612
Ekurhuleni Metro Municipality	GT	1.000	0.806	0.628	0.989	1.435
eThekwini Metro Municipality	KZN	1.000	0.804	0.616	0.990	1.409
Ugu District Municipality	KZN	0.642	0.578	0.515	0.636	1.880
Umgungundhlovu District Municipality	KZN	1.000	0.805	0.617	0.988	1.414
Uthukela District Municipality	KZN	0.369	0.321	0.269	0.364	1.616
Uthungulu District Municipality	KZN	0.834	0.719	0.611	0.826	1.674
Capricorn District Municipality	LIM	1.000	0.806	0.615	0.989	1.407
West Coast District Municipality	WC	1.000	0.812	0.665	0.989	1.543
Central Karoo District Municipality	WC	1.000	0.805	0.616	0.989	1.413

The metropolitan municipalities in the Eastern Cape (Nelson Mandela Bay and Buffalo City) are performing well below efficiently. Within the district municipalities, there are a large number of WSA's who are performing efficiently. The average efficiency level of the district and metropolitan municipalities is 0.8397. It is deduced that, apart from a few municipalities which require attention and guidance (like Uthukela), the general level of efficiency in municipal water service provision in South Africa is high. This deduction is, however, based on a small sample size which can inflate estimates of efficiency. A larger sample of size, of approximately 25-30 metropolitan and district municipalities, would have yielded a better picture of the true efficiency levels.

The municipality with the highest bias corrected efficiency is Tshwane (metropolitan). It is followed closely by O.R. Tambo (district). The lowest performing municipality in terms of bias corrected efficiency is Uthukela (district). The metropolitan municipalities situated in the Eastern Cape (Nelson Mandela Bay and Buffalo City) have bias corrected efficiency estimates which are less than 0.5; indicating substantial inefficiency.

The confidence interval estimates for the fully efficient municipalities overlap to a great extent, making it difficult to attach confidence to the ranking of the municipalities, but perhaps can be taken as a potential indicator of the levels which each municipality's efficiency could reach (or fall to).

6.4.2 Local Municipalities

Local municipalities in South Africa operate at a much smaller scale than their metropolitan counterparts. The amount of water distributed and size of population served by local municipalities is typically much smaller than those of the metropolitan municipalities. Each district municipality is made up of a number of local municipalities and thus comparison between the two is pointless. Table 6.2 provides the results of the DEA analysis conducted on the group of 37 local municipalities for which data was available. The average efficiency for the local municipalities is 0.7291, indicated a moderate overall efficiency level. The province with the highest average value of the efficiency estimates is the Eastern Cape (0.9033). This comes in contrast to the efficiency estimates for the metropolitan municipalities in this region. The province with the lowest average efficiency estimate value is the Northern Cape with an average efficiency estimate of 0.6364. The provinces with only a single observed local municipality (Limpopo, North West and KwaZulu-Natal) were not included in the assessment of average efficiency. The Eastern Cape has three municipalities that are classified as efficient, followed by Gauteng with two efficient local municipalities and the Free State and North West Provinces each have one fully efficient municipality. The municipality with the lowest estimated efficiency level is the Mogalakwena local municipality in Limpopo with an estimated efficiency of 0.2850.

Table 6.2: Efficiencies of local municipalities using DEA, 95% bias corrected confidence intervals are provided for each efficiency estimate.

Municipality		E^{VRS}	E_{bc}	$CL^{bc}_{0.025}$	$CL^{bc}_{0.097}$	$\lvert bias \rvert / \hat{\sigma}_b$
Camdeboo Local Municipality	EC	1.000	0.781	0.650	0.984	1.807
Makana Local Municipality	EC	0.613	0.543	0.484	0.604	2.031
Ndlambe Local Municipality	EC	1.000	0.808	0.702	0.983	2.036
Baviaans Local Municipality	EC	1.000	0.774	0.639	0.985	1.728
Letsemeng Local Municipality	FS	0.925	0.847	0.773	0.913	1.890
Mantsopa Local Municipality	FS	0.556	0.497	0.444	0.546	1.955
Tswelopele Local Municipality	FS	0.919	0.831	0.758	0.906	2.022
Matjhabeng Local Municipality	FS	0.919	0.765	0.641	0.906	1.827
Setsoto Local Municipality	FS	0.451	0.404	0.363	0.444	2.045
Moqhaka Local Municipality	FS	1.000	0.819	0.712	0.986	2.075
Ngwathe Local Municipality	FS	0.778	0.667	0.577	0.768	1.805
Metsimaholo Local Municipality	FS	0.506	0.429	0.370	0.498	1.900
Mafube Local Municipality	FS	0.594	0.506	0.433	0.587	1.626
Emfuleni Local Municipality	GT	1.000	0.767	0.599	0.984	1.616
Midvaal Local Municipality	GT	0.446	0.379	0.321	0.439	1.743
Lesedi Local Municipality	GT	0.374	0.328	0.292	0.368	2.051
Nokeng Tsa Taemane Local Municipality	GT	0.931	0.811	0.724	0.916	2.117
Kungwini Local Municipality	GT	1.000	0.769	0.622	0.984	1.661
Mogale City Local Municipality	GT	0.602	0.503	0.422	0.592	1.758
Randfontein Local Municipality	GT	0.757	0.651	0.571	0.745	1.918

Municipality		E^{VRS}	E_{bc}	$CL^{bc}_{0.025}$	$CL^{bc}_{0.097}$	$\lvert bias \rvert / \hat{\sigma}_b$
Westonaria Local Municipality	GT	0.617	0.549	0.485	0.607	1.899
Merafong City Local Municipality	GT	0.795	0.673	0.571	0.782	1.765
Umhlathuze Local Municipality	KZN	0.770	0.649	0.528	0.759	1.542
Mogalakwena Local Municipality	LIM	0.285	0.241	0.203	0.281	1.784
Msukaligwa Local Municipality	MP	0.424	0.373	0.335	0.418	2.033
Govan Mbeki Local Municipality	MP	0.642	0.544	0.465	0.632	1.881
Steve Tshwete Local Municipality	MP	0.557	0.471	0.406	0.548	1.879
Dr JS Moroka Local Municipality	MP	0.988	0.851	0.724	0.973	1.775
Mbombela Local Municipality	MP	0.916	0.794	0.684	0.902	1.869
Rustenburg Local Municipality	NW	1.000	0.768	0.621	0.984	1.691
Richtersveld Local Municipality	NC	0.500	0.430	0.349	0.496	1.432
Nama Khoi Local Municipality	NC	0.388	0.342	0.306	0.382	2.112
Hantam Local Municipality	NC	0.925	0.793	0.688	0.912	1.836
Karoo Hoogland Local Municipality	NC	0.840	0.701	0.575	0.827	1.623
Khai-Ma Local Municipality	NC	0.529	0.448	0.369	0.522	1.561
Cederberg Local Municipality	WC	0.596	0.540	0.495	0.586	2.080
Drakenstein Local Municipality	WC	0.832	0.711	0.618	0.818	1.989

The bias corrected efficiency estimates provide similar results to those discussed above. The Eastern Cape still has the highest average efficiency estimates and the local municipality with the lowest efficiency estimate is still Mogalakwena local municipality in Limpopo (0.2410). The difference in results comes in the assessment of the most efficient local municipalities. The three most efficient local municipalities according to their bias corrected estimates are Dr JS Moroka local municipality in Mpumalanga (0.851), Letsemeng local municipality in the Free State (0.847) and Tswelopele local municipality in the Free State (0.831).

The results of the assessment shown in Table 6.2 could be improved by the inclusion of more municipalities from Limpopo, North West and KwaZulu-Natal, all of which have only one municipality included in this assessment.

6.4.3 Local and Metropolitan Municipalities

Combining the local and metropolitan municipalities in an efficiency evaluation is not advised as these municipalities operate in vastly different environments (Steering Committee for the Review of Commonwealth/State Service Provision, 1997). Acknowledging this difference, the results of the combined evaluation are given in Table 6.3. The results indicate that the metropolitan municipalities typically outperform the local municipalities. The average efficiency for metropolitan municipalities is 0.986, while the average of the local municipalities is 0.717. This underlines the fact that operating environments do affect the efficiency evaluations of municipal water providers. Metropolitan municipalities will typically serve larger populations and have better infrastructure than local municipalities which provide water services to citizens in small towns and rural villages. These operating differences serve as an explanation for differences in measured relative efficiency, but may be exogenous rather than endogenous (subject to management control and variation).

As was the case for the local municipalities, the Eastern Cape had the highest average efficiency (0.918). This result comes from the fully efficient Camdeboo, Ndlambe and Baviaans local municipalities. The Buffalo City metropolitan municipality achieved an efficiency rating of 0.999; an interesting result as this municipality only achieved a 0.481 efficiency value when assessed previously (grouped with district and metropolitan municipalities). This result provides further evidence of the divide between the metropolitan and local municipalities. The province which achieved the lowest average efficiency was Mpumalanga (0.684). Limpopo Province and North West Province each had a single municipality included in the evaluation which excluded them from inter-provincial comparisons.

The best performing municipalities according to the bias-corrected estimates are the Buffalo City metropolitan municipality (0.869), Letsemeng local municipality (0.855) and Tswelopele Local Municipality (0.840). These results are similar to those found in the evaluation of local municipalities. The bias corrected efficiencies presented in Table 6.3 show that the fully efficient municipalities are more affected by bias corrections than those municipalities which are not fully efficient. The average bias correction made for fully efficient municipalities is 0.202, while the average bias correction for the municipalities which are not fully efficient is 0.082.

These results indicate that the operating environment influences the efficiency estimates for water service authorities in South Africa. The metropolitan municipalities tend to operate more efficiently than the local municipalities.

Table 6.3: Efficiencies of local and metropolitan municipalities using DEA, 95% bias corrected confidence intervals are provided for each efficiency estimate.

Municipality		E^{VRS}	E_{bc}	$CL^{bc}_{0.025}$	$CL^{bc}_{0.097}$	$\lvert bias \rvert / \hat{\sigma}_b$
Nelson Mandela Bay Metro Municipality	EC	0.901	0.784	0.677	0.891	1.768
Camdeboo Local Municipality	EC	1.000	0.792	0.658	0.987	1.722
Makana Local Municipality	EC	0.613	0.549	0.492	0.606	1.991
Ndlambe Local Municipality	EC	1.000	0.820	0.709	0.988	1.900
Baviaans Local Municipality	EC	1.000	0.788	0.646	0.988	1.608
Buffalo City Metro Municipality	EC	0.999	0.869	0.754	0.986	1.859
Letsemeng Local Municipality	FS	0.925	0.855	0.785	0.916	1.824
Mangaung Metro Municipality	FS	1.000	0.809	0.687	0.987	1.836
Mantsopa Local Municipality	FS	0.556	0.503	0.454	0.550	1.903
Tswelopele Local Municipality	FS	0.919	0.840	0.770	0.908	1.963
Matjhabeng Local Municipality	FS	0.861	0.734	0.627	0.850	1.760
Setsoto Local Municipality	FS	0.451	0.409	0.372	0.445	1.992
Moqhaka Local Municipality	FS	1.000	0.832	0.732	0.986	2.049
Ngwathe Local Municipality	FS	0.778	0.676	0.590	0.768	1.760
Metsimaholo Local Municipality	FS	0.506	0.436	0.380	0.500	1.868
Mafube Local Municipality	FS	0.594	0.514	0.438	0.588	1.521

Municipality		E^{VRS}	E_{bc}	$CL^{bc}_{0.025}$	$CL^{bc}_{0.097}$	$^{\|bias\|}/\,\hat{\sigma}_b$
City Of Johannesburg Metro Municipality	GT	1.000	0.777	0.600	0.986	1.516
City Of Tshwane Metro Municipality	GT	1.000	0.798	0.647	0.987	1.661
Ekurhuleni Metro Municipality	GT	1.000	0.779	0.618	0.986	1.554
Emfuleni Local Municipality	GT	1.000	0.840	0.701	0.986	1.759
Midvaal Local Municipality	GT	0.446	0.385	0.328	0.440	1.691
Lesedi Local Municipality	GT	0.374	0.333	0.298	0.369	1.987
Nokeng Tsa Taemane Local Municipality	GT	0.931	0.822	0.736	0.919	2.017
Kungwini Local Municipality	GT	1.000	0.782	0.638	0.987	1.590
Mogale City Local Municipality	GT	0.578	0.492	0.422	0.571	1.770
Randfontein Local Municipality	GT	0.757	0.660	0.582	0.747	1.820
Westonaria Local Municipality	GT	0.617	0.556	0.497	0.609	1.869
Merafong City Local Municipality	GT	0.795	0.685	0.589	0.784	1.723
eThekwini Metro Municipality	KZN	1.000	0.777	0.599	0.987	1.511
Umhlathuze Local Municipality	KZN	0.512	0.437	0.374	0.505	1.820
Mogalakwena Local Municipality	LIM	0.285	0.245	0.208	0.282	1.719
Msukaligwa Local Municipality	MP	0.424	0.377	0.340	0.420	1.934
Govan Mbeki Local Municipality	MP	0.628	0.548	0.490	0.619	2.113
Steve Tshwete Local Municipality	MP	0.557	0.479	0.417	0.550	1.837
Dr JS Moroka Local Municipality	MP	0.927	0.801	0.695	0.914	1.915
Mbombela Local Municipality	MP	0.885	0.786	0.708	0.872	2.057
Rustenburg Local Municipality	NW	1.000	0.784	0.643	0.986	1.608

Municipality		E^{VRS}	E_{bc}	$CL^{bc}_{0.025}$	$CL^{bc}_{0.097}$	$^{\lvert bias \rvert}/\hat{\sigma}_b$
Richtersveld Local Municipality	NC	0.500	0.437	0.354	0.497	1.365
Nama Khoi Local Municipality	NC	0.388	0.346	0.311	0.384	2.040
Hantam Local Municipality	NC	0.925	0.805	0.697	0.913	1.733
Karoo Hoogland Local Municipality	NC	0.840	0.713	0.583	0.829	1.524
Khai-Ma Local Municipality	NC	0.529	0.455	0.374	0.523	1.488
Cederberg Local Municipality	WC	0.596	0.546	0.503	0.588	2.006
Drakenstein Local Municipality	WC	0.825	0.720	0.640	0.813	2.033

6.5 Ranking Efficient Municipalities and Outlier Detection

A method for ranking the efficient municipalities given in section 6.4 was developed by Andersen and Petersen (1993). The procedure described by Andersen and Petersen (1993) allows for an investigator to differentiate between the municipalities which form the frontier (have and efficiency of 1). The method is termed "super-efficiency" and is determined for each efficient municipality.

The super-efficiency of an efficient municipality is determined by excluding the municipality from the DEA evaluation, and re-forming the frontier in the absence of the extracted efficient unit. This will obviously effect the change the frontier. The excluded municipality is then re-introduced and its efficiency estimate is calculated against the new frontier (formed in its absence). The efficiency estimate of this municipality is calculated in the same manner as all other municipalities. When the efficiency is calculated in this way for the excluded municipality, the efficiency estimates will now be larger than 1. This is a consequence of the fact that this municipality will now fall beyond the frontier which was constructed in its absence. Efficiencies calculated in this manner are termed "super-efficiencies". Bogetoft and Otto (2010) indicate that the drawback of using the super efficiency measure can result in infeasible solutions to the linear program. The super efficiencies are defined as ∞ in these cases. These "infinite" super efficiencies are a result of there being no firms against which the municipality can be gauged. These firms are called hyper efficient.

The super-efficiency values for a set of municipalities can also be used as a simplistic method of determining outliers in a DEA evaluation. Outliers can have an enormous effect on the results of a DEA evaluation and must be removed if at all possible. Bogetoft and Otto (2010) argue that a super-efficiency value of 3 to 4 can indicate the presence of an outlier. The theoretical underpinnings of this method are limited (Bogetoft and Otto, 2010). However, it provides a simple method for identification of outliers. Hyper efficient firms are disregarded in this type of outlier detection as there is insufficient information available to establish whether these are outliers.

The ranking of the efficient municipalities for the district and metropolitan municipality evaluation is given in Table 6.4.

Table 6.4: Super-efficiencies for the district and metropolitan municipalities.

Municipality		**Super-efficiency**
City Of Johannesburg Metro Municipality	GT	∞
eThekwini Metro Municipality	KN	∞
Central Karoo District Municipality	WC	18.562
Capricorn District Municipality	LIM	9.841
Umgungundhlovu District Municipality	KZN	5.609
Ekurhuleni Metro Municipality	GT	1.969
West Coast District Municipality	WC	1.762
O.R. Tambo District Municipality	EC	1.616
City Of Tshwane Metro Municipality	GT	1.341

The results in Table 6.4 indicate that there are two hyper-efficient metropolitan municipalities, Johannesburg and eThekwini. Also there are three district municipalities which are potential outliers: Central Karoo, Capricorn and Umgungundhlovu. These municipalities have super-efficiencies of 18.562, 9.841 and 5.609 respectively, all of which exceed the 3 to 4 threshold for potential outliers. These extreme values indicate that there is a high likelihood that these municipalities are outliers; a consequence of extremely high efficiency or of poor record keeping and measurement. For this reason, the data for these municipalities needs to be audited to ensure their accuracy, before they can be included in the analysis.

The remaining municipalities have typical super-efficiency values, and are ranked according to this value. The results are that, apart from the two hyper-efficient municipalities and the three potential outliers, the Ekurhuleni metropolitan municipality performs best followed by the West Coast and O.R. Tambo district municipalities. The city of Tshwane metropolitan municipality has the lowest rated super-efficiency value. These rankings can also be used to assess the municipalities in a benchmarking initiative.

The super-efficiencies of the efficient municipalities for the local municipality evaluation are given in Table 6.5.

Table 6.5: Super-efficiencies for the local municipalities.

Municipality		**Super-efficiency**
Emfuleni Local Municipality	GT	∞
Kungwini Local Municipality	GT	2.436
Baviaans Local Municipality	EC	1.977
Rustenburg Local Municipality	NW	1.863
Ndlambe Local Municipality	EC	1.509
Camdeboo Local Municipality	EC	1.503
Moqhaka Local Municipality	FS	1.301

The super-efficiency values for the efficient local municipalities indicate that there is a single hyper-efficient municipality, namely Emfuleni local municipality. The values of the super-efficiencies for this selection of local municipalities indicate that there are no obvious outliers (all super-efficiencies are less than 3). Furthermore, a ranking can now be placed on the efficient municipalities. Regarding Emfuleni as hyper-efficient, and ranking of the remaining efficient local municipalities according to their super-efficiency values, results in Kungwini local municipality being ranked best, followed by Baviaans, Rustenburg, Ndlambe, Camdeboo and Moqhaka.

6.6 Conclusion

Better data collection and reporting systems will permit more municipalities to be captured in the efficiency assessments.

The results of the analysis of data currently available on South Africa's municipal water service providers indicate that the efficiency evaluation for the local municipalities is more ''well behaved" than that of the district and metropolitan municipalities. This superior behavior of local municipalities is evidenced (in the analysis of the local municipalities) by the lack of obvious outliers in the data, a smaller proportion of efficient municipalities and more reasonable super-efficiency estimates, but this result could also due to of a lack of sufficient data at a district and metropolitan municipal level. The number of municipalities included in the evaluation of the district and metropolitan municipalities was 15 which exceeds minimum requirement as proposed by Dyson *et al*. (2001), which is 12 = (2 × 2 × 3), but only just.

