CHAPTER 8: MEASURING THE EFFICIENCY OF LOCAL GOVERNMENT EXPENDITURE: AN FDH ANALYSIS OF A SAMPLE OF SOUTH AFRICAN MUNICIPALITIES

Jugal Mahabir[105]

8.1 Introduction

Local government is part of a three-sphere system of government. The Constitution mandates municipalities to provide several important services, including the four basic services of water, sanitation, refuse and electricity. These services are enshrined in the country's Bill of Rights and are a social and legal right of all citizens. To fulfil their obligations, municipalities need to be well funded and spend efficiently to provide high-quality services to their communities.

National government has injected considerable funds into the local government sphere through the local government equitable share (LES) and conditional grants. However, despite these resources, municipalities continue to perform poorly in the delivery of basic services to their communities, as the frequent service delivery protests in recent years attest.

The social contract theory validates government to collect taxes from constituents to finance the delivery of public goods, and tax payers and citizens have the political and democratic right to know if these resources are used efficiently. Funds used inefficiently will have negative effects on the quantity and quality of public goods and service delivery to communities. It is thus vital to evaluate the spending efficiency of municipalities to provide "citizens with the necessary information to monitor their political representatives and get good value for their money" (Afonso and Fernandes, 2003).

The recent global financial crisis and recession in the South African economy has affected municipalities' own-revenue sources, which are generally cyclical in nature. In addition, the recession has had an impact on national tax revenues, resulting in potentially lower growth rates in intergovernmental transfers, while some municipalities are already fiscally stressed due to increased demand for services and other economic and socio-demographic circumstances. Therefore, the current resources available to the municipal sphere, no matter how limited, must be spent efficiently to ensure the optimum outcomes as prescribed by policy and the demand of communities.

8.1.1 Problem statement

Ultimately, local government and the services mandated to them exist because of the perceived efficiency of a decentralised system of government to improve and enhance service delivery. The argument for fiscal decentralisation is the perceived efficiency gain from delegating certain services and taxation powers to lower levels of government, which are seen as being better placed to deliver such services. While theoretically justified, such an argument needs to be validated through empirical research.

8.1.2 Background to the research

In 2009 the Financial and Fiscal Commission (the Commission) adopted a five-year research strategy to help attain its vision of enhancing the developmental impact of public resources through the financial and fiscal system in South Africa.

105 Financial and Fiscal Commission (FFC).

The four thematic areas identified to serve as fundamental principles and goals within the research strategy, which would direct the institution in attaining this vision, are:

- policy outcomes
- accountable institutions
- equitable growth and distribution of resources
- flexible responses.

This research is particularly aligned to the Commission's accountable institutions principle, as it evaluates the efficiency of municipalities in their use of public funds and will assist local constituents and national government to assess the performance of municipalities and to hold them accountable.

Each year, the Commission adopts a research theme that informs the submission on the Division of Revenue (DOR) for the given year. The focus for the 2012/13 DOR centred on three key themes of contemporary policy and debate: payback time, progressive realisation and developmental state. This research forms part of the payback time theme. It quantifies the current wastage of resources in the local government system, particularly at a time when austerity measures need to be balanced with optimal service delivery. National debt can be reduced by cutting expenditures or increasing taxes, and government has propagated the former as its policy to consolidate its fiscal position. Therefore, to fulfil the goals of government, local government needs to use resources optimally.

This chapter aims to measure the efficiency of municipalities in South Africa and to estimate the factors that affect this efficiency. The chapter will also assess the changes in efficiency over a four-year period (from the 2005/06 to the 2008/09 financial years); provide a method to benchmark municipal performance in terms of the efficient use of resources; and assess the most efficient methods of providing municipal services, specifically free basic services (FBS).

South Africa is now in its twelfth year of a fully functioning local government. Yet, the literature review found no research that has applied efficiency measurement techniques to analyse the spending performance of South African municipalities, while the analysis of efficiency over a given period is also relatively thin. Ranking municipal performance – by identifying the country's efficient municipalities and understanding the dynamics and fundamentals that inform municipal performance – will assist policy-makers to allocate scarce resources optimally to ensure sustained and adequate quality of services to communities. The research also hopes to be a springboard for more comprehensive research on the efficiency of local government, which will ultimately improve the overall performance of a very significant sphere in South Africa.

After a brief description of local government's legal functions and revenue powers and the spending patterns of municipalities, economic efficiency is defined and the methodology used is described. The free disposable hull (FDH) estimation technique and the tobit regression analysis are then outlined and their results analysed. Lastly, the conclusion includes recommendations and identifies areas for further research.

8.2 Structure and Expenditure Performance of Local Government

8.2.1 The constitutional and legal framework of local government

The Constitution

South Africa is a constitutional democratic state established as a unified country following the first democratic elections in 1994. The country operates in a unitary and decentralised system of government that constitutes three "distinctive, interdependent and interrelated" spheres: the national, provincial (made up of nine provinces) and the local (made up of 283 municipalities[106]) spheres (The Constitution of South Africa Act, No. 106 of 1996). The constitution establishes and sets the guiding principles for the cooperative governance system in South Africa. Chapter 7 of the Constitution recognises and outlines the role of local government in the country. Section 155(1) establishes three broad categories of municipalities: Category A (met-

106 This is valid as of 2010.

ropolitan) municipalities, category B (local) municipalities and category C (district) municipalities. Following an initial period of development and consolidation, in 1998 local government was officially established, consisting of a total of 853 municipalities with several transitional councils in former homeland areas. Since then several demarcation processes have resulted in the current local government configuration of 283 municipalities, including six metropolitan municipalities (usually abbreviated to metros) and 231 local municipalities that fall under 46 district municipalities.[107] The Constitution assigns metros exclusive executive and legislative powers within their area of jurisdiction, while these are shared between the local and district municipalities

Service delivery mandates and the division of powers and functions of municipalities

In accordance with the country's decentralised system of governance, social and basic services are provided either concurrently or exclusively by the three spheres of government. Generally, policy norms and standards are devised at national level and then implemented at sub-national level. While provinces implement most social services, such as education and