The high proportion of efficient municipalities and potential outliers indicates that the sample size may be too small. The fact that district municipalities comprise of collections of local municipalities, renders efficiency evaluations between the district and local municipalities meaningless. The data reported for district municipalities is simply an aggregate of the data reported for the constituent local municipalities. The reality therefore is that the operating environments of district and metropolitan municipalities is vastly different; with the result that they fall into different "efficiency camps".

A possible solution to this would be to classify metropolitan municipalities and local municipalities which contain large towns as "urban" municipalities, and the remaining local municipalities as "rural" and conduct an efficiency evaluation on these two distinct groups. This may solve the problem of insufficient data, as well as providing a method to compare municipalities which have similar operating environments.

The super-efficiency (non-parametric estimation based) ratings were preferred on the grounds that their calculation was more closely linked to the principles of a frontier than those based on estimating parameters of pre-specified models, and that their validity was able to be tested through applying a bias correction procedure and recalculating the DEA indices in a way that both reveals outliers and distinguishes equivalently (maximally) efficient water service providers.

CHAPTER 7

APPLYING EFFICIENCY MONITORING – A POLICY PERSPECTIVE AND A CONCLUDING NOTE

7.1 Introduction

As the South African tax and tariff paying population matures in its new and evolving democratic setting, it can be expected that they will become more conscious about getting value for their tax and tariff payments – much like the current citizens of most developed countries. Obtaining value for money is a multi-faceted goal. One facet of it is requiring high levels of efficiency performance from the provider of the service. Another is spending the money on the most highly valued attributes of the service – a feature covered in other WRC Research reports.

In the water service sector, customers of South African municipalities are served by Blue Drop and Green Drop monitoring and rating assessments of their the municipal potable and waste water service. These performance assessments indicate where the capacity to deliver, or the record of delivery, is inadequate. The reference for adequacy is a set of national benchmarks and standards. This information serves to indicate how well particular municipalities are shaping up against the national standards (or targets) for water service delivery. Failures to attain the required performance standard may occur for many reasons, including shortfalls in municipal infrastructure, management and technical personnel, and it is good that South Africa has a centralised monitoring system that can expose these problems.

Through the annual Blue Drop and Green Drop rating and monitoring reports the public has been made aware of many inadequacies in water service delivery. In many cases this monitoring and reporting may well be all that is needed to incentivise municipalities to maintain or improve the standard of water service they supply, but experience in the public service more generally, suggests monitoring and reporting is not sufficient incentive to bring about improvement in all cases (Curristine, Lonti and Joumard, 2007), and other additional incentives may be required to sustain the improvement.

From a monitoring the performance perspective, the meeting national guidelines and complying with Water Law is not all that can be done with respect to assessing water service delivery performance. Useful as the Blue Drop and Green Drop performance monitoring services are, they only present a partial picture of how the various municipalities are performing. There are also other potential performance failures or successes in which the public are interested and which influence social welfare. One such performance area is the efficiency in use of the municipal water service delivery infrastructure and resources[2]. A municipality may meet the required national standard of delivery, but in doing so it may be using more (budget) resources than would be efficient. To assess the efficiency of use of (budget) resources, another type of performance assessment is needed – an efficiency performance assessment.

The rationale for, or contribution of, efficiency and other performance assessments can assist the public sector contribution to society in a number of ways, for instance, by:

- creating more precise awareness about results of service delivery within government
- providing more and improved information on government aims and how government policy contributes to these aims
- signalling to key decision makers which parts of the service are working well and which not
- improving the information available to parliament and the public on public delivery management and efficiency performance (Curristine, Lonti and Joumard, 2007).

At the operational level, regular performance monitoring and assessment has the potential to identify persistent inefficiency in service delivery. Once identified, this monitoring information can be used to induce self-managed and motivated improvement by those identified as inefficient, attract regulatory support to improve the inefficient parts of the service and to identify role models and mentors (those who sustain high levels of efficiency) to assist less efficient performers. All other things being equal, a low relative efficiency rating would indicate that the water service output being provided by a municipality is inferior to what is being provided by other municipalities, adjusted for the inputs available to them and output delivery challenge faced by them.

[2] Yet another important assessment of municipal water service delivery performance remaining to be done is that of demand (customer) satisfaction. In addition to the supply side assessments (Blue Drop, Green Drop, low water loss waste and efficiency ratings) there is a strong case for a demand side assessment to be incorporated into performance assessment.

Critical to the attractiveness of employing efficiency benchmarking is that the indices do what they claim, viz., indicate relative efficiency. If they do not, the generation of such indices may induce to the very opposite of what they are supposed to – they may misdirect regulatory management effort or induce misallocations of resources. Chapter 7 will review the principles underlying the calculation of relative efficiency indices, identify problems that may undermine the rationale for using these indices to guide regulatory management and propose some guidelines for addressing these problems.

7.2 Policy instruments appropriate to addressing identified performance inadequacy (like inefficiency)

International experience suggests there are many instruments and policies by which to initiate effort to improve efficiency in public sector service delivery (Curristine, Lonti and Joumard, 2007). Efficiency in the international context is defined with reference to cost per unit of service (output). These policies include the devolution of responsibility to the local level, decentralisation of management and responsibility, introducing competitive pressure in supply (not necessarily through privatisation), changing the composition, size and human resource management of the workforce, and the bringing of management and budget practices and procedures more closely into line with the achievement of targeted results and objectives (Curristine, Lonti and Joumard, 2007). Experience within the OECD countries indicates there to be no (single) blueprint for improving efficiency in the water services sector or other sections of the public service (Curristine, Lonti and Joumard, 2007). For this reason, OECD countries have adopted a wide range of instruments to assist in the policy of improving efficiency.

Despite the wide range of efficiency improving initiatives and instruments employed across the many services of the public sector, there is little evidence so far accumulated that particular ones consistently work. The lack of evidence is due to there being insufficient budget being allocated to investigate the efficiency impacts on particular policy instruments, the changing of many external factors as the same time as the instruments are applied, and the absence of pre-change measures of performance against which to compare post-change performance (Curristine, Lonti and Joumard, 2007).

What OECD country experience does reflect is that within the government bureaucracy tasked with service delivery, great care needs to be exercised in designing the signals and consequences linked to performance information reporting. The signalling can be both positive (bonuses and promotion incentives) and negative (punishments like budget and staff cutting) and induce positive or (unintentional) negative behavioural change. The rational for introducing performance measuring and monitoring is based on the inducement to positive change only.

Within the OECD countries the major vehicle for linking performance benchmarking information to efficiency performance improvement has been through budget processes (Curristine, Lonti and Joumard, 2007). The strategic objective pursued through this linkage was to shift the focus of public sector service decision making away from bureaucratic effort to win as much money as possible from the Treasury for the budget, to increasing what can be achieved with the money that has been allocated (Curristine, Lonti and Joumard, 2007). However, the linkage of performance benchmarking to punishments, such as budget cuts, has induced some negativity toward benchmarking in some countries (Curristine, Lonti and Joumard, 2007).

To sum up, there are costs incurred to developing and maintaining systems for collecting and reporting the performance of water service providers, and these costs need to be weighed up against the unproven, but likely, benefits of implementing such systems.

7.3 Reviewing what the efficiency performance measures (indices) are supposed to indicate and achieve

7.3.1 Technical versus allocation efficiency

For the purpose of water services provision two types of efficiency indices can be identified – technical (production) and allocation (input price related). The sum of these two types of efficiency is termed overall or cost efficiency. The difference between the two types is described in Chapter 2 of this report, but because it is important, the difference is further expanded upon in this Chapter – with reference to the model of water service production in an input space represented in Figure 7.1.

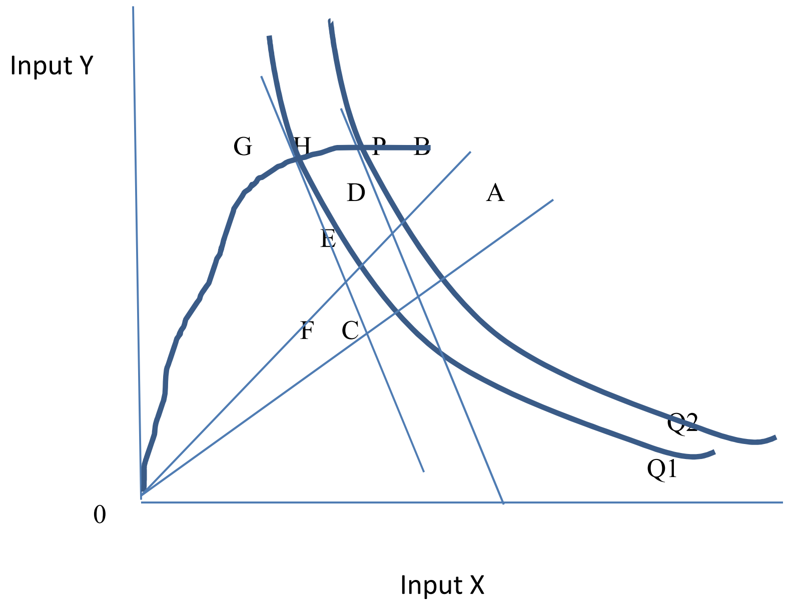

Figure 7.1: Cost or overall efficiency

In Figure 7.1 a production function in inputs X and Y is described by a set of isoquants (for Q1 and Q2). The inputs employed and outputs produced for two municipalities are shown by points A and B and the relative prices of inputs (per unit costs) are reflected by the slope of the total cost lines GF and HE. These prices are assumed to remain constant as output increases (the municipalities are price takers), so this slope remains constant as output and total cost expenditure increases. The minimum cost of inputs in order to produce Q1 output is defined by the point of tangency of the total cost line and the isoquant Q1, point G along the GF line. The minimum cost of inputs in order to produce Q2 output is similarly defined by point of tangency between the total cost line and isoquant Q2, point H along the HE line.

Given the price ratios of the inputs described by the cost lines and the production function described by the isoquants, an overall or cost efficiency can be defined – a situation where the employment combination of inputs that minimise costs and those inputs are utilised in a way that maximises production of the desired (targeted) output. This efficiency optimum is described by the locus of points along the production ray OGH in Figure 7.1.

The concept of relative efficiency is a measure that compares actual input hiring and deployment production choices with options that lie on this production ray (OGH). If the actual input employment choice of the two municipalities in producing Q1 and Q2 outputs lie off the production ray (at points such as at A and B), the total costs of these municipalities will be higher than would be efficient (as reflected by points on the production ray OGH). This inefficiency is in part due to the total cost lines going through points A and B being positioned to the right of those total cost lines going through points G and H. The measures of (in)efficiency of municipalities A and B respectively are F0/A0 and E0/B0. The closer these measures are to unity, the greater the efficiency of municipality, that is, the more successful it is in minimising production cost.

Cost efficiency as described in Figure 7.1 is a combination of two types of efficiency – incorrect technological choice and over employment of inputs. The over employment of inputs inefficiency is the excess input hired over the minimum required, as represented by points along the isoquants, and is labelled by the literature as technical inefficiency. For municipality A, the relevant measure of technical efficiency is 0C/0A and for municipality B it is 0D/0B. Both municipalities A and B are technically inefficient because their efficiency indices are less than one. Of the two municipalities, B is more technically efficient than A because its index is higher.

7.3.2 Both technical and allocation efficiencies are relevant

Both technical and input price related (also called allocation) efficiency are relevant. Technical efficiency is relevant when the choice of technology has been made by which to produce water services, and guidance is sought on how to minimise the inputs used in order to produce a target output, or maximise an output from a given set of inputs. A high technical efficiency indicates that, given the technology being employed by the municipality, the municipality is not wasting inputs by employing an unnecessary excess of them in production. It follows that technical inefficiency implies the overemployment of inputs required to do the job, given the technology chosen.

A complicating factor with interpreting technical efficiency is that high cost efficiency indices may also be consistent with a municipality that over-reaches its water service infrastructure. In this situation the municipality's apparent impressive current technical efficiency may be coincident with impending future technical efficiency failure. Given the long lead time required to correct for water infrastructure shortfalls, sustainable technical efficiency may require that investment in water service infrastructure maintains a certain level of spare or excess capacity, but excess capacity can reflect as inefficiency when the latter is estimated.

It follows that the analysis of efficiency performance measures needs to take into consideration the sufficiency of the water service infrastructure to meet planned future needs. The interpretation of the efficiency index should be done against the background of whether the municipality has the infrastructure capacity needed for the planned future to provide the water service it is constitutionally mandated to deliver. In order to accommodate the excess planning requirement for infrastructure inputs into production, a time horizon has to be set for analysing efficiency performance measures that encompass planned future delivery obligations and responsibilities.

Allocation (input price related) efficiency is defined with respect to the choice of technology. Some technologies require greater proportions of some inputs in the total that are hired. For instance, some technologies may require more capital to be hired relative to labour. In terms of the cost of inputs and the production function, some technologies are more efficient than others. Allocation inefficiency occurs when a municipality chooses a technology (and input proportion) that does not minimise the total cost of producing a given level of output (because it does not take into account the relative unit costs of the inputs). In the model shown in Figure 7.1, both municipalities are allocation inefficient – they employ too much of input X relative to input Y in production. The relevant measures of allocation efficiency are F0/C0 for Municipality A and E0/D0 for Municipality B. Choosing the right technology is an important efficiency cost minimising concern. If a municipality is allocation inefficient, the problem that must be addressed is changing the technology (the proportions of inputs hired relative to each other) or revisiting the unit costs paid for the various inputs employed.

7.3.3 Which is the preferred statistical efficiency frontier to benchmark by – SFA or DEA?

Technical and allocation efficiency are combined in the relative efficiency indices generated through the estimation of efficiency frontiers. The locus of cost efficient input combinations to produce water service (points along the production ray OGH in Figure 7.1) can be transformed into a production efficiency frontier or a total cost efficiency frontier. As a production efficiency frontier it describes the maximum output that can be produced with different expenditures on inputs, given the relative costs of the inputs used in production. As a cost efficiency frontier it describes the minimum total expenditure on inputs that needs to be allocated in order to produce different levels of output.

The SFA and OLS estimation methodologies employed in this study predict the costs efficient combinations as a cost efficiency frontier, while the DEA estimation methodology employed in this study predict the cost efficient combinations as a production efficiency frontier.

Assuming a Cobb-Douglas form, the SFA was applied to estimate the aggregate total cost function (sum of potable and waste water management operational costs) in the inputs total system volume input for the municipality, the length of the mains for the municipality, the average unit cost of labour for the municipality and the average unit cost of other operating expenditure (repairs, etc.) for the municipality. Applying a linear programming methodology, the DEA was applied to estimate the maximum output that could be achieved in the form of system input volume (in kl), length of mains (in km) and number of connections, given the inputs total operating expenditure and total number of employees. The parametric form of the production frontier in the COLS evaluation was predetermined and not as flexible as that of the DEA approach. The COLS evaluation incorporated input prices, which took into account the cost of employees, a factor which was not included in the DEA evaluations.

The alternative statistical methodologies by which the combination of technical and allocation efficiency may be estimated (predicted) will frequently yield inconsistent ratings (Cubbin and Tzanidakis 1998). Such inconsistency was also found in this study. In order to assess how the results of each efficiency evaluation compare with each other the rankings were compared. In this comparison the ranks of the DEA results were established using the ranking of the super-efficiency values discussed in Chapter 6, and the COLS based ranking discussed in Chapter 5. The rankings were arranged from largest to smallest with a ranking of 1 indicating the most efficient municipality according to the evaluation method used. A municipality which was classified as hyper-efficient in the DEA evaluation was automatically ranked as 1. The comparison between the ranks of the evaluations of the district and metropolitan municipalities is provided in Table 7.1. The two municipalities which did not have sufficient data to be included in the COLS evaluation were omitted from this comparison. It is clear from Table 7.1 that the rankings of the district and metropolitan municipalities differ substantially for each of the evaluation methods. The Spearman's rank correlation was calculated to be 0.3202 indicating a weak correlation between the ranks of the two evaluation methods.

Table 7.1: Comparison between efficiency estimates using DEA and COLS for district and metropolitan municipalities.

Municipality		**DEA**		**COLS**	
		$E_{\sup er}^{VRS}$	**Rank**	*eff*	**Rank**
Nelson Mandela Bay Metro Municipality	EC	0.5692	11	0.5281	10
Buffalo City Metro Municipality	EC	0.4814	12	0.3759	13
O.R. Tambo District Municipality	EC	1.6165	6	0.5469	9
Mangaung Metro Municipality	FS	0.6997	9	0.6657	5
City Of Johannesburg Metro Municipality	GT	∞	1	0.798	3
City Of Tshwane Metro Municipality	GT	1.3413	7	0.6219	7
Ekurhuleni Metro Municipality	GT	1.9693	4	0.6326	6
eThekwini Metro Municipality	KZN	∞	1	0.6094	8
Ugu District Municipality	KZN	0.6424	10	1.000	1
Uthukela District Municipality	KZN	0.3687	13	0.5258	11
Uthungulu District Municipality	KZN	0.8342	8	0.9172	2
West Coast District Municipality	WC	1.7623	5	0.5086	12
Central Karoo District Municipality	WC	18.5617	3	0.6777	4

The comparison reveals some similarities in the efficiency values and ranks. The rankings and efficiency estimates of the Nelson Mandela Bay and Buffalo City metropolitan municipalities were very similar in both evaluations. The same could be said for the efficiency estimate of the Mangaung metropolitan municipality – very similar in both evaluations.

The comparison between the efficiency estimates of the local municipalities using DEA and SFA is given in Table 7.2. As it was for the district and metropolitan municipalities, the comparison of efficiency estimates and rankings for the local municipalities shows that the two evaluation methods provide considerably different results. The Spearman's rank correlation between the COLS rankings and the DEA rankings is 0.1716, indicating a very weak correlation between the ranks. The lower Spearman's rank correlation statistic indicates that there is less consistency between the rankings for the local municipalities than there is for the district and metropolitan municipalities. There were a few municipalities which had similar efficiency estimates for both evaluation procedures. The Makana (Eastern Cape), Mafube (Free State), Mogale (Gauteng), Merafong (Gauteng), Steve Tshwete (Mpumalanga), Hantam (Northern Cape) and Karoo Hoogland (Northern Cape) all had similar efficiency estimates (within 0.03). The rankings between these municipalities, however, were often significantly different.

The analysis of the rankings reveals significant differences between the efficiency ranking systems. For example, in the SFA evaluation the Khai-Ma local municipality (Northern Cape) achieved the highest ranking and was therefore classified as the most efficient local municipality, while in the DEA evaluation the same municipality was classified as the 29th most efficient (out of 37 considered). Similar results are observed for municipalities which achieve high efficiency rankings in the DEA evaluation. For example, in the DEA evaluation, the Dr JS Moroka local municipality (Mpumalanga) achieved a ranking of 8th (efficiency estimate of 0.9877), while at the same time was ranked as second worst (efficiency estimate of 0.2025) in the SFA evaluation.

Table 7.2: Comparison between efficiency estimates using DEA and SFA for local municipalities.

Municipality		DEA		SFA	
		$E^{VRS}_{\sup er}$	Rank	*eff*	Rank
Camdeboo Local Municipality	EC	1.5026	6	0.9995	3
Makana Local Municipality	EC	0.6132	23	0.5832	30
Ndlambe Local Municipality	EC	1.5092	5	0.6058	27
Baviaans Local Municipality	EC	1.9774	3	0.6976	24
Letsemeng Local Municipality	FS	0.9252	10	0.7592	20
Mantsopa Local Municipality	FS	0.5555	28	0.4532	33
Tswelopele Local Municipality	FS	0.9187	13	0.7866	17
Matjhabeng Local Municipality	FS	0.9193	12	0.7829	18
Setsoto Local Municipality	FS	0.4508	32	0.6810	25
Moqhaka Local Municipality	FS	1.3010	7	0.7325	22
Ngwathe Local Municipality	FS	0.7780	18	0.7221	23
Metsimaholo Local Municipality	FS	0.5064	30	0.9280	5
Mafube Local Municipality	FS	0.5942	26	0.5839	29
Emfuleni Local Municipality	GT	∞	1	0.7539	21
Midvaal Local Municipality	GT	0.4457	33	0.9709	4
Lesedi Local Municipality	GT	0.3740	36	0.5717	31
Nokeng Tsa Taemane Local Municipality	GT	0.9307	9	0.8506	8
Kungwini Local Municipality	GT	2.4363	2	0.9995	2
Mogale City Local Municipality	GT	0.6020	24	0.6031	28
Randfontein Local Municipality	GT	0.7567	20	0.8264	13

Municipality		DEA		SFA	
		$E_{\sup er}^{VRS}$	**Rank**	***eff***	**Rank**
Westonaria Local Municipality	GT	0.6166	22	0.8284	12
Merafong City Local Municipality	GT	0.7950	17	0.7799	19
Umhlathuze Local Municipality	KZN	0.7698	19	0.4332	34
Mogalakwena Local Municipality	LIM	0.2850	37	0.4053	35
Msukaligwa Local Municipality	MP	0.4242	34	0.1648	37
Govan Mbeki Local Municipality	MP	0.6416	21	0.8027	15
Steve Tshwete Local Municipality	MP	0.5571	27	0.5651	32
Dr JS Moroka Local Municipality	MP	0.9877	8	0.2025	36
Mbombela Local Municipality	MP	0.9155	14	0.6141	26
Rustenburg Local Municipality	NW	1.8626	4	0.8484	9
Richtersveld Local Municipality	NC	0.5000	31	0.8867	7
Nama Khoi Local Municipality	NC	0.3881	35	0.8399	10
Hantam Local Municipality	NC	0.9245	11	0.9131	6
Karoo Hoogland Local Municipality	NC	0.8399	15	0.8124	14
Khai-Ma Local Municipality	NC	0.5292	29	0.9995	1
Cederberg Local Municipality	WC	0.5959	25	0.8338	11
Drakenstein Local Municipality	WC	0.8319	16	0.7889	16

The differences between the results of these evaluations are concerning, but not unexpected. A similar study of efficiency in water service delivery conducted by Cubbin and Tzanidakis (1998) compared the results of DEA and a regression analysis approach similar to the COLS and SFA approaches used in this study. The results of Cubbin and Tzanidakis (1998) were similar to those obtained here – large differences between the rankings of the two approaches. The two evaluation methods do not reflect the same efficiencies.

The most obvious explanation for the differences is that the two methods incorporated different information in their calculation – the parametrically derived estimates incorporated input pricing whereas the non-parametrically derived estimates did not.

One of the main objectives of this study was to reflect on the methodology by which to calculate relative efficiency measures. In terms of methodological appeal this study prefers and recommends the adoption of the DEA method for rating technical efficiency. It does so because Data Envelopment Analysis (DEA) more closely relates to the primal production optimisation problem being modelled (in Figure 7.1), viz. maximising output with a given level of inputs (but see Cubbin and Tzanidakis 1998).

7.3.4 The effect of including additional municipalities

Given that the validity of the methodologies can be improved by the addition of variables, it is important to also consider what would be the likely consequences of adding variables. This study argues that the inclusion of additional municipalities into the models would be unlikely to have a major effect on the efficiency estimates. If new data is included into the SFA model, this is likely to change the estimation of the frontier and, as a consequence the efficiency estimates. If the included observation is not an outlier, it is only likely that it will have a very small effect on the model and the efficiency estimates. In the DEA model, as long as the new municipality does not form part of the frontier (i.e. does not have an efficiency estimate of 1) the inclusion of an additional municipality will have no effect on the construction of the frontier and thus will have no effect on the efficiency estimates of the original municipalities. If the new municipality does redefine the frontier (have an efficiency estimate of 1) this can have a big effect on the original efficiency estimates. This impact must be considered when introducing new elements (municipalities) into the evaluation.

7.4 Factors that may cause efficiency indices to mislead policy

There is considerable scope for efficiency indices to mislead policy because the phenomenon they claim to 'explain' as being due to variations in efficiency is due to other factors. Three such other factors are: input data inadequacy, temporary change and differing externally determined potentials (explanatory circumstances or factors on production and allocation choice not captured in the economic models applied). The problem of data adequacy has many facets, including lack of information on key inputs and too small a sample size. The problem of temporary change is that the statistic calculated may simply reflect a short-term temporary position and not underlying inefficiency or efficiency. The problem of differing potentials is that it may be more sensible to compare like with like (oranges with oranges) in order to draw reliable (and sensible) deductions about efficiency.

In order to guard against erroneously drawing conclusions about persistent (chronic) inefficiency that are really reflecting a temporary supply phenomenon, there needs to be a degree of persistence about the inefficiency found before it should trigger regulatory intervention effort aimed at correcting the inefficiency.

The usefulness (for regulation) of water service relative efficiency assessments are shaped by the way these problem factors are addressed.

7.5 Five guidelines for addressing the factors that may cause efficiency performance assessment to fail in their objective

In order to address factors that cause efficiency performance monitoring to fail, five guidelines are proposed for applying efficiency indices to inform regulatory management of South Africa's municipal water service provision.

Guideline One – in order to assess efficiency there must be sufficient data available, and the data must be generated through an adequate system of water service accounting, reported annually, and be subject to an independent validation process (audit).

<u>(a) An adequate accounting system by which to gathering the information</u>

In order to calculate performance indices an adequate accounting system must be in operation. With respect to the inputs, an accrual accounting system is required that records costs of the water service (as distinct from other services provided) rather than cash flows or general costs incurred (for multiple services). With respect to the outputs, the measurement of water service provided should capture all the outputs, including both the important quantitative and qualitative parts of the services provided. Some of the output may be very difficult to measure because it is intangible, for instance, advice provided to consumers on ways to conserve scarce water or pay less for their water service.

(b) Reporting of the range of data needed for use in performance information analysis

(i) *What information on inputs and outputs*

Once the costs of inputs and delivery of outputs are measured a decision has to be made on the scope of the information about input cost and output type that the performance assessment requires, and how and when this information should be made available (reported). In this connection there are output targets that have to be selected, e.g. potable water service or waste water service. International experience indicates that it is preferable to have few rather than many targets because too many targets dilute the focus of attention and make prioritisation difficult (Curristine, Lonti and Joumard, 2007).

(ii) *When the information should be made available*

The information has to be reported in a standardised (uniform) way so that valid cross-water service provider (municipal) comparisons can be made. With guidance these measurement and information requirement challenges should not be too difficult to overcome in the water services sector. Supplying the information could be made part of the required annual budgetary process. Regular performance assessment permits improvement or the lack of it to be determined.

(iii) *Balancing the cost of implementing performance assessment against the benefits*

There has to be a balance struck between increasing the scope for performance assessment and imposing excessively on water service providers to supply information annually. The wider the range of input and output information captured, the greater the potential range of inputs and outputs that can be utilised in the efficiency modelling process, but the greater the cost imposed on the managers of the water service provision.

(c) Sufficient sample size

The higher the number of water service providers (municipalities) included in the performance monitoring exercise, the greater the range of inputs and outputs that can be incorporated in the models estimated. Small sample sizes reduce the scope for identifying inefficient municipalities – increasing the possibility that municipalities are assessed as efficient in water service provision only because the sample size was so small that they could not be shown up as inefficient.

(d) Authentic (accurate) data reported

As important as anything else about performance monitoring is that the data used to calculate the efficiencies must be authentic (credible). As with all calculations, the results are only as accurate and reliable as the accuracy of the numbers used in the calculations.

Some municipalities in South Africa expressed concerns at the Stone Cradle National Seminar on Water Service tariffs, held on 14 June 2013, with performance monitoring in South Africa on the grounds that the data that would be used would be too unreliable – 'garbage in garbage out'. Specifically with respect to relative efficiency benchmarking, the concern expressed was that the water service outputs reported and used in the pilot demonstration assessment were in many cases lower than what was really being delivered. It was alleged that large sections of the service were, in fact, non-operational for substantial periods and that this inflated selected efficiency ratings relative to others. In order to address this concern and preserve the integrity of the results, some form of information verification is clearly required (Gibson, August 2013).

Guideline Two – the municipalities must be grouped into equivalent efficiency potential cohorts

The circumstances faced by the various municipalities often differ dramatically, and for various reasons, many of these differences cannot be reflected (captured) in the inputs (as a distinct variable). The presence of these differences provide a non-efficiency (circumstantial difference) explanation for divergence between the municipalities in the efficiency ratings, so the difference in efficiency index may not, in fact, be due to a difference in efficiency but a difference in some other factors (the circumstances facing the municipality).

One solution to this problem (the inevitable inadequacies and omissions of the data and modelling process), is to stratify the Water Service Providers into like groups, that better equate their circumstances, *prima facia.* In this way the water service provision challenges within the group are made more comparable/standard.

The case for this stratification, though, is not overwhelming. One of the features of the estimation methodologies is their capacity to identify inefficiency is enhanced the greater the number of observations generated within each grouping or cohort (more is better). In this way the methodology can be said to encourage efficiency to be compared between municipalities facing substantially different circumstances. However, there are risks in applying the methodology in this way, because there is advantage in comparing 'oranges with oranges', for instance, comparing metropolitan municipalities with each other and district municipalities with each other, etc. The reason is that the various groups face different potentials for efficiency, and the relevant policy issue is more one related to realising cost potential than it is about the revealing of different cost potentials. Metropolitan municipalities are commonly perceived to enjoy a number of cost and personnel advantages – due to economies of scale, state of economic development and technical expertise available and history of water resource development. For this reason there is a good case for analysing relative efficiency within groups of roughly equivalent cost potentials for water service delivery.

Guideline Three – input and output variable selection must be that most appropriate for estimating the efficiency benchmarks (frontier) of interest

The selection of variables as inputs and outputs is not as straightforward a matter as it may seem at first glance. What some model estimators may deem inputs, others may deem outputs, and vice versa. Intuitively the length of pipes used to reticulate the potable water is an input into production, but in our DEA application (and many others) it was selected (in practice) as an output because it reflects the reach of the delivery of the water service. The classification of measures as inputs or outputs influences the results.

Guideline Four – a weighting needs to be given to relative efficiency performance in an overall performance benchmarking assessment

Relative efficiency, technical or allocation, measures how well inputs hired are being used to deliver water service and how cost sensitive the budget allocation is on the inputs used to deliver water services. It is one element in an overall assessment of production performance, and should be seen as such. A comprehensive performance assessment also includes three other types, viz.:

- how well the delivery of the water service provider matches up to guidelines for quality of service set nationally – in the South African context, assessed through Blue and Green Drop ratings
- the portions of the water service delivered for which there is under- revenue collected, including that part for which no revenue is collected (because the water is 'lost', *inter alia*) and

- the level of consumer welfare generated and what perceptions consumers have about the quality and value for money of the service they receive.

These four different types of performance assessment should feed into and condition each other. Each type of performance assessment reveals something unique, so they need to be interpreted in conjunction with each other. A weak performance in one area may be a motive for taking a more detailed look at the way the service is being managed, or its cause may be readily explained by considering the whole (all four assessments), or its cause may be an (obvious) temporary phenomenon . Each case needs to be considered on its own merits. In principle, there is no reason to expect trade-offs to occur between the different assessments of performance, but the possibility of such trade-offs should not be entirely excluded.

Guideline Five – There must be credible and timely performance assessment feedback and clarity provided on the link to planning and management interventions to improve performance

The embracement and operation success of performance information may be enhanced by politicians citing it in voter campaigns and senior government officials citing it as a decision making influence. More generally, buy-in from the key stakeholders in the water service sector (including particularly local and national levels of government) is crucial to the success in using performance information as a vehicle for bringing about positive change in the water service. A long-term goal pursuing approach is advocated, which follows a learning-by-doing route, employs piloting schemes to test for any unintended behavioural consequences and is flexible enough to adjust to the local context (Curristine, Lonti and Joumard, 2007).

Some OECD country guidelines for making performance information relevant are (Curristine, Lonti and Joumard 2007):

- adapting implementation to the prevailing political and institutional context and other relevant (evolving) market circumstances
- instituted it as part of the government planning and reporting requirement
- integrating it into the budget process, but not automatically linked to budget allocations (care has to be taken not to create perverse incentive outcomes, but the direct involvement of the Treasury is critical)
- making sure independent assessment of the performance information is made

- incentivising civil servants and politicians to aligned themselves with the goal of improving measured efficiency performance.

In order to incorporate performance information into the budget process may require a restructuring of the way budgets are configured and the requirements of the budgetary process, as well as a certain amount of training. To get the system up and running requires additional resources so that the necessary expertise is available to set up and train the personnel who will operate the performance assessment system. Currently budgets are structured in accordance with institutional and functional boundaries. They are not allocated according to result categories. For this reason, not only is it difficult to relate costs to service actually delivered, but the making of this link is not actually encouraged by the budget process and annual cycle. Some may argue that this is as it should be. Making a linkage between performance assessment or improvement and budget allocation is consistent with the goals of a performance assessment system, but there are risks to and limits in making this linkage

7.6 Conclusion

The South Africa municipal water services providing sector is suited to the monitoring of efficiency performance benchmarking. There are a large number of separate water service providers and their performances differ widely, but there is sufficient commonality between them to reject the uniqueness and welfare counter arguments against the implementation of performance benchmarking (Chapter 1). This conclusion is a qualified one. There are many challenges that need to be resolved before the efficiency indices generated can be relied upon to guide policy and regulatory effort to improve efficiency performance; the most important of which are input data adequacy and authenticity. This study was able to access relevant (the best available) data for just over 50 municipalities in order to demonstrate how SFA could be used to rate their efficiency performance. With a broader range of data and a bigger sample size a statistically more reliable and preferred model could be estimated.

The study recommends that there take place discussion with and between the various water service authorities (municipalities) on how the data problem can and should be addressed and how the problem of measured inefficiency should be addressed once it is identified. The rationale for statistical performance benchmarking depends on there being a process initiated subsequent to measurement whereby efficiency performance is improved. This process has to be managed sensitively, and has to take into account that there are limits on the scope for engineering change. If managed insensitively, can send negative signals and undermine employee morale.

The introduction of efficiency performance benchmarking, and other types of benchmarking, often begins with high and unrealistic expectations of what such performance assessments and reporting can deliver (Curristine, Lonti and Joumard 2007). The simplicity of the idea is appealing – a measuring system that informs all providers of the service annually who is efficient and who not – thereby enabling a set of interventions to be initiated aimed at improving efficiency amongst the inefficient.

How should the interventions be effected – by a system of rewards and punishments? The rewards can be of many types, but the most common and enduring instrument of correction that has been adopted appears to be the budget. This instrument for correction would be problematic in South Africa because the budget decision making takes place in a political setting and is subject to many constraints. In the water services sector legal and national policy guidelines in South Africa constrain how budgets may be allocated in the provision of water services, and limit the scope for regulatory interventions. For this reason, other instruments for encouraging efficiency performance improvements may be needed, but they must be credible and effective. The experience of some OECD countries was that it took up to five years to establish a fully- fledged operational performance measurement structure (Curristine, Lonti and Joumard, 2007).

There is much that has been achieved in South Africa in the monitoring of municipal water service performance through Blue Drop and Green Drop ratings (and the Non-Revenue water proportions of the total service). This study has demonstrated that the monitoring of efficiency in water service delivery is also very feasible. It has demonstrated that relative efficiency can readily be monitored and cost efficiency indices calculated by applying either parametric or non-parametric methodologies – corrected ordinary least squares (COLS), stochastic frontier analysis (SFA) and or data envelopment analysis (DEA). Depending on what the objective of the efficiency assessment is one or other of these methodologies is preferred. This study prefers the DEA method because it more closely aligns with the primal maximising problem facing the municipal providers of water services.

Reviewing the experience of others in implementing performance monitoring, it appears that, if monitoring and improvement arrangements of relative efficiency are to be successfully implemented, buy-in will have to be obtained from those municipalities being rated. Part of the requirement for this buy-in will be agreement on the formula and data input used to calculate the efficiency indices and the groupings appropriate for their calculation (who they are to be compared against).

Five guidelines for the implementation of an efficiency performance system are advocated. They are that:

- sufficient data must be available, and the data must be generated through an adequate system of water service accounting, reported annually, and be subject to an independent validation process (audit)
- the municipalities must be grouped into equivalent efficiency potential cohorts
- input and output variable selection must be that most appropriate for estimating the efficiency benchmarks (frontier) of interest
- a weighting needs to be given to relative efficiency performance in an overall performance benchmarking assessment, and
- there must be credible and timely performance assessment feedback and clarity provided in the link to planning and management interventions to improve performance.

REFERENCES

Aida, K., William, W., Jesús, T. and Toshiyuki, S. (1998). 'Evaluating water supply services in Japan with RAM: a range-adjusted measure of efficiency'. Omega, International Journal of Management Science, 26(2), 207-232.

Aigner, D. J. and Chu, S. F. (1968). 'On Estimating the Industry Production Function'. The American Economic Review, 58(4), pp. 826-839.

Aigner, D., Lovell, C. and Schmidt, P. (1977). 'Formulation and estimation of stochastic frontier production function models'. Journal of Econometrics, 6(1), 21-37.

Akoto, W. 2009. Valuing preferences for freshwater inflows into selected western and Eastern Cape estuaries. Master's Dissertation. Port Elizabeth: Nelson Mandela Metropolitan University.

Akter, S., Brouwer, R., Chowdhury, S., and Aziz, S. 2007. Testing reliability and construct validity of in-kind WTP responses in contingent valuation. Available from the World Wide Web: http://www.prem-online.org/archive/16/doc/PREM%20WP%2007-07.pdf [Accessed 13 November 2011]

Allanson, B.R. and Baired, D.1999. Estuaries of South Africa. Cape Town: Cambridge university press.

Ammons, D. N. (1999). 'A proper mentality for benchmarking'. Public Administration Review, 59(2), 105-109.

Ammons, D. N., Coe, C. and Lombardo, M. (2001). 'Performance-comparison projects in local government: Participants' perspectives'. Public Administration Review, 61(1), 100-110.

Andersen, P. and Petersen, N. C. (1993). 'A procedure for ranking efficient units in data envelopment analysis'. Management Science, 39(10), 1261-1264.

Arrow, K., Leamer, E.E., Portney, P.R., Radner, R, Schuman, H., Solow R. 1993. Report of the NOAA Panel on Contingent Valuation. Federal Register. Volume 58, Number 10. Pages 4601 to 4614. Available from the World Wide Web: http://www.darp.noaa.gov/library/pdf/cvblue.pdf [Accessed 15 May 2011]

Ashton, J. K. (2000). ‘Cost efficiency in the UK water and sewerage industry’. Applied Economics Letters, 7(7), 455-458.

Baird, D. 2006. Swartkops River Estuary, Algoa Bay, Eastern Cape, South Africa. Available from the World Wide Web: http://nest.su.se/mnode/africa/S_Africa/Swartkops/swartkopsbud.htm [Accessed 24 February 2011]

Banker, R., Charnes, A. and Cooper, W. (1984). ‘Some models for estimating technical and scale inefficiencies in data envelopment analysis’. Management Science, 30(9), 1078-1092.

Barbier, E. B., Acreman, M. and Knowler, D. 1997. Economic valuation of wetlands: a guide for policy makers and planners. Gland: Ramsar Convention Bureau.

Bate, G., Smailes, P., and Adams, J. 2004. A water quality index for use with diatoms in the assessment of rivers. ISSN 0378-4738 = Water SA Vol. 30 No.

Battese, G. E. (1991). Frontier production functions and technical efficiency: A survey of empirical applications in agricultural economics. accessed: 27 November 2012. URL: http://www.une.edu.au/business-school/working-papers/econometrics/emetwp50.pdf DRAFT

Battese, G. E. and Coelli, T. J. (1988). ‘Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data’. Journal of Econometrics, 38(3), 387-399.

Berg, S. and Lin, C. (2008). ‘Consistency in performance rankings: the Peru water sector’. Applied Economics, 40(6), 793-805.

Binning, K. & Baird, D. 2001. Survey of heavy metals in the sediments of the Swartkops River Estuary. Port Elizabeth, South Africa. Water SA 27(4): 461-466.

Bogetoft, P. and Otto, L. (2010). Benchmarking with DEA, SFA, and R. Vol. 157 of International Series in Operations Research & Management Science. Springer Science+Business Media, LLC.

Bogetoft, P. and Otto, L. (2013). Benchmarking with DEA and SFA. R package version 0.23.

Bottasso, A. and Conti, M. (2003). Cost inefficiency in the English and Welsh water industry: An heteroskedastic stochastic cost frontier approach. University of Essex.

Bowerman, M. and Ball, A. (2000). 'The modernisation and importance of government and public services: Great expectations: Benchmarking for best value'. Public Money & Management, 20(2), 21-26.

Braadbaart, O. (2007). 'Collaborative benchmarking, transparency and performance: Evidence from The Netherlands water supply industry'. Benchmarking: An International Journal, 14(6), 677-692.

Bruder, K. A. and Gray, E. M. (1994). 'Public-sector benchmarking: A practical approach'. Public Management, 76(9), S9-S14.

Burke, J. (2004). Why are provinces and territories promoting performance measurement in municipalities?. Provincial-Territorial Charrette on Municipal Performance and its Measurement (May 2004).

Cazals, C., Florens, J.-P. and Simar, L. (2002). 'Nonparametric frontier estimation: a robust approach'. Journal of Econometrics, 106(1), 1-25.

Chadwick H. 2001. Organizational Research: Determining Appropriate Sample Size. Information Technology and Performance Journal, Volume 19, Number 1, 43-50.

Charnes, A., Cooper, W. and Rhodes, E. (1978). 'Measuring the efficiency of decision making units'. European Journal of Operational Research, 2(6), 429-444.

Christensen, L. R. and Greene, W. H. (1976). 'Economies of scale in U.S. electric power generation'. Journal of Political Economy, 84(4), pp. 655-676.

Cochran, W. G. 1963. Sampling Techniques. 2nd Ed., New York: John Wiley and Sons, Inc.

Cochran, W.G. (1977). Sampling techniques. New York: Wiley and Sons.

Coelli, T. and Henningsen, A. (2012). frontier: Stochastic Frontier Analysis. R package version 0.997-12. URL: http://CRAN.R-project.org/package=frontier

Cohen, J and Cohen, P. 1975. Analytic Strategies: Simultaneous, Hierarchical, and Stepwise Regression. Available from the World Wide Web: http://www.nd.edu/~rwilliam/stats1/x95.pdf [Accessed 6 November 2011]

Cooper, W. and Tone, K. (1997). 'Measures of inefficiency in data envelopment analysis and stochastic frontier estimation'. European Journal of Operational Research, 99(1), 72-88.

Cooper, W. W., Seiford, L. M. and Tone, K. (2006). Introduction to Data Envelopment Analysis and Its Uses: With DEA-Solver Software and References. Springer Science+Business Media, Inc.

CSIR. 2010. A CSIR perspective on water in South Africa – 2010. CSIR Report No. CSIR/NRE/PW/IR/2011/0012/A. ISBN: 978-0-7988-5595-2.

Cubbin, J. and Tzanidakis, G. (1998). 'Regression versus data envelopment analysis for efficiency measurement: an application to the England and Wales regulated water industry'. Utilities Policy, 7(2), 75-85.

Curristine T., Lonti Z. and Joumard I. (2007) 'Improving Public Sector Efficiency: Challenges and Opportunities'. OECD Journal on Budgeting, Volume 7 (1): 1-41.

Davids, A.S. 2002. GIS using 1996 census data. Human Sciences Research council. Pretoria.

Debreu, G. (1951). 'The coefficient of resource utilization'. Econometrica, 19(3), pp. 273-292.

Department of Economic Development and Enviromental Affairs. 2009. The Eastern Cape State of the Environment Report (2nd Edition)

Department of Water Affairs and Forestry (DWAF). 2001. Managing the water quality effects of settlements: Legal considerations for managing pollution from settlements. Department of Water Affairs and Forestry, Pretoria.

Department of Water Affairs and Forestry (DWAF). 2006. Eastern Cape Provincial Water Sector Plan 2007/08 to 2011/12. Available from the World Wide Web: http://www.dwa.gov.za/masibambane/documents/strategies/pwsp/ec/ec-prov-watersector-plan%202007.pdf [Accessed 21 April 3 2011]

Department of Water Affairs and Forestry (DWAF). 2008a. Directorate Water Services Regulation: Annual publication 2007/08. Draft No. 8. DWAF, Pretoria.

Department of Water Affairs and Forestry (DWAF). 2008b. National Water Services Regulation Strategy. Penultimate draft. DWAF, Pretoria.

Department of Water Affairs and Forestry (DWAF). 2009. Blue Drop System, Drinking Water Quality Regulation. Available from the World Wide Web: http://www.dwaf.gov.za/dir_ws/DWQR/ [18/10/2009]. [Accessed 21 February 2011]

Department of Water Affairs and Forestry. 2009. Water affairs and forestry. Available from the World Wide Web: http://www.info.gov.za/aboutsa/water.htm#intro [Accessed 25 February 2010]

Department of Water Affairs. 1986. Management of the resources of the Republic of South Africa.

Development Bank of Southern Africa. 2009. Municipal compliance with water services policy: A challenge for water security. Development Planning Division Working Paper Series No. 10.

Dorsch, J. J. and Yasin, M. M. (1998). 'A framework for benchmarking in the public sector: Literature review and directions for future research'. International Journal of Public Sector Management, 11(2), 91-115.

DWA (2011a). Blue Drop Handbook. 1st edn.

DWA (2011b). Green Drop Handbook. 1st edn.

DWAF (2009). Water Quality Regulation: A strategy for incentive based regulation. Blue and Green Drop certification.

DWAF, SALGA & WRC (2005). National Water Services Benchmarking Initiative: Promoting best practice. Benchmarking outcomes for 2004/5.

DWAF, SALGA and WRC (2005). National Water Services Benchmarking Initiative: Promoting best practice. Benchmarking outcomes for 2004/5. December 2005.

DWAF, SALGA and WRC (2007). National Water Services Benchmarking Initiative: Promoting best practice. Benchmarking outcomes for 2005/2006. May 2007.

DWAF, SALGA and WRC (2008). National Water Services Benchmarking Initiative: Report on 2006/2007 Performance. August 2008.

Dyson, R., Allen, R., Camanho, A., Podinovski, V., Sarrico, C. and Shale, E. (2001). 'Pit-falls and protocols in DEA'. European Journal of Operational Research, 132(2), 245-259.

Eales, K. 2009. Water services in South Africa, 1994-2000. In Schreiner, B and Hassan, R (Eds), Water policy in South Africa: Resources for the future.

Efron, B. and Tibshirani, R. (1993). An introduction to the bootstrap. London, United Kingdom: Chapman-Hall.

Farrell, M. J. (1957). 'The measurement of productive efficiency'. Journal of the Royal Statistical Society. Series A (General), 120(3), pp. 253-290.

Filippini, M., Hrovatin, N. and Zoric, J. (2008). 'Cost efficiency of Slovenian water distribution utilities: an application of stochastic frontier methods'. Journal of Productivity Analysis, 29, 169-182.

Fink, A. 2003. How to sample in surveys. Thousand Oaks, California: Sage Publications, Inc.

Folz, D. H. (2004). 'Service quality and benchmarking the performance of municipal services'. Public Administration Review, 64(2), 209-220.

Fraquelli, G. and Moiso, V. (2005). Cost efficiency and economies of scale in the Italian water industry. XVII Conferenza società italiana di economia pubblica. Finanziamento del settore pubblico.

Ganley, J. A. and Cubbin, J. S. (1992). Public Sector Efficiency Measurement: Applications of Data Envelopment Analysis. Elsevier Science Publishers B. V.. Amsterdam, The Netherlands.

Ganley, J. and Cubbin, J. (1992). Public sector efficiency measurement: applications of data envelopment analysis. Elsevier Science Publishers.

García-Valñasa, M. A. and Muñiza, M. A. (2007). 'Is DEA useful in the regulation of water utilities? A dynamic efficiency evaluation of water utilities'. Applied Economics, 39(2), 245-252.

Gibson, J. (August 2013) National Benchmarking Initiative team member. Personal Communication.

Greene, W. H. (2008). The Measurement of Productive Efficiency and Productivity Growth. Fried, H. O., Knox Lovell, C. A. and Schmidt, S. S. (eds). Oxford University Press, Inc. Chapter: The Economic Approach to Efficiency Analysis, pp. 92-250.

Gujarati, D. 2003. Basic Econometrics. New York: McGraw Hill.

Hair, J.F., Anderson, R.E., Tatham R.L. and Black, W.C. 1998. Multivariate data Analysis. 5th edition. New Jersey: Prentice-Hall.

Hanley, N. and Spash, C.L. 1993. Cost-Benefit Analysis and the Environment. Vermont: Edward Elgar Publishing.

Hill, R.C, Griffiths, W.E, Lim G.C. 2008. Principles of Econometrics. 3rd Edition.

Hilmer, T and Bate, G.C. 1987. Hydrocarbon levels in the Swartkops estuary: a preliminary study. Department of Botany, Institute for Coastal Research, University of Port Elizabeth, P. 0. Box 1600, Port Elizabeth 6000, South Africa. ISSN 0378-4738 = Water SA Vol. 13. No. 3.

Hook and Reel Angling Club. Swartkops, South Africa. Available from the World Wide Web: http://www.hookandreel.org/ [Accessed 21 April 1 2011]

Hosking, S.G, Woodridge, TH, Dimopoulos, Mlangeni, M, Lin, C-H, Sale, M and Du Preez. 2004. The valuation of changes to estuary services in South Africa as a result of changes to freshwater inflow. Department of Economics and Economic History and Department of Zoology/University of Port Elizabeth. Report to the Water Research Commission. WRC Report No: 1304/1/04. ISBN: 1-77005-278-x.

Hosking, S.G., Du Preez, M., and Sale, M.C. 2009. Application of the contingent valuation method to estimate a recreational value for the freshwater inflows into the Kowie and the Kromme Estuaries. Available from the World Wide Web: http://www.scielo.org.za/pdf/wsa/v35n3/a04v35n3.pdf [Accessed 23 February 2010]

Hosking, S.G., Du Preez, M., Cambell, E.E., Woodridge, T.H. and Du Plessis, L.L. 2002. Evaluating the Environmental use of water-selected case studies in the Eastern and Southern Cape. Pretoria: Water Research Commission. Report Number 1045/1/102.

Israel, Glenn D. 1992. Sampling the evidence of extension program impact. Program Evaluation and Organizational Development, IFAS, University of Florida. PEOD-5.

Johnston, J. (1960). Statistical cost analysis. Economics handbook series. McGraw-Hill.

Jondrow, J., Lovell, C. K., Materov, I. S. and Schmidt, P. (1982). 'On the estimation of technical inefficiency in the stochastic frontier production function model'. Journal of Econometrics, 19, 233- 238.

King, D.M. Mazzotta, M.J. 2000. Ecosystem Valuation. Available from the World Wide Web: http://www.ecosystemvaluation.org/contingent_valuation.htm [Accessed 23 February 2010]

Knox Lovell, C. (1993). The Measurement of Productive Efficiency: Techniques and Applications. Fried, H. O., Knox Lovell, C. A. and Schmidt, S. S. (eds). Oxford University Press. Chapter: Production frontiers and productive efficiency, pp. 3-67.

Kreiberg, D. 2010. Diagnostic testing in Eviews. Available from the World Wide Web: http://www.docstoc.com/docs/21468796/Overview-of-the-classical-linear-regression-model [Accessed 3 November 2011]

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. Educational and Psychological Measurement, 30, 607-610.

Maddala, G.S. 2001. Introduction to Econometrics. 3rd edition. Chichester: John Wiley and Son Ltd.

Magobiane, S. 2012. Willingness to Pay for water quality changes in the Swartkops estuary. Unpublished M Com dissertation, Nelson Mandela Metropolitan University, Port Elizabeth, South Africa.

Mallory, S.J.L, and Pashkin, A.E. 2008. Risk Assessment and Water Resources Planning for specific high-profile events. Available from the World Wide Web: http://www.waterresources.co.za/downloads/ExtremeWaterMosow.pdf [Accessed 21 April 3 2011]

McKenzie, R., Siqalaba, Z. N. and Wegelim, W. A. (2012). The State of Non-Revenue Water in South Africa (2012). Report TT 522/12. Water Research Commission.

Mckenzie, R.; Siqalaba, Z.N. & Wegelin, W.A. (2012). The State of Non-Revenue Water in South Africa (2012). Water Research Commission, Report TT 522/12.

Meeusen, W. and van Den Broeck, J. (1977). 'Efficiency estimation from Cobb-Douglas production functions with composed error'. International Economic Review, 18(2), pp. 435-444. DRAFT

Mendenhall, W. & Sinchich, T. 1996. A second course in statics – Regression analysis. 5th Edition. New Jersey: Prentice Hall.

Meyer, W.N. 2007. The Economics of Water: water for life; sanitation for dignity. 1st Edition. Pretoria: Van Schaik Publishers.

Mittelstaedt, Jr., R. E. (1992). 'Benchmarking: how to learn from best-in-class practices?'. National Productivity Review, 11, 301+.

Mlangeni, M.M. 2007. A contingent valuation of river water into the Swartkops. Kariga, Mngazi and Mngazana estuaries in the Eastern Cape. Master's Dissertation. Port Elizabeth: Nelson Mandela Metropolitan University.

Moraka, W., Bhagwan, J., Mackintosh, G. and Stevens, F. (2011). Improving water services provision through the National Municipal Benchmarking Initiative: Benchmarking our way to better services, more effectively, more efficiently. IMESA 2011 Conference.

Moraka, W., Bhagwan, J., Mackintosh, G. and Stevens, F. (2012). Improving municipal water services provision through the National Benchmarking Initiative. WISA 2012 Conference.

Murillo-Zamorano, L. R. (2004). 'Economic efficiency and frontier techniques'. Journal of Economic Surveys, 18(1), 33-77.

Oberholster, P.J. and Ashton, P.J. 2008. State of the nation: An overview of the current status of water quality and eutrophication in South African rivers and reservoirs. CSIR Report No. CSIR/NRE/WR/IR/2008/0075/C. Council for Scientific and Industrial Research: Pretoria, South Africa. p. 14.

Oelofse, SHH and Godfrey, L. 2008. Towards improved waste management services by local government – A waste governance perspective. Science real and relevant: 2nd CSIR Biennial Conference, CSIR International Convention Centre Pretoria, 17&18 November 2008, pp 9.

OMBI (2011). About: The Ontario Municipal Benchmarking Initiative. Accessed: 20 June 2012. URL: www.ombi.ca.

Ondrich, J. and Ruggiero, J. (2001). 'Efficiency measurement in the stochastic frontier model'. European Journal of Operational Research, 129(2), 434-442.

Pybus, P., Sham, B., Schoeman, G., Mogale, T. and Connolly, J. (2005). Pilot initiative to facilitate benchmarking in the water services sector. Report. Water Research Commission (WRC). KV165/05.

R Core Team (2013). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: http://www.R-project.org/

Roux, S.P, De Lange, W, Oelofse, S.H.H. 2010. The rising costs of both sewage treatment and the production of potable water associated with increasing levels of pollution in a portion of the Crocodile-West Marico water management area (A Case Study). Available from the World Wide Web: http://researchspace.csir.co.za/dspace/bitstream/10204/4209/1/Roux_2010.pdf [Accessed 27 February 2011]

Ruggiero, J. (1999). 'Efficiency estimation and error decomposition in the stochastic frontier model: A Monte Carlo analysis'. European Journal of Operational Research, 115(3), 555-563.

Ruggiero, J. (2011). Frontiers in Major League Baseball: Nonparametric Analysis of Performance Using Data Envelopment Analysis. Sports Economics, Management and Policy. Springer New York.

Sapa (2011). Hold us to account, Manuel urges MPs. Fin24 (5 September 2011). Accessed: 21 June 2012. URL: http://www.fin24.com/Economy/Hold-us-to-account-Manuel-urges-MPs-20110905

Schmidt, P. and Lovell, C. K. (1979). 'Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers'. Journal of Econometrics, 9(3), 343-366.

Sebugwawo, M. (2011). Service delivery protests in South Africa: Lessons for municipalities. Afesis-corplan. Accessed: 21 June 2012. URL: http://www.afesis.org.za/Local-Governance-Articles/

Shepard, R. W. (1970). Theory of Cost and Production Functions. Princeton University Press.

Simar, L. and Wilson, P. W. (1998). ‘Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models’. Management Science, 44(1), pp. 49-61.

Simar, L. and Wilson, P. W. (1999). ‘Estimating and bootstrapping Malmquist indices’. European Journal of Operational Research, 115(3), 459-471.

Simar, L. and Wilson, P. W. (2000). ‘A general methodology for bootstrapping in non-parametric frontier models’. Journal of Applied Statistics, 27(6), 779-802.

Simar, L. and Wilson, P. W. (2008). The Measurement of Productive Efficiency and Productivity Growth. Fried, H. O., Knox Lovell, C. A. and Schmidt, S. S. (eds). Oxford Univeristy Press, Inc. Chapter: Statistical Inference in Nonparametric Frontier Models: Recent Developments and Perspectives, pp. 421-521.

Simar, L. and Wilson, P. W. (2011). Handbook on Data Envelopment Analysis. 2nd edn. Cooper, W. W., Seiford, L. and Zhu, J. (eds). Springer Science+Business Media. Chapter: Performance of the Bootstrap for DEA Estimators and Iterating the Principle, pp. 241-272.

Singh, M. R., Mittal, A. K. and Upadhyay, V. (2011). ‘Benchmarking of North Indian urban water utilities’. Benchmarking: An International Journal, 18(1), 86-106.

Social Research Methods. 1993. Available from the World Wide Web: http://www.socialresearchmethods.net/Gallery/Kong/Contingent%20valuation.html#1.2.2 [Accessed 25 March 2011]

Spendolini, M. J. (1992). The Benchmarking Book. AMACOM. New York, NY.

StatsSA (2011a). P9114 – financial census of municipalities 2010/2011. URL: http://www.statssa.gov.za/publications/statspastfuture.asp?PPN=P9114

StatsSA (2011b). P9115 – non-financial census of municipalities 2010/2011. URL: http://www.statssa.gov.za/publications/statspastfuture.asp?PPN=P9115

Steering Committee for the Review of Commonwealth/State Service Provision (1997). Data Envelopment Analysis: A technique for measuring the efficiency of government service delivery. Commonwealth of Australia. AGPS, Canberra.

Stone Cradle Workshop (June 2013). National Seminar of Tariffs , Costs, Revenue and Regulation in the municipal water services. Stone Cradle, Pretoria, South Africa.

Swartkops Integrated Environmental Management Plan. (2009). C.A.P.E. Estuaries management Programme. Available from the World Wide Web: http://www.nelsonmandelabay.gov.za/Documents.aspx?objID=85&cmd=download&docID=672. [Accessed 21 March 2011]

Tait, S. and Marks, M. (2011). 'You strike a gathering, you strike a rock: Current debates in the policing of public order in South Africa'. SA Crime Quarterly, 38, 15-18. Institute for Security Studies.

Taylor, J. (2009). 'Strengthening the link between performance measurement and decision making'. Public Administration, 87(4), 853-871.

Thanassoulis, E. (2000a). 'DEA and its use in the regulation of water companies'. European Journal of Operational Research, 127(1), 1-13.

Thanassoulis, E. (2000b). 'The use of data envelopment analysis in the regulation of UK water utilities: Water distribution'. European Journal of Operational Research, 126(2), 436-453.

The Water Wheel (2011a). 'Authorities show commitment but Green Drops still elusive'. The Water Wheel (November/December 2011), 10(6), 27-29.

The Water Wheel (2011b). 'Blue Drop: Slow but steady progress in restoring SA's drinking water quality'. The Water Wheel (September/October 2011), 10(5), 18-20.

Tillema, S. (2010). 'Public sector benchmarking and performance improvement: what is the link and can it be improved?'. Public Money and Management, 30(1), 69-75.

Triantafillou, P. (2007). 'Benchmarking in the public sector: a critical conceptual framework'. Public Administration, 85(3), 829-846.

Tupper, H. C. and Resende, M. (2004). 'Efficiency and regulatory issues in the Brazilian water and sewage sector: an empirical study'. Utilities Policy, 12(1), 29-40.

Wooldridge, M. 2005. Introductory Econometrics: A Modern Approach. South-Western College Publishing: Thomson.

World Water Assessment Programme Report. 2006. Water: a shared responsibility. Available from the World Wide Web: http://unesdoc.unesco.org/images/0014/001444/144409E.pdf [Accessed 19 February 2010]

APPENDICES

APPENDIX A: ARE CONSUMERS PREPARED TO PAY FOR IMPROVEMENTS IN THE EFFICIENCY OF WATER SERVICE DELIVERY? A CASE STUDY OF WASTE WATER DISCHARGES INTO THE SWARTKOPS ESTUARY

1.1 Introduction

Up until recently most of the attention with respect to municipal water service delivery in South Africa has been focussed on increasing the access of the population to the service, and in this regard much has been achieved (The Water Wheel, April 2008: 33). Nationally the proportion of households with piped water inside a dwelling rose from 32.3% in 2001 to 47% in 2007 and a flushing toilet connected to the sewerage system from 49.1% in 2001 to 55.1% in 2007(The Water Wheel, April 2008: 33).

More recently, more questions have been raised about the efficiency of this delivery, due to reports of wastage, such as 63% of the water within the Msunduzi Municipality being lost to the revenue system (Delivery, June 2011:6). In order to assess the potential for cost savings and water tariff cuts, and also detect where there is gross inefficiency occurring, regulators of water service providers, need to have a method by which to undertake comparative efficiency exercises. Unless it a water service performance is measured it is impossible to improve delivery because one cannot objectively compare before and after situations. One method by which to compare efficiency is through the generation of a relative efficiency measures or an efficiency frontier.

This project will develop and identify such a measure, apply it to a sample of municipalities and show how it may be used to identify inefficient (high cost) water service providers. In pursuing this aim it will complement WRC Project K5/2087 which has as its primary focus of attention demand side efficiency – the sensitivity of municipalities in their provision of water service to demand and the need or scope for a neutral and external national regulator of water services. Such regulators exist, for instance, in England and Wales in the form of the Office of Water Services (OFWAT) and in Germany in the form of Landerarbeitsgemeinschaft Wasser (LAWA). These regulators are not only be interested in demand side efficiency, but also in supply side efficiency – to assess how cost effective is the municipality's provision of water services.

One of the ways of assessing cost effectiveness is the methodology called performance benchmarking analysis (Cubbin and Tzanidakis, 1998; Marques, Berg and Yane, 2011). This project aims to fill the gap in South Africa in supply side relative efficiency assessment of municipal water service providing institutions, by proposing a tool for benchmarking supply efficiency and applying it to a sample of municipalities.

The primary rationale for identifying inefficiency is to reduce the wastage of public funds and resources, and free up resources to be used on other good social causes (elsewhere) – both public resources (the government contribution) and private resources (payments by the water service receivers). The primary intervention agency for managing the relative efficiency performance benchmarking system is a water service regulator. The regulation of the water service providing institutions requires that there be some efficiency measures generated.

As motivation for a municipal water service regulator to introduce efficiency performance monitoring, Chapter One estimates the public willingness to pay to fix inadequate waste water management by Nelson Mandela Bay Municipality, using the contingent valuation method.

1.2 The Swartkops case study

1.2.1 Why it is relevant

Population growth and development in and around estuaries, as well as the coastline in general, is placing indirect pressure on the health and functioning of these ecosystems (Department of Environmental Affairs and Tourism 2001). Of particular concern for human and ecosystem health are the entry of contaminants in the aquatic environment (Department of Water Affairs 1986: 27; Oberholster & Ashton 2008). The reduced quality of water in flows undermines the residential and holiday recreational appeal of estuaries, as well as lack of capacity to support subsistence livelihoods (Hosking, Du Preez, Sale 2009; Water Assessment Programme Report 2006; CSIR 2010).

The poor quality water disposed into rivers and estuaries not only limits its utilisation value but it also places an added economic burden on society through the primary treatment costs and the secondary impacts on the economy, e.g. higher treatment costs, human health (loss in economic activity), and bacterial contamination or bioaccumulation of heavy metals in agricultural goods (CSIR 2010).

A CSIR (2010) study argues the ongoing decline in quality of water in South Africa's estuaries points to South Africa's outdated and inadequate water treatment and sewage treatment plant infrastructure and unskilled operators as reasons, inter alia. The majority of sewage emanating from South Africa's urban areas is not treated properly prior to discharge (CSIR 2010).

The re-usability of return flows depends on the user's requirements and the affordability of the necessary treatment to meet these requirements (Roux, De Lange & Oelofse 2010). Shortages in capacity at these facilities to deal with increased demand requires costly upgrades of existing infrastructure, the building of new wastewater treatment facilities and/or improvement of technologies available to meet user's requirements (Roux, De Lange & Oelofse 2010). The technologies and infrastructure used a decade ago, or less, to treat relatively good quality water are inadequate to cope with current demand (Roux, De Lange & Oelofse 2010).

The ability of municipalities and other entities to install, manage and maintain new wastewater infrastructure has been constrained by large deficits in engineering and technical skills, inadequate capital and operating funds, and skewed compliance incentives (Development Bank of Southern Africa 2009). Incompetently operated treatment works, as well as operating above design capacity, result in wastewater discharges that pose health risks to downstream users. Not only do poorly operated treatment works pose health risks, but also damage the environment, and raise the cost of water treatment and environmental rehabilitation (Development Bank of Southern Africa 2009; DWAF 2008b:17).

The way waste water management services are provided and maintained lie at the heart of the water pollution problem (DWAF 2001). Though the rehabilitation and expansion of waste treatment infrastructure is capital intensive, the release of inadequately or untreated wastewater into surface water resources will have negative effects on the environment and water users, and has to be addressed with urgency (Oelofse & Godfrey 2008).).

Pollution reduces the quality and therefore the economic value of water available (Roux, De Lange & Oelofse 2010). This reduction in water quality impacts on users abstracting water directly from the streams for irrigation and domestic use purposes and may even affect the value of property adjacent to polluted streams (Roux, De Lange & Oelofse 2010). The most noticeable impact in economic terms is on the costs incurred by municipal and private entities responsible for water purification for potable use (Roux, De Lange & Oelofse 2010).

In the long run, there are numerous negative consequences of poor control of pollution, including negative health impacts on informal users, poor quality agricultural produce and increased costs associated with providing the public with potable water. These costs will ultimately have to be recovered from the end users (general public and industry) and will lead to a rise in living costs (Roux, De Lange & Oelofse 2010).

How much estuary user's value can be gained by improving water quality into the Swartkops estuary to safe levels for swimming, fishing and boating? This study addressed this question by eliciting user's willingness to pay for the implementation of a proposed project to improve water quality in the Swartkops estuary to safe levels for swimming, fishing and boating, a method known as the contingent valuation method. The contingent valuation method was applied to elicit user's willingness to pay for the implementation of a proposed project to improve water quality. Information for the literature study was obtained through secondary sources, including internet sources, journals articles, previous research reports, magazines, textbooks and other sources on the subject. Data for this application was collected through a pre-coded contingent valuation method survey questionnaire and administered through face to face interviews at the selected site.

The contingent valuation method is a technique to establish the value of a good or service that is not bought or sold in an actual market (King & Mazzotta 2000). The CVM is used to estimate economic values for all kinds of ecosystem and environmental services. It can be used to estimate both use and non-use values (King & Mazzotta 2000). The CVM establishes the economic value of the good by asking the users of an environmental good to state their willingness to pay for a hypothetical scenario to prevent, or bring about certain changes in the current condition of the environmental good (King & Mazzotta 2000). The users' WTP is aggregated to establish a total willingness to pay (TWTP) for the population of the users of the environmental good (King & Mazzotta 2000).

The data collected from the CVM survey was analysed using regression analysis for the selected estuary. Two statistical models will be used to generate predicted mean and median WTP values. These are the Ordinary least Squares (OLS) and Tobit models. The OLS and Tobit models have continuous dependent variables explaining the variation in respondents WTP (Mendenhall & Sincich 1996). The Tobit Model is commonly used in CVM studies to describe the relationship between WTP (the dependent variable) which is non-negative, and a vector of explanatory variables. The use of the method of ordinary least squares (OLS) for this type of analysis leads to parameter estimates being biased and inconsistent, and incorporates negative predicted WTP values which is incorrect from a theoretical perspective (Mendenhall & Sincich 1996; Hill, Griffiths & Lim, 2008).

1.2.3 The study site

The Swartkops estuary is situated near Port Elizabeth and discharges into Algoa Bay (Figure 1.1, Baird 2006). In addition to the Swartkops River it has three tributaries that flow into it. These are the Motherwell and Markman canals, which are storm-water canals draining residential and industrial township areas respectively, and the Chatty River, draining an area where informal settlements have been established. The estuary is approximately 15 km long and has a total surface area of 4 km^2 (Baird 2006), it is located (33°52' S; 25°38' E) in a warm temperate region and is considered to be in a fair condition (Hosking, Wooldrige, Dimopoulos, Mlangeni, Lin, Sale & Du Preez 2004).

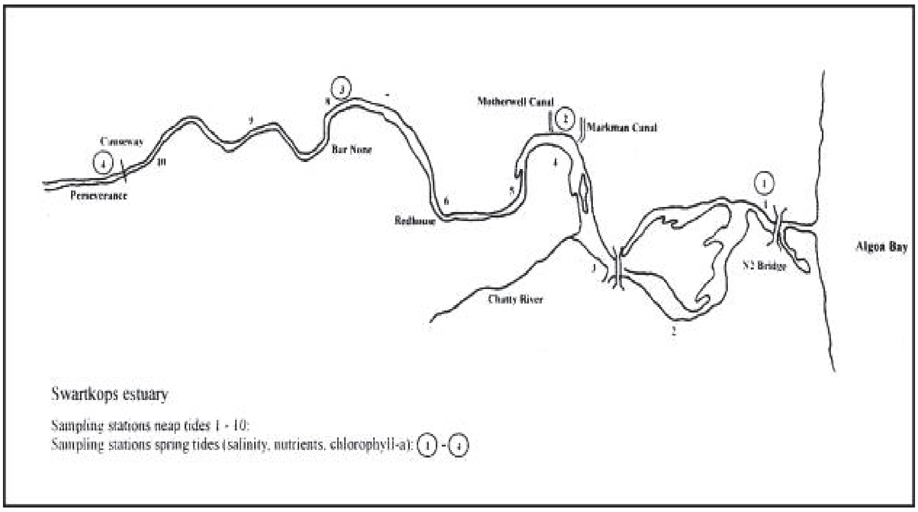

Figure 1.1: The Swartkops estuary *Source*: Baird (2006)

1.2.4 Users of the Swartkops estuary

Estuaries are productive systems that provide valuable supplies of goods and services, varying from fisheries to recreational activities (Lamberth and Turpie 2003). The Swartkops River estuary is no different. Its use ranges from subsistence to recreational. The most common recreational activities in the Swartkops estuary are boating, angling, picnicking, bird watching, walking and jogging along the banks, enjoyment of aesthetics and the scenery and water based sports competitions and team building activities (Mlangeni 2007). Subsistence uses include bait collection, fishing and providing assistance to other users in return for money or food (Mlangeni 2007).

1.2.5 Loss of water quality in the Swartkops

The water quality of the Swartkops River is critically impacted by several anthropogenic sources (Bate, Smailes & Adams 2004), including agriculture, a wool processing factory, three sewerage treatment works, storm water runoff from informal settlements and discharges from light industries, such as leather tanning (Bate, Smailes & Adams 2004).

Due to the absence of human settlement in the upper catchment of the KwaZunga the water quality of the Swartkops River is good, with low levels of pollutants and sediments (Institute for Water Research 2004). As the river passes the urban area below the confluence of the Elands and KwaZunga the water quality deteriorates for a range of reasons. These include:

(1) Increasing but fluctuating coliform bacterial counts due to urban return flows, notably from inadequate sanitation provision.

(2) Increasing variety of salts and other mineral contents, due to the natural condition of the groundwater from the primary aquifer in the lower part of the catchment below Uitenhage.

(3) Increasing nutrient and other content due to agricultural and industrial (arsenic) return flows.

(4) Increasing sediment loads due to in-channel sand mining and deforestation (Institute for Water Research 2004).

Uncontaminated return flows are vital to limit the effects of rising salinity and nutrient levels for agriculture and industrial users (Development Bank of Southern Africa, 2009). The poor quality of discharged municipal effluent increases eutrophication and bacterial contamination of water resources, including ground water (Eales 2009). Waterborne diseases can also spread as a result of bacterial contamination from poorly treated sewage.

The Swartkops estuary receives residential, recreational and industrial pollution to varying degrees due to its closeness to the towns of Port Elizabeth and Uitenhage. Much of the area around the estuary is used for industrial purposes, which include the Swartkops Power Station, the Algorax carbon factory and the South African Transport Services yards. Sewage and industrial effluent are not directly discharged into the estuary, but input above the tidal limit of the river in the form of sewage outfalls, wool washing factory discharge near Uitenhage and a tannery discharge at Perseverance (Hilmer & Bate 1987).

2.1 Applying the CVM

2.1.1 Questionnaire design

The first step of CVM is to set up a realistic and credible market that can be presented to the respondents of the questionnaire, in order to elicit their WTP values. A hypothetical project (Appendix A) was formulated to improve and sustain water quality to safe levels for swimming, fishing and boating in the Swartkops estuary. The project was to increase effectiveness of the Nelson Mandela Municipality Wastewater Treatment works along the Swartkops River and replace faulty joints and pipes in the system. The respondents (users) needed to know how these changes in water quality provision would affect their lives, in order to assist them in the estimation of WTP to implement this project.

The payment card elicitation method was used to elicit WTP for the implementation of a project for the valuation scenario. The method was chosen because of the conservative approach recommended. It was expected to yield conservative values and reduce starting point bias by providing a range of options to choose one alternative (Hosking et al. 2004). The payment options ranged from zero to R4001+ per annum (Appendix A). Additional waste water charges levied on all water services users was chosen as the method of payment, as this was a familiar method of payment to the respondents.

The questionnaire included a variety of questions that help to interpret the responses to the valuation question. The questionnaire was broken down into four sections. Section one gathered socio-economic and household characteristics. Section two collected information on the respondents' attitude towards water quality in the estuary. Section three gathered information on the behavioural characteristics of the respondents. Section four was the valuation section. Its purpose was to obtain the WTP value for water quality changes in the Swartkops estuary. A copy of the questionnaire is found in the Appendix A. It has 16 questions and they are discussed below. A full account of the guidelines and motivation followed in drawing up the questionnaire is reported in Magobiane (2012)

2.1.2 Target population and sample

The target population for this study was those who were located in close proximity to the estuary and residents of Nelson Mandela Bay with a demand for the estuary services, directly or indirectly. No records are kept of their identities and they only reveal their propensity to consume estuary services when they use them. For this reason identifying them *ex ante* was almost impossible. The majority of them were visitors to the area and were not permanent residents of the land adjacent to the estuary. As a consequence, statistically preferred respondent selection procedures could not be applied.

Based on previous research done by Hosking et al. (2004), the household figures in Table 1.1 were used to estimate the sample population of people who use the Swartkops estuary annually.

Table 1.1: Estimated total population of users of households using the Swartkops Estuary per annum by main user category.

Estuary use	**No. household users annually and (%)**
Proximity/viewing	1 000 (19, 2%)
Anglers	2500 (48, 1%)
Bait collectors	100 (1, 9%)
Boaters	1000 (19, 2%)
Bird watching	100 (1, 9%)
Picnicking	200 (3, 8%)
Swimming	100 (1, 9%)
Other (walking)	200 (3, 8%)
Total	**5200 (100%)**

Source: Tiger Bay Boat Registration Office (2003), Hosking et al (2004)

An estuary's ability to provide these services is reliant on the estuary's water quality, amongst other things. For this reason, the users of these services should, in principle, be willing to pay to improve and maintain good water quality. People who actively used the estuary to derive these services were characterized as active users. Other people reveal their demand for estuary services in a more subtle manner, such as those who just enjoy the view of the estuary. These people are passive users of the estuary. Others may not use the estuary either actively or passively, but may derive utility from knowing that it exists, and that they or their children can visit it in the future. These non-current users may also have a willingness to pay to improve and maintain water quality. These active, passive and non-current users make up the target population.

This study used a stratified intercept sampling method to select respondents. The population was divided into subgroups, or strata and a sample was then selected from each group (Fink 2003). Estuary users were divided into strata on the basis of their present or intended use of the estuary. Fishers, boaters, swimmers and birders were categorized as recreational users. Commercial fishers, restaurant operators, B&B and guest house owners were categorized as commercial users. Those people who depended on the estuary for subsistence were categorized as subsistence users. Those who were engaged in activities, not related to active use, were categorized as passive users. The proportion of users in each stratum varied according to the most common activity.

Sample size

Cochran's (1977) formula for determining appropriate sample size was used to determine ideal minimum sample size for the Swartkops estuary survey. To apply this formula we assumed maximum variability (p=0.5) because of the large and diverse population of users. A 95% confidence level was desired and ± 8% level of precision. The resulting sample size is calculated from equations 1.1 and 1.2.

$$n_0 = Z^2 pq / (e)^2 \qquad (1.1)$$

$$= (1.96)^2 (.5)(.5) / (0.08)^2$$

$$= 150$$

Based on the estimated target population size (N), the sample size was computed as follows:

$$n = n_0 / (1 + (n_0 - 1)) / N \qquad (1.2)$$

$$= 150 / (1 + (150 - 1)) / 5200$$

$$= 146$$

In anticipation of less than 100% response rate, the minimum sample size was adjusted upwards. Based on similar research done by Akoto (2009), Lin (2005) and Van Der Westhuizen (2007) a response rate of around 90% was anticipated. The following calculation was used to determine the drawn sample size required to produce the minimum sample size adjusted for anticipated response rate:

Where anticipated response rate = 90%.

Where minimum sample size (n) = 146.

Where n_1= sample size adjusted for response rate.

n_1= 146/0.90 = 162 users of households using the Swartkops estuary

Sample selection

There were 174 questionnaires administered on the intercept method of sample selection. Chapter five summarises the responses to the survey instrument and applies stages three, four and five of the CVM. These stages entail calculating mean and median WTP, estimating a bid curve, predicting mean and median WTP and aggregating a TWTP. Two statistical models were used for this purpose, namely the OLS and Tobit models.

The relevant descriptive statistics are provided below, namely, socio-economic and household characteristics, attitudinal responses, characterisation of behaviour and contingent valuation WTP responses.

Of the respondents 55% were visitors from around Nelson Mandela Bay, 43% were permanent residents and 2% were holiday visitors. The gender participation rate of the respondents was split 80% for males and 20% for females. The average age of the respondents was 43 years. On average there were 3 people in each household. Of the respondents 31% were black, 51% white, 16% coloured and 3% Indian. The majority of respondents (51%) were employed, with 37% and 12% being unemployed and retired respectively. Most of the respondents had some form of formal education with only 7% having none. The respondents pre-tax income ranged from zero to R500 000 plus. The majority (70%) of the respondents' pre-tax income ranged between zero and R100 000 (Figure 1.2), with the highest frequency of respondents (21%) being between R1 and R10 000.

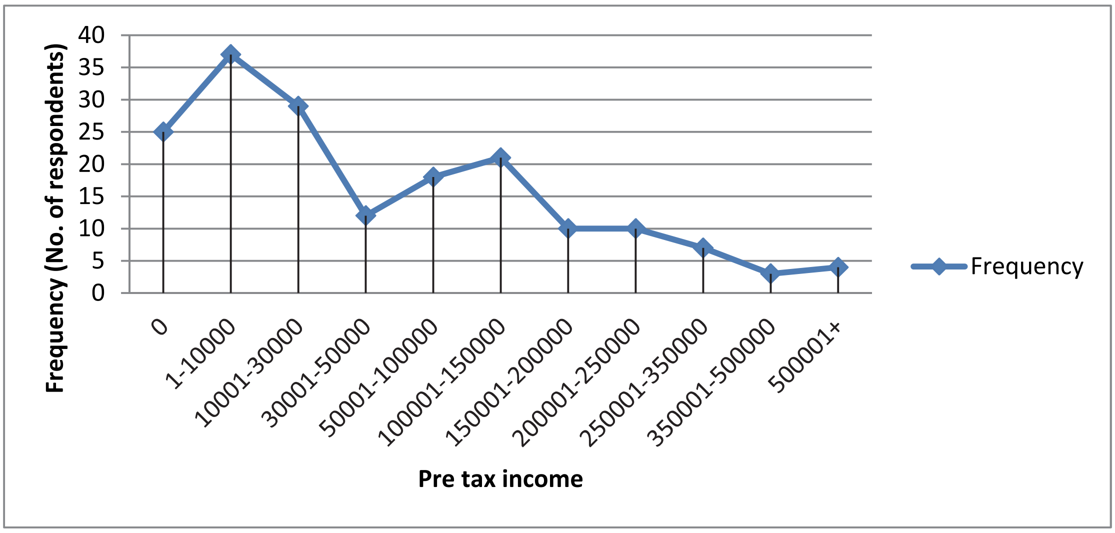

Figure 1.2: Total monthly pre-tax income of Household

The majority of the respondents (75%) thought the local municipality was not doing enough to maintain water quality in the Swartkops estuary. Only 56% felt that the water quality in the estuary was suitable for recreational activities. A large percentage (69%) of respondents also felt that the water quality was not suitable for commercial activities and 56% felt the water quality was suitable for subsistence activities.

The average distance travelled by the respondents from their residences to the estuary was 25 km, with the furthest distance being 1200 km. The main use for the estuary among the respondents was recreation, with 65% of them using the estuary for fishing, boating, etc. 16% for subsistence use and 1% for commercial use. The rest of the respondents (18%) did not use the estuary directly, but were driven by altruistic motives and wanted the estuary to be available for others (children, future generations and other people elsewhere).

Excluding permanent residence (43%), the majority (27%) of the respondents visited the estuary 3 to 4 months in a year, 10% visited it 1 to 2 months in a year, 4% visited it 3 weeks in a year, 4% visited it 2 weeks in a year, 9% visited it 1 week in a year and 3% visited once a year (Figure 1.3)

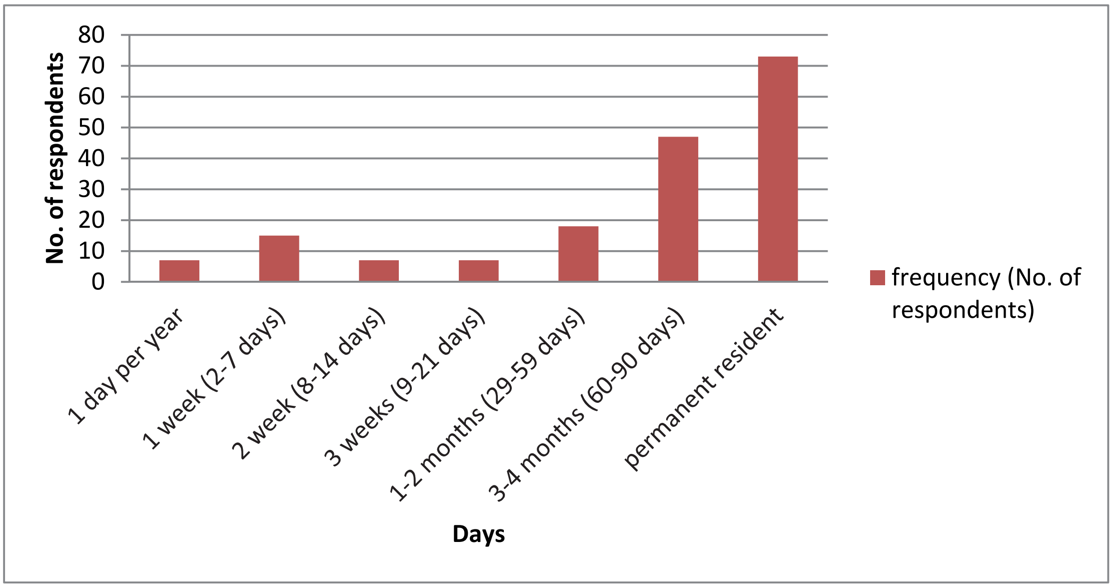

Figure 1.3: Average number of days spent visiting the Swartkops Estuary

The respondents were presented with activities of the Swartkops estuary and asked to rank the relative importance they placed on each activity of the estuary, on a scale of 1 to 4 (1 is extremely important and 4 is unimportant). The question was intended to elicit information on what services the user population of the Swartkops estuary found to be most important. The majority (59%) of respondents felt that being able to use the Swartkops estuary for recreational acitivities was extremely important. Most of them (51%) also felt that commercial activities in the estuary would be harmful to the estuary and therefore were unimportant activities. Of the respondents 47% felt that it was important that those who were underpriviliged be able to use the estuary for subsistence activities. Of the users 81% felt that it was either extermely or very important that the the estuary be accessible for others (Figure 1.4).

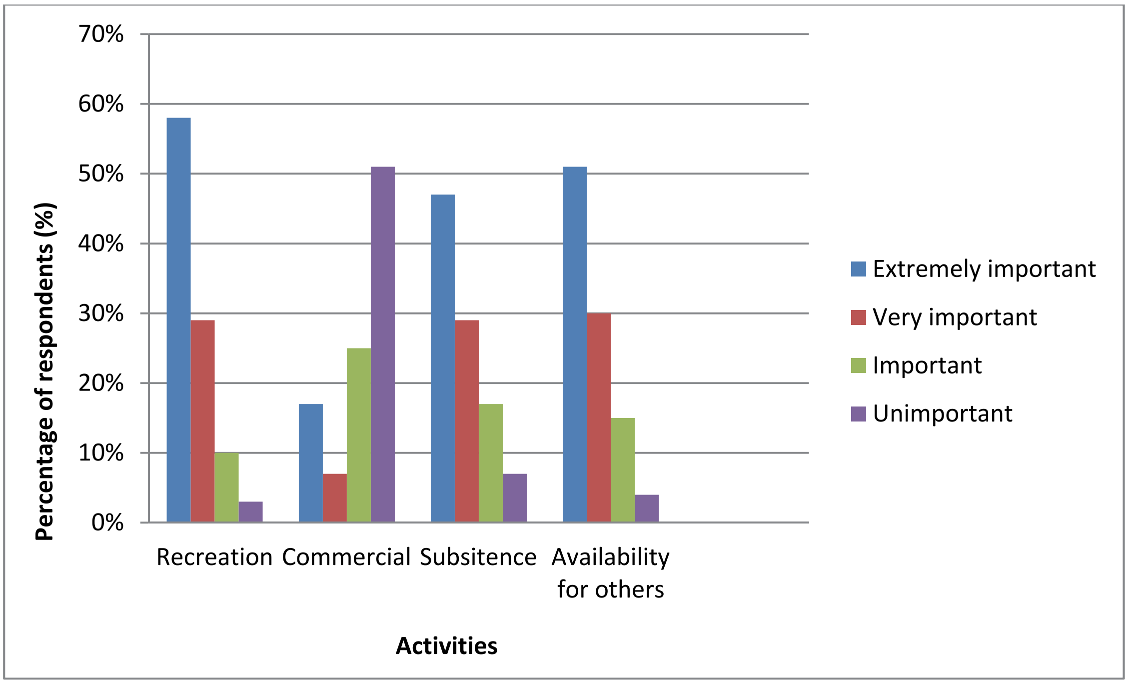

Figure 1.4: The relative importance attached to activities of the estuary

Table 1.2: Percentage of respondents who gave positive WTP responses and zero WTP

Amount WTP per annum	Frequency	Percentage of total (%)
0	59	34%
(1-10)	23	13%
(11-20)	9	5%
(21-30)	2	1%
(31-50)	8	5%
(51-100)	21	12%
(101-200)	21	12%
(201-500)	20	12%
(501-1000)	7	4%
(1001-2000)	0	0%
(2001-3000)	1	0.50%
(3001-4000)	1	0.50%
(4001+)	3	1%
Total	163	100

There were four protest bids. The majority of respondents, who had protest bids, felt that it was the responsibility of the local municipality to maintain water quality in the Swartkops river estuary. They felt that they already paid enough in the form of taxes, rates, fishing and boating licences, etc. and that these resources should be used more efficiently. Only 3 (1%) respondents' had bids in excess of R4001. These were treated as outliers and removed from the mean and median value calculation. The sample mean and median WTP for water quality improvements in the Swartkops estuary was calculated to be R128 and R15 per annum respectively.

Fourteen explanatory variables were identified as likely to influence WTP responses. These were income, age, distance, population group, marital status, gender, employment status (unemployed and retired), visitor or resident, annual usage (some use (1-4 months) and frequent use (permanent resident)), perception of water quality and education level (secondary, and diploma, degree,..., other).

The complete Tobit model could not be estimated as the matrix was ill-conditioned. As a result the complete Tobit model was estimated as a linear model (Table 5.3). Using backwards stepwise approach from this model, the reduced Tobit model was estimated. This model was found to include ln-inc, population group, employment status and annual usage as the only significant variables. The adjusted R-squared was 37% (Table 1.3). All the variables had the expected signs except annual use.

Table 1.3: Reduced Tobit model for the Swartkops estuary

Model	**Reduced Tobit model**			
Dependent variable	**LN_WTP**			
Method	**ML-Censored Normal (TOBIT) (Quadratic hill climbing)**			
Variable	**Coefficient**	**Std. Error**	**z-Statistic**	**Prob.**
C	-0.6341	1.2336	-0.514	0.6072
LN_INC	0.3393	0.0764	4.441	0.0000
POPGR	1.0061	0.4689	2.146	0.0319
EMP2	-1.2105	0.5798	-2.088	0.0368
EMP3	-0.9928	0.7195	-1.380	0.1676
ANNUSE1	-0.8838	0.6169	-1.432	0.1520
ANNUSE2	-1.3849	0.6222	-2.226	0.0260
R-squared	0.3984			
Adjusted R-squared	0.3712			
Log likelihood	-306.00			

The Tobit model is routinely used when the values of the observed dependent variable are exclusively non-negative and are clustered at zero and is also preferred over OLS model for predictive purposes. The WTP bid function for the Swartkops estuary user population was found to be:

Equation 1.1:

Ln-WTP = -0.6341 + 0.3393 * ln-inc + 1.0061*popgr -1.2105*EMP2 - 0.9928*EMP3 - 0.8838* ANNUSE1 - 1.3849*ANNUSE2.. (1.1)

The interpretation of the regression coefficients change if they are transformed to log-linear functional form. After transformation, the coefficient of the explanatory variable represents the percentage contribution that the explanatory variable makes to the dependant variable when there is a one unit increment in the independent variable (Mendenhall & Sincich 1996; Hill, Griffiths, Lim 2008). The percentage change is obtained by transforming the coefficient (Mendenhall & Sincich 1996):

Percentage change = ($e^{\beta i}$ – 1) x 100 .... (1.2)

Table 1.4: Coefficient interpretation

		Model	
Variable	**Interpretation**	**OLS**	**Tobit**
Ln-inc	A 1% increase in income will result in a percentage increase in WTP	0.18%	0.33%
Popgrp	The percentage increase in WTP due to the respondent being white.	150.85%	173.49%
EMP2 (unemployed)	The percentage decrease in WTP due to the respondent being unemployed	-68.31%	-70.19%
EMP3 (retired)	The percentage decrease in WTP due to the respondent being unemployed	-62.15%	-62.94%

ANNUSE1 (infrequent use, <= 21days)	The percentage decrease in WTP due to a change (increase) in number of days spent at the estuary	-47.10%	-58.67%
ANNUSE2 (some use, 1-4 months)	The percentage decrease in WTP due to a change (increase) in number of days spent at the estuary	-64.53%	-74.96%

In the two models (OLS and Tobit) population group and employment status both have signs that are consistent with expectations, while annual use has a negative relationship with WTP, contrary to what was expected. It was expected that respondents' WTP would be positively related to the number of days spent at the estuary, but this was not the case. The result suggests that the greater the number of days spent at the estuary annually leads to a decrease in WTP.

A plausible explanation for this result is that the majority of the respondents sensitive to the water quality problem (therefore theoretically inclined to have a higher WTP) no longer use the estuary and as a result were not interviewed. The questions that remain to be answered are: (1) Where were these respondents? (2) Why was there a negative relationship between number of days spent at the estuary and WTP for the respondents that remained using the estuary?

The majority of users who used the estuary frequently and whose use thereof was affected by water quality over the years opted to find substitutes to the Swartkops estuary, e.g. Sunday River. The remaining users were those whose activities at the estuary were not affected as much by water quality, or had no other option but to use the estuary in its condition. Their use adapted or accommodated the deteriorated water quality over the years. As a result improving and maintaining river water quality was not of sufficient importance to warrant paying any money, because they had grown accustomed to the poor water quality over the years.

Theoretically the number of days spent at an estuary should be positively related to WTP for water quality improvement. A user who has a high demand for the estuary should be willing to pay more to improve water quality in the estuary. However, when water quality does not matter (adapted users) or is not of sufficient importance, the number of days spent at the estuary does not necessarily mean that their WTP should be higher for a water quality improvement project. The respondents who use the Swartkops estuary more per annum have adapted to the poor water quality.

When water quality perception is an insignificant variable (reduced Tobit model) in determining WTP for a water quality improvement project, it should be of no surprise that the number of days spent at the estuary may not be positively correlated to WTP for a water quality improvement project. This explains the negative relationship that annual use has with WTP.

Using the WTP bid function (equation 1.2), the predicted mean and median WTP values for the Swartkops estuary are R39.19 and R13.24 respectively

The predicted bids were converted to a population bid (TWTP) by multiplying the predicted mean or median WTP bid per annum by the number of user households. Given the preference for conservative values (see Chapter three), it was decided to use the predicted median bid to calculate the total population WTP. TWTP per annum for the implementation of a proposed project to improve water quality to safe levels for swimming, fishing and boating in the Swartkops estuary was calculated to be R68848. Using the mean bid in the calculation increases TWTP to R203632 per annum. Using the broader definition of target population TWTP per annum was calculated to be R3481987 (median bid) and R10298688 (mean bid) per annum.

Validity assessment

The contingent valuation method is an imperfect technique. Perfection does not exist, even in actual markets (Boyle 2003: 42), but in terms of content validity requirements (complying with the majority of the NOAA Blue-Ribbon panel guidelines, 1993) it is argued the Swartkops estuary has a model that moderately supports construct validity and a CV questionnaire that meets the elusive criterion of “working”.

1.3 Conclusion

The Swartkops estuary is a valuable recreational and ecological asset, but has been subject to negative water quality effects and influences (Binning & Baird 2001), thereby undermining some of its recreational, subsistence and environmental appeal (Roux, De Lange and Oelofse 2010). Chapter One has demonstrated, with the aid of the CVM, that there is a demand for better sewage management than is currently provided, in order to improve water quality in the Swartkops estuary to safe levels for swimming, fishing and boating. The contingent valuation method is suited to this kind of demonstration.

Using the predictive function specified in Equation 1.2, the mean and median WTP for the Swartkops estuary were predicted to be R39.19 and R13.24 respectively, and the TWTP per annum of a narrowly defined user population of the Swartkops estuary was calculated at R68848 (median bid) and R203632 (mean bid). This result confirms the hypothesis, that improved freshwater quality flowing into the Swartkops estuary has a positive economic value. Using a more broadly defined interested population TWTP per annum was calculated to be R3481987 (median bid) and R10298688 (mean bid).

This finding provides a case study demand-side motivation for improving the efficiency of the waste water service management provided by the Nelson Mandela Bay municipality. The public are prepared to pay extra to prevent inadequate water service provision that cause contamination of their estuaries, but it still occurs. The public are prepared to pay for improved efficiency in water service provision, but its occurrence needs to be highlighted, so remedial action is undertaken by the municipality. One way of highlighting it is through monitoring efficiency performance and publicly reporting on it.

APPENDIX B: The Survey instrument for Appendix A

No:_____

Date of interview:________

Name of Respondent:__________________________

Questionnaire to Assess Water Quality Improvements: Contingent Valuation for a water quality improvement project for the Swartkops River

Part 1: Socio-Economic and Household Characteristics

1. Respondents gender

____ 1. Male

____ 2. Female

2. Race of Respondent

____ 1. Black

____ 2. White

____ 3. Coloured

____ 4. Indian

____ 5. Other

3. Age …… years

4. Marital Status

____ 1. Single

____ 2. Married

____ 3. Widowed/Divorced/Separated

5. Education

_____ 1. None

_____ 2. Primary

_____ 3. Secondary

_____ 4. Technical, Diploma

_____ 5. University Degree

_____ 6. Post Graduate Degree

_____ 7.Others (please specify)…………..…………………….

6. Employment Status

____ 1. Employed

____ 2. Unemployed

____ 3. Retired

7. How many people make up your household? ________

8. Total annual pre tax income of Household

Pre Tax Income (Rands)	
0	0
1- 10 000	1
10 001-30 000	2
30 001-50 000	3
50 001-100 000	4
100 001-150 000	5
150 001-200 000	6
200 001-250 000	7
250 001-350 000	8
350 001-500 000	9
500 001+	10

9. Visitor or Resident?

____ 1. Visitor

____ 2. Permanent resident

____ 3. Holiday resident

Part 2: Attitudinal section

10. Do you think the water quality in the Swartkops estuary is suitable for the following activities?

Recreational

____ 1. Yes

_____ 2. No

Commercial

_____ 1. Yes

_____ 2. No

Subsistence

_____ 1. Yes

_____ 2. No

11. Do you think the local municipality are doing enough to maintain the water quality in the Swartkops estuary?

_____ 1. Yes

_____ 2. No

Part 3: Characterisation of behaviour

12. On average how often do you visit the Swartkops estuary per year?

Days	
1 day per year	1
1 week (2-7 days)	2
2 weeks (8-14 days)	3
3 weeks (15-21 days)	4

1-2 months (29-59 days)	5
3-4 months (60-90 days)	6
Permanent resident	7

13 Indicate the distance from your place of residence to the Swartkops River estuary.

13.1 About ……………..km.

14. What is the main purpose of your use of the estuary?

____ 1. Recreation (Fishing, boating, swimming, bird watching, relaxing and enjoying the scenery)

____ 2. Commercial (fishing and bait collection)

____ 3. Subsistence (bait collecting, fishing, employment from recreation users)

____ 4. Availability for others (ones children, future generations and other people elsewhere)

15.1 Rank the relative importance you attach to the following activities of the estuary on a scale of 1-4 (1 is extremely important and 4 is unimportant)

Activity	Extremely important	Very important	important	unimportant
Recreation	1	2	3	4
Commercial	1	2	3	4
Subsistence	1	2	3	4
Availability for others	1	2	3	4

Part 4: Contingent Valuation Section

The Swartkops River passes through a highly urbanised and industrialised region of the Eastern Cape and forms an important part of Port Elizabeth and areas surrounding it. It is a valuable recreational and ecological asset. As a result of rapidly expanding urban areas, the water quality of the Swartkops estuary is critically impacted by several effects and influences of these developments. The sources of impact include: agriculture, wool processing factory, three sewerage treatment works, and runoff from informal settlements and discharges from light industries. The Swartkops is one of the most threatened freshwater systems in South Africa.

Water quality in the Swartkops estuary among several researchers has been rated as good. But as recently as May 2010 a warning was issued to the public not to swim or fish in the Swartkops River as water quality had deteriorated as a result of broken sewage pipes and stormwater drains. The Redhouse River mile as a consequence was moved due to dangerously high levels of pollution. This research proposes a project to improve and sustain the Swartkops estuary water quality rating, thus bringing the water quality of the Swartkops estuary to swimmable and fishable levels. This improvement will involve increasing the effectiveness of Nelson Mandela Municipality Wastewater Treatment Works along the Swartkops River and replacing of faulty joints and pipes in the system. This project will be financed through higher waste water charges levied on all water service users in the Nelson Mandela Bay Municipality.

Valuation Question:

In answering the valuation question it is important that you consider all your existing financial obligations before making your decision, and that you answer truthfully as possible. Please also remember that you should take into account not only your current use, but also yours and your families' potential use as well as anyone else's that is important to you. It is also important to remember that there are other alternatives to the Swarkops estuary; such as the Sunday river estuary within proximity.

16. How much are you willing to pay for the implementation of a proposed project to improve water quality to safe levels for swimming, fishing and boating.

Amount willing to pay (Rand)	
0	0
0-10	1
11- 20	2
21-30	3
31-50	4
51-100	5
101-200	6
201-500	7
501-1000	8
1001-2000	9
2001-3000	10
3001-4000	11
4001+ (specify)	12

16.1 If your answer was 0, what are your reasons?

____ 1. I do not have enough money to pay for river water quality improvement and maintenance.

____ 2.I do not believe that improving and maintaining river water quality was of sufficient importance to warrant paying any money

____ 3 I did not believe the scenario

____ 4 I object to having to pay money for maintaining water quality in rivers

___ 5 Other reason not reflected in the questionnaire

Do you have any other comments you would like to contribute on this public issue?

____

____

____

APPENDIX C: THE DATA USED IN THE FRONTIER ANALYSIS ESTIMATION

Table 4.7: Analysis of connections of Municipalities

Municipality	Total Connections	Connections (metered)	Connections (unmetered)
NELSON MANDELA BAY METROPOLITAN MUNICIPALITY	227042	217791	9251
CACADU DISTRICT MUNICIPALITY	113831	97375	16456
CAMDEBOO LOCAL MUNICIPALITY	11287	10672	615
MAKANA LOCAL MUNICIPALITY	22371	20326	2045
NDLAMBE LOCAL MUNICIPALITY	17289	14267	3022
BAVIAANS LOCAL MUNICIPALITY	4159	3871	288
KOUGA LOCAL MUNICIPALITY	25602	18125	7477
AMATHOLE DISTRICT MUNICIPALITY	228555	170746	57809
MBHASHE LOCAL MUNICIPALITY	860	805	55
MNQUMA LOCAL MUNICIPALITY	2600	1470	1130
GREAT KEI LOCAL MUNICIPALITY	2703	2379	324
AMAHLATHI LOCAL MUNICIPALITY	5332	3242	2090
BUFFALO CITY LOCAL MUNICIPALITY	195479	153566	41913
NGQUSHWA LOCAL MUNICIPALITY	1467	741	726
NKONKOBE LOCAL MUNICIPALITY	15517	4546	10971
O.R. TAMBO DISTRICT MUNICIPALITY	334138	14940	319198
MBIZANA LOCAL MUNICIPALITY	44841	250	44591
NTABANKULU LOCAL MUNICIPALITY	26659	134	26525
NGQUZA HILL LOCAL MUNICIPALITY	49485	601	48884
PORT ST JOHNS LOCAL MUNICIPALITY	28077	169	27908
NYANDENI LOCAL MUNICIPALITY	53458	290	53168
MHLONTLO LOCAL MUNICIPALITY	42176	153	42023
KING SABATA DALINDYEBO LOCAL MUNICIPALITY	89442	13343	76099
XHARIEP DISTRICT MUNICIPALITY	23809	20713	3096
LETSEMENG LOCAL MUNICIPALITY	13750	11902	1848
Dr RUTH SEGOMOTSI MOMPATI DISTRICT MUNICIPALITY	191435	135496	55939
MANGAUNG LOCAL MUNICIPALITY	167954	115727	52227
MANTSOPA LOCAL MUNICIPALITY	14725	12081	2644
LEJWELEPUTSWA DISTRICT MUNICIPALITY	166210	132062	34148
TSWELOPELE LOCAL MUNICIPALITY	13169	8916	4253
MATJHABENG LOCAL MUNICIPALITY	94279	72225	22054

Municipality	Total Connections	Connections (metered)	Connections (unmetered)
THABO MOFUTSANYANE DISTRICT MUNICIPALITY	140733	91044	49689
SETSOTO LOCAL MUNICIPALITY	26429	3504	22925
DIHLABENG LOCAL MUNICIPALITY	0		
FEZILE DABI DISTRICT MUNICIPALITY	26023	12137	13886
MOQHAKA LOCAL MUNICIPALITY	46082	40339	5743
NGWATHE LOCAL MUNICIPALITY	34406	31697	2709
METSIMAHOLO LOCAL MUNICIPALITY	36680	32250	4430
MAFUBE LOCAL MUNICIPALITY	16853	15850	1003
CITY OF JOHANNESBURG METROPOLITAN MUNICIPALITY	559367	451951	107416
CITY OF TSHWANE METROPOLITAN MUNICIPALITY	423414	423414	0
EKURHULENI METROPOLITAN MUNICIPALITY	497570	430462	67108
SEDIBENG DISTRICT MUNICIPALITY	277069	265510	11559
EMFULENI LOCAL MUNICIPALITY	229610	220255	9355
MIDVAAL LOCAL MUNICIPALITY	24776	22573	2203
LESEDI LOCAL MUNICIPALITY	22682	22682	0
METSWEDING DISTRICT MUNICIPALITY	59567	54334	5233
NOKENG TSA TAEMANE LOCAL MUNICIPALITY	18064	16520	1544
KUNGWINI LOCAL MUNICIPALITY	41503	37814	3689
WEST RAND DISTRICT MUNICIPALITY	1918546	1897319	21227
MOGALE CITY LOCAL MUNICIPALITY	60489	57609	2880
RANDFONTEIN LOCAL MUNICIPALITY	45521	42860	2661
WESTONARIA LOCAL MUNICIPALITY	34813	28220	6593
MERAFONG CITY LOCAL MUNICIPALITY	67723	58630	9093
ETHEKWINI METROPOLITAN MUNICIPALITY	442721	442721	0
UGU DISTRICT MUNICIPALITY	38967	35939	3028
UMDONI LOCAL MUNICIPALITY	8411	8013	398
UMZUMBE LOCAL MUNICIPALITY	2143	1593	550
EZINQOLENI LOCAL MUNICIPALITY	133	63	70
HIBISCUS COAST LOCAL MUNICIPALITY	26447	24879	1568
UMGUNGUNDHLOVU DISTRICT MUNICIPALITY	162446	142730	19716
UTHUKELA DISTRICT MUNICIPALITY	151180	35479	115701

Municipality	Total Connections	Connections (metered)	Connections (unmetered)
EMNAMBITHI-LADYSMITH LOCAL MUNICIPALITY	90708	21287	69421
INDAKA LOCAL MUNICIPALITY	15118	3548	11570
UMTSHEZI LOCAL MUNICIPALITY	30236	7096	23140
NTAMBANANA LOCAL MUNICIPALITY	9071	2129	6942
IMBABAZANE LOCAL MUNICIPALITY	6047	1419	4628
UTHUNGULU DISTRICT MUNICIPALITY	108919	63006	45913
UMHLATHUZE LOCAL MUNICIPALITY	24213	24213	
CAPRICORN DISTRICT MUNICIPALITY	246428	151209	95219
WATERBERG DISTRICT MUNICIPALITY	103955	75833	28122
MOGALAKWENA LOCAL MUNICIPALITY	18089	16129	1960
GERT SIBANDE DISTRICT MUNICIPALITY	214972	175717	39255
MSUKALIGWA LOCAL MUNICIPALITY	19963	19963	0
GOVAN MBEKI LOCAL MUNICIPALITY	69954	58429	11525
NKANGALA DISTRICT MUNICIPALITY	276681	241455	35226
STEVE TSHWETE LOCAL MUNICIPALITY	42222	36465	5757
DR JS MOROKA LOCAL MUNICIPALITY	52373	43626	8747
EHLANZENI DISTRICT MUNICIPALITY	270429	167186	103243
MBOMBELA LOCAL MUNICIPALITY	79876	39775	40101
BOJANALA PLATINUM DISTRICT MUNICIPALITY	317563	240172	77391
RUSTENBURG LOCAL MUNICIPALITY	134983	108316	26667
NGAKA MODIRI MOLEMA DISTRICT MUNICIPALITY	128256	75079	53177
MAFIKENG LOCAL MUNICIPALITY	40321	27933	12388
RAMOTSHERE MOILOA LOCAL MUNICIPALITY	23387	10208	13179
Dr KENNETH KAUNDA DISTRICT MUNICIPALITY	177703	157603	20100
TLOKWE LOCAL MUNICIPALITY	37758	33634	4124
NAMAKWA DISTRICT MUNICIPALITY	30889	29930	959
RICHTERSVELD LOCAL MUNICIPALITY	2758	2713	45
NAMA KHOI LOCAL MUNICIPALITY	12174	11870	304
HANTAM LOCAL MUNICIPALITY	6940	6940	0
KAROO HOOGLAND LOCAL MUNICIPALITY	2758	2504	254
KHAI-MA LOCAL MUNICIPALITY	3136	3112	24

Municipality	Total Connections	Connections (metered)	Connections (unmetered)
PIXLEY KA SEME DISTRICT MUNICIPALITY	20952	20028	924
THEMBELIHLE LOCAL MUNICIPALITY	3159	3008	151
SIYATHEMBA LOCAL MUNICIPALITY	3159	3008	151
FRANCES BAARD DISTRICT MUNICIPALITY	9477	9024	453
SOL PLAATJIE LOCAL MUNICIPALITY	3159	3008	151
JOHN TAOLO GAETSEWE DISTRICT MUNICIPALITY	16528	14516	2012
GA-SEGONYANA LOCAL MUNICIPALITY	3159	3008	151
WEST COAST DISTRICT MUNICIPALITY	72666	71016	1650
CEDERBERG LOCAL MUNICIPALITY	13098	12606	492
CAPE WINELANDS DISTRICT MUNICIPALITY	167542	158158	9384
DRAKENSTEIN LOCAL MUNICIPALITY	54025	51652	2373
EDEN DISTRICT MUNICIPALITY	453952	445924	8028
CENTRAL KAROO DISTRICT MUNICIPALITY	15178	14906	272

Table 4.8: Volumes of water delivered and lengths of mains

Municipality	Length of Mains	System input Volume
NELSON MANDELA BAY METROPOLITAN MUNICIPALITY	4541	94036000
CACADU DISTRICT MUNICIPALITY	2277	18217682
CAMDEBOO LOCAL MUNICIPALITY	226	6590000
MAKANA LOCAL MUNICIPALITY	447	2040000
NDLAMBE LOCAL MUNICIPALITY	346	3368070
BAVIAANS LOCAL MUNICIPALITY	83	619612
KOUGA LOCAL MUNICIPALITY	512	5600000
AMATHOLE DISTRICT MUNICIPALITY	1927668	79193991
MBHASHE LOCAL MUNICIPALITY	86	1060093
MNQUMA LOCAL MUNICIPALITY	1183	4960589
GREAT KEI LOCAL MUNICIPALITY	556251	713283
AMAHLATHI LOCAL MUNICIPALITY	361536	1377401
BUFFALO CITY LOCAL MUNICIPALITY	3910	62652039
NGQUSHWA LOCAL MUNICIPALITY	572960	4721746
NKONKOBE LOCAL MUNICIPALITY	431607	3208991
O.R. TAMBO DISTRICT MUNICIPALITY	1823	71832332
MBIZANA LOCAL MUNICIPALITY	175	525210

Municipality	Length of Mains	System input Volume
NTABANKULU LOCAL MUNICIPALITY	192	507740
NGQUZA HILL LOCAL MUNICIPALITY	177	1201456
PORT ST JOHNS LOCAL MUNICIPALITY	111	796750
NYANDENI LOCAL MUNICIPALITY	358	683400
MHLONTLO LOCAL MUNICIPALITY	358	682640
KING SABATA DALINDYEBO LOCAL MUNICIPALITY	453	67435136
XHARIEP DISTRICT MUNICIPALITY	476	6301728
LETSEMENG LOCAL MUNICIPALITY	275	1895728
Dr RUTH SEGOMOTSI MOMPATI DISTRICT MUNICIPALITY	3704	84358852
MANGAUNG LOCAL MUNICIPALITY	3401	79085845
MANTSOPA LOCAL MUNICIPALITY	128	5273007
LEJWELEPUTSWA DISTRICT MUNICIPALITY	3009	37168428
TSWELOPELE LOCAL MUNICIPALITY	263	2863509
MATJHABENG LOCAL MUNICIPALITY	1571	34304919
THABO MOFUTSANYANE DISTRICT MUNICIPALITY	2666	54490905
SETSOTO LOCAL MUNICIPALITY	380	8560905
DIHLABENG LOCAL MUNICIPALITY		10000000
FEZILE DABI DISTRICT MUNICIPALITY	2680	37083069
MOQHAKA LOCAL MUNICIPALITY	922	11744810
NGWATHE LOCAL MUNICIPALITY	688	8304317
METSIMAHOLO LOCAL MUNICIPALITY	734	15060000
MAFUBE LOCAL MUNICIPALITY	337	1973942
CITY OF JOHANNESBURG METROPOLITAN MUNICIPALITY	11296	502956153
CITY OF TSHWANE METROPOLITAN MUNICIPALITY	9141	265845467
EKURHULENI METROPOLITAN MUNICIPALITY	10077	322249616
SEDIBENG DISTRICT MUNICIPALITY	5393	97107804
EMFULENI LOCAL MUNICIPALITY	4592	79559382
MIDVAAL LOCAL MUNICIPALITY	496	12053525
LESEDI LOCAL MUNICIPALITY	304	5494897
METSWEDING DISTRICT MUNICIPALITY	1191	26352214
NOKENG TSA TAEMANE LOCAL MUNICIPALITY	361	2162214
KUNGWINI LOCAL MUNICIPALITY	830	24190000
WEST RAND DISTRICT MUNICIPALITY	3646	52690678
MOGALE CITY LOCAL MUNICIPALITY	685	27257044
RANDFONTEIN LOCAL MUNICIPALITY	910	8740000
WESTONARIA LOCAL MUNICIPALITY	696	6404434

Municipality	Length of Mains	System input Volume
MERAFONG CITY LOCAL MUNICIPALITY	1354	10289200
ETHEKWINI METROPOLITAN MUNICIPALITY	11643	332941393
UGU DISTRICT MUNICIPALITY	3883	32092567
UMDONI LOCAL MUNICIPALITY	600	4577570
UMZUMBE LOCAL MUNICIPALITY	351	3044395
EZINQOLENI LOCAL MUNICIPALITY	51	401749
HIBISCUS COAST LOCAL MUNICIPALITY	2242	22215446
UMGUNGUNDHLOVU DISTRICT MUNICIPALITY	1655	58492009
UTHUKELA DISTRICT MUNICIPALITY	3024	43889500
EMNAMBITHI-LADYSMITH LOCAL MUNICIPALITY	1814	26462500
INDAKA LOCAL MUNICIPALITY	302	4197500
UMTSHEZI LOCAL MUNICIPALITY	605	8942500
NTAMBANANA LOCAL MUNICIPALITY	181	2644500
IMBABAZANE LOCAL MUNICIPALITY	121	1642500
UTHUNGULU DISTRICT MUNICIPALITY	1839	42557901
UMHLATHUZE LOCAL MUNICIPALITY	145	42557901
CAPRICORN DISTRICT MUNICIPALITY	3684	33555165
WATERBERG DISTRICT MUNICIPALITY	513608	21072102
MOGALAKWENA LOCAL MUNICIPALITY	152	9202102
GERT SIBANDE DISTRICT MUNICIPALITY	4314	29676453
MSUKALIGWA LOCAL MUNICIPALITY	413	3751034
GOVAN MBEKI LOCAL MUNICIPALITY	1399	22240419
NKANGALA DISTRICT MUNICIPALITY	4584	93565714
STEVE TSHWETE LOCAL MUNICIPALITY	717	15602431
DR JS MOROKA LOCAL MUNICIPALITY	225	23400000
EHLANZENI DISTRICT MUNICIPALITY	4701	38133933
MBOMBELA LOCAL MUNICIPALITY	1005	29099341
BOJANALA PLATINUM DISTRICT MUNICIPALITY	6351	57121439
RUSTENBURG LOCAL MUNICIPALITY	2700	35581439
NGAKA MODIRI MOLEMA DISTRICT MUNICIPALITY	2565	17339376
MAFIKENG LOCAL MUNICIPALITY	806	10928631
RAMOTSHERE MOILOA LOCAL MUNICIPALITY	468	6410746
Dr KENNETH KAUNDA DISTRICT MUNICIPALITY	3554	45475577
TLOKWE LOCAL MUNICIPALITY	755	13365577
NAMAKWA DISTRICT MUNICIPALITY	268	6260177
RICHTERSVELD LOCAL MUNICIPALITY	55	364443

Municipality	Length of Mains	System input Volume
NAMA KHOI LOCAL MUNICIPALITY	243	3469823
HANTAM LOCAL MUNICIPALITY	149	831408
KAROO HOOGLAND LOCAL MUNICIPALITY	55	803102
KHAI-MA LOCAL MUNICIPALITY	63	791401
PIXLEY KA SEME DISTRICT MUNICIPALITY	598	7313606
THEMBELIHLE LOCAL MUNICIPALITY	63	1353678
SIYATHEMBA LOCAL MUNICIPALITY	63	1611075
FRANCES BAARD DISTRICT MUNICIPALITY	189	32602015
SOL PLAATJIE LOCAL MUNICIPALITY	63	32602015
JOHN TAOLO GAETSEWE DISTRICT MUNICIPALITY	314	7086180
GA-SEGONYANA LOCAL MUNICIPALITY	63	871615
WEST COAST DISTRICT MUNICIPALITY	1871	20417029
CEDERBERG LOCAL MUNICIPALITY	262	2611029
CAPE WINELANDS DISTRICT MUNICIPALITY	3298	30807514
DRAKENSTEIN LOCAL MUNICIPALITY	1080	15975000
EDEN DISTRICT MUNICIPALITY	3238	23383245
CENTRAL KAROO DISTRICT MUNICIPALITY	280	3611000

Table 4.9: Analysis of billed consumption

Municipality	Total billed consumption	Total unbilled consumption	Billed metred consumption
NELSON MANDELA BAY METROPOLITAN MUNICIPALITY	58914000	35122000	56414000
CACADU DISTRICT MUNICIPALITY	14422867	3794815	12582867
CAMDEBOO LOCAL MUNICIPALITY	4260000	2330000	2620000
MAKANA LOCAL MUNICIPALITY	1780000	260000	1580000
NDLAMBE LOCAL MUNICIPALITY	3233899	134171	3233899
BAVIAANS LOCAL MUNICIPALITY	448967	170645	448967
KOUGA LOCAL MUNICIPALITY	4700000	900000	4700000
AMATHOLE DISTRICT MUNICIPALITY	43875176	35318815	40095776
MBHASHE LOCAL MUNICIPALITY	603808	456285	603808
MNQUMA LOCAL MUNICIPALITY	1663828	3296761	1663828
GREAT KEI LOCAL MUNICIPALITY	527000	186283	527000
AMAHLATHI LOCAL MUNICIPALITY	11155044	-9777643	11155044

Municipality	Total billed consumption	Total unbilled consumption	Billed metred consumption
BUFFALO CITY LOCAL MUNICIPALITY	37620667	25031372	33841267
NGQUSHWA LOCAL MUNICIPALITY	359829	4361917	359829
NKONKOBE LOCAL MUNICIPALITY	1985000	1223991	1985000
O.R. TAMBO DISTRICT MUNICIPALITY	68948919	2883413	68948919
MBIZANA LOCAL MUNICIPALITY	475580	49630	475580
NTABANKULU LOCAL MUNICIPALITY	458230	49510	458230
NGQUZA HILL LOCAL MUNICIPALITY	1023624	177832	1023624
PORT ST JOHNS LOCAL MUNICIPALITY	771820	24930	771820
NYANDENI LOCAL MUNICIPALITY	673120	10280	673120
MHLONTLO LOCAL MUNICIPALITY	611850	70790	611850
KING SABATA DALINDYEBO LOCAL MUNICIPALITY	64934695	2500441	64934695
XHARIEP DISTRICT MUNICIPALITY	2987337	3314391	2263337
LETSEMENG LOCAL MUNICIPALITY	1007337	888391	1007337
Dr RUTH SEGOMOTSI MOMPATI DISTRICT MUNICIPALITY	42380737	41978115	42364873
MANGAUNG LOCAL MUNICIPALITY	39305132	39780713	39305132
MANTSOPA LOCAL MUNICIPALITY	3075605	2197402	3059741
LEJWELEPUTSWA DISTRICT MUNICIPALITY	24331515	12836913	24331515
TSWELOPELE LOCAL MUNICIPALITY	2404671	458838	2404671
MATJHABENG LOCAL MUNICIPALITY	21926844	12378075	21926844
THABO MOFUTSANYANE DISTRICT MUNICIPALITY	22253893	32237012	22253893
SETSOTO LOCAL MUNICIPALITY	4263893	4297012	4263893
DIHLABENG LOCAL MUNICIPALITY	4000000	6000000	4000000
FEZILE DABI DISTRICT MUNICIPALITY	28384734	8698335	28384734
MOQHAKA LOCAL MUNICIPALITY	8297246	3447564	8297246
NGWATHE LOCAL MUNICIPALITY	6240000	2064317	6240000
METSIMAHOLO LOCAL MUNICIPALITY	12476000	2584000	12476000
MAFUBE LOCAL MUNICIPALITY	1371488	602454	1371488
CITY OF JOHANNESBURG METROPOLITAN MUNICIPALITY	310677660	192278493	289415013
CITY OF TSHWANE METROPOLITAN MUNICIPALITY	195426657	70418810	195426657
EKURHULENI METROPOLITAN MUNICIPALITY	193973397	128276219	193973397
SEDIBENG DISTRICT MUNICIPALITY	58187727	38920077	38057957

Municipality	Total billed consumption	Total unbilled consumption	Billed metred consumption
EMFULENI LOCAL MUNICIPALITY	44248813	35310569	24120973
MIDVAAL LOCAL MUNICIPALITY	8896469	3157056	8896469
LESEDI LOCAL MUNICIPALITY	5042445	452452	5040515
METSWEDING DISTRICT MUNICIPALITY	15400000	10952214	15400000
NOKENG TSA TAEMANE LOCAL MUNICIPALITY	1600000	562214	1600000
KUNGWINI LOCAL MUNICIPALITY	13800000	10390000	13800000
WEST RAND DISTRICT MUNICIPALITY	39123606	13567072	36349114
MOGALE CITY LOCAL MUNICIPALITY	20180070	7076974	19219114
RANDFONTEIN LOCAL MUNICIPALITY	6830000	1910000	6830000
WESTONARIA LOCAL MUNICIPALITY	4500000	1904434	4500000
MERAFONG CITY LOCAL MUNICIPALITY	7613536	2675664	5800000
ETHEKWINI METROPOLITAN MUNICIPALITY	208119455	124821938	208119455
UGU DISTRICT MUNICIPALITY	20191068	11901499	16399764
UMDONI LOCAL MUNICIPALITY	4212280	365290	3932681
UMZUMBE LOCAL MUNICIPALITY	737059	2307336	513629
EZINQOLENI LOCAL MUNICIPALITY	375176	26573	309749
HIBISCUS COAST LOCAL MUNICIPALITY	13538259	8677187	11252583
UMGUNGUNDHLOVU DISTRICT MUNICIPALITY	21653346	36838663	21653346
UTHUKELA DISTRICT MUNICIPALITY	14016127	29873373	14016127
EMNAMBITHI-LADYSMITH LOCAL MUNICIPALITY	8409676	18052824	8409676
INDAKA LOCAL MUNICIPALITY	1401613	2795887	1401613
UMTSHEZI LOCAL MUNICIPALITY	2803225	6139275	2803225
NTAMBANANA LOCAL MUNICIPALITY	840968	1803532	840968
IMBABAZANE LOCAL MUNICIPALITY	560645	1081855	560645
UTHUNGULU DISTRICT MUNICIPALITY	31059590	11498311	31059590
UMHLATHUZE LOCAL MUNICIPALITY	31059590	11498311	31059590
CAPRICORN DISTRICT MUNICIPALITY	19523972	14031193	19523972
WATERBERG DISTRICT MUNICIPALITY	15267840	5804262	13794540
MOGALAKWENA LOCAL MUNICIPALITY	5625640	3576462	5450340
GERT SIBANDE DISTRICT MUNICIPALITY	21460297	8216156	21460297
MSUKALIGWA LOCAL MUNICIPALITY	2993141	757893	2993141
GOVAN MBEKI LOCAL MUNICIPALITY	16617156	5623263	16617156
NKANGALA DISTRICT MUNICIPALITY	43936013	49629701	41314653
STEVE TSHWETE LOCAL MUNICIPALITY	10325501	5276930	10325501
DR JS MOROKA LOCAL MUNICIPALITY	6572166	16827834	6335806

Municipality	Total billed consumption	Total unbilled consumption	Billed metred consumption
EHLANZENI DISTRICT MUNICIPALITY	23818742	14315191	19025745
MBOMBELA LOCAL MUNICIPALITY	17044393	12054948	13715396
BOJANALA PLATINUM DISTRICT MUNICIPALITY	35237394	21884045	35237394
RUSTENBURG LOCAL MUNICIPALITY	21697394	13884045	21697394
NGAKA MODIRI MOLEMA DISTRICT MUNICIPALITY	11851892	5487484	11851892
MAFIKENG LOCAL MUNICIPALITY	8588629	2340002	8588629
RAMOTSHERE MOILOA LOCAL MUNICIPALITY	3263263	3147483	3263263
Dr KENNETH KAUNDA DISTRICT MUNICIPALITY	37179005	8296572	37179005
TLOKWE LOCAL MUNICIPALITY	12729005	636572	12729005
NAMAKWA DISTRICT MUNICIPALITY	3895266	2364911	3738339
RICHTERSVELD LOCAL MUNICIPALITY	220606	143837	220606
NAMA KHOI LOCAL MUNICIPALITY	1690736	1779087	1690736
HANTAM LOCAL MUNICIPALITY	685613	145795	685613
KAROO HOOGLAND LOCAL MUNICIPALITY	656049	147053	656049
KHAI-MA LOCAL MUNICIPALITY	642262	149139	485335
PIXLEY KA SEME DISTRICT MUNICIPALITY	5821019	1492587	5821019
THEMBELIHLE LOCAL MUNICIPALITY	968288	385390	968288
SIYATHEMBA LOCAL MUNICIPALITY	1148136	462939	1148136
FRANCES BAARD DISTRICT MUNICIPALITY	15624720	16977295	14920435
SOL PLAATJIE LOCAL MUNICIPALITY	15624720	16977295	14920435
JOHN TAOLO GAETSEWE DISTRICT MUNICIPALITY	6258468	827712	6258468
GA-SEGONYANA LOCAL MUNICIPALITY	2236143	-1364528	2236143
WEST COAST DISTRICT MUNICIPALITY	16962244	3454785	16962244
CEDERBERG LOCAL MUNICIPALITY	1711244	899785	1711244
CAPE WINELANDS DISTRICT MUNICIPALITY	25940790	4866724	25717025
DRAKENSTEIN LOCAL MUNICIPALITY	14884000	1091000	14852000
EDEN DISTRICT MUNICIPALITY	17564843	5818402	15695816
CENTRAL KAROO DISTRICT MUNICIPALITY	1704000	1907000	1704000

Table 4.10: Total expenditure and total fixed assets of municipalities

Municipality	Total Expenditure	Fixed Assets
NELSON MANDELA BAY METROPOLITAN MUNICIPALITY	778405	13409672
CACADU DISTRICT MUNICIPALITY	15284	461898
CAMDEBOO LOCAL MUNICIPALITY	10869	76890
MAKANA LOCAL MUNICIPALITY	36118	1646243
NDLAMBE LOCAL MUNICIPALITY	12168	108052
BAVIAANS LOCAL MUNICIPALITY	2424	38224
KOUGA LOCAL MUNICIPALITY	55273	454938
AMATHOLE DISTRICT MUNICIPALITY	180481	1996570
MBHASHE LOCAL MUNICIPALITY	169	81335
MNQUMA LOCAL MUNICIPALITY	0	198776
GREAT KEI LOCAL MUNICIPALITY	0	44572
AMAHLATHI LOCAL MUNICIPALITY	0	102092
BUFFALO CITY LOCAL MUNICIPALITY	550783	12795928
NGQUSHWA LOCAL MUNICIPALITY	0	228529
NKONKOBE LOCAL MUNICIPALITY	0	53928
O.R. TAMBO DISTRICT MUNICIPALITY	569502	3696717
MBIZANA LOCAL MUNICIPALITY	0	302340
NTABANKULU LOCAL MUNICIPALITY	0	29521
NGQUZA HILL LOCAL MUNICIPALITY	0	72595
PORT ST JOHNS LOCAL MUNICIPALITY	0	343524
NYANDENI LOCAL MUNICIPALITY	0	373880
MHLONTLO LOCAL MUNICIPALITY	0	46572
KING SABATA DALINDYEBO LOCAL MUNICIPALITY	0	1428884
XHARIEP DISTRICT MUNICIPALITY	0	29784
LETSEMENG LOCAL MUNICIPALITY	15012	187537
Dr RUTH SEGOMOTSI MOMPATI DISTRICT MUNICIPALITY	41076	495681
MANGAUNG LOCAL MUNICIPALITY	399343	4878142
MANTSOPA LOCAL MUNICIPALITY	29037	452995
LEJWELEPUTSWA DISTRICT MUNICIPALITY	0	142610
TSWELOPELE LOCAL MUNICIPALITY	13734	277636
MATJHABENG LOCAL MUNICIPALITY	254834	1073961
THABO MOFUTSANYANE DISTRICT MUNICIPALITY	0	31752
SETSOTO LOCAL MUNICIPALITY	95141	401708
DIHLABENG LOCAL MUNICIPALITY	27394	1997730

Municipality	Total Expenditure	Fixed Assets
FEZILE DABI DISTRICT MUNICIPALITY	0	183216
MOQHAKA LOCAL MUNICIPALITY	54318	1220613
NGWATHE LOCAL MUNICIPALITY	48325	677298
METSIMAHOLO LOCAL MUNICIPALITY	134983	985503
MAFUBE LOCAL MUNICIPALITY	21358	303201
CITY OF JOHANNESBURG METROPOLITAN MUNICIPALITY	4308760	42930190
CITY OF TSHWANE METROPOLITAN MUNICIPALITY	1705723	19061221
EKURHULENI METROPOLITAN MUNICIPALITY	2670979	50365760
SEDIBENG DISTRICT MUNICIPALITY	0	337646
EMFULENI LOCAL MUNICIPALITY	740172	2058168
MIDVAAL LOCAL MUNICIPALITY	94582	2097728
LESEDI LOCAL MUNICIPALITY	56980	737053
METSWEDING DISTRICT MUNICIPALITY	0	15208
NOKENG TSA TAEMANE LOCAL MUNICIPALITY	36258	431062
KUNGWINI LOCAL MUNICIPALITY	57332	417581
WEST RAND DISTRICT MUNICIPALITY	0	219924
MOGALE CITY LOCAL MUNICIPALITY	207093	5732669
RANDFONTEIN LOCAL MUNICIPALITY	90020	2792722
WESTONARIA LOCAL MUNICIPALITY	98705	1393237
MERAFONG CITY LOCAL MUNICIPALITY	165972	2838585
ETHEKWINI METROPOLITAN MUNICIPALITY	3664121	38651565
UGU DISTRICT MUNICIPALITY	633853	1482268
UMDONI LOCAL MUNICIPALITY	0	602263
UMZUMBE LOCAL MUNICIPALITY	0	111057
EZINQOLENI LOCAL MUNICIPALITY	0	35513
HIBISCUS COAST LOCAL MUNICIPALITY	0	1183469
UMGUNGUNDHLOVU DISTRICT MUNICIPALITY	41244	839215
UTHUKELA DISTRICT MUNICIPALITY	260884	709042
EMNAMBITHI-LADYSMITH LOCAL MUNICIPALITY	0	896164
INDAKA LOCAL MUNICIPALITY	0	55822
UMTSHEZI LOCAL MUNICIPALITY	0	698769
NTAMBANANA LOCAL MUNICIPALITY	0	69099
IMBABAZANE LOCAL MUNICIPALITY	0	70128
UTHUNGULU DISTRICT MUNICIPALITY	227305	1102252
UMHLATHUZE LOCAL MUNICIPALITY	410898	5058392
CAPRICORN DISTRICT MUNICIPALITY	54019	1240908

Municipality	Total Expenditure	Fixed Assets
WATERBERG DISTRICT MUNICIPALITY	0	164063
MOGALAKWENA LOCAL MUNICIPALITY	74382	1004739
GERT SIBANDE DISTRICT MUNICIPALITY	0	344765
MSUKALIGWA LOCAL MUNICIPALITY	63217	324015
GOVAN MBEKI LOCAL MUNICIPALITY	207354	3714877
NKANGALA DISTRICT MUNICIPALITY	0	748874
STEVE TSHWETE LOCAL MUNICIPALITY	94135	6793073
DR JS MOROKA LOCAL MUNICIPALITY	87268	1506717
EHLANZENI DISTRICT MUNICIPALITY	0	386229
MBOMBELA LOCAL MUNICIPALITY	185215	5869157
BOJANALA PLATINUM DISTRICT MUNICIPALITY	0	218773
RUSTENBURG LOCAL MUNICIPALITY	328344	2183795
NGAKA MODIRI MOLEMA DISTRICT MUNICIPALITY	0	348092
MAFIKENG LOCAL MUNICIPALITY	56632	939881
RAMOTSHERE MOILOA LOCAL MUNICIPALITY	13695	198085
Dr KENNETH KAUNDA DISTRICT MUNICIPALITY	0	183200
TLOKWE LOCAL MUNICIPALITY	61563	1975355
NAMAKWA DISTRICT MUNICIPALITY	5	97318
RICHTERSVELD LOCAL MUNICIPALITY	7760	162719
NAMA KHOI LOCAL MUNICIPALITY	29322	402353
HANTAM LOCAL MUNICIPALITY	8063	87472
KAROO HOOGLAND LOCAL MUNICIPALITY	3195	194760
KHAI-MA LOCAL MUNICIPALITY	6777	67018
PIXLEY KA SEME DISTRICT MUNICIPALITY	0	41140
THEMBELIHLE LOCAL MUNICIPALITY	4085	200850
SIYATHEMBA LOCAL MUNICIPALITY	8313	258683
FRANCES BAARD DISTRICT MUNICIPALITY	0	123867
SOL PLAATJIE LOCAL MUNICIPALITY	143237	1210533
JOHN TAOLO GAETSEWE DISTRICT MUNICIPALITY	33187	128890
GA-SEGONYANA LOCAL MUNICIPALITY	21622	990131
WEST COAST DISTRICT MUNICIPALITY	54314	459877
CEDERBERG LOCAL MUNICIPALITY	22317	452045
CAPE WINELANDS DISTRICT MUNICIPALITY	0	551021
DRAKENSTEIN LOCAL MUNICIPALITY	97889	3592608
EDEN DISTRICT MUNICIPALITY	1060	711928
CENTRAL KAROO DISTRICT MUNICIPALITY	2222	49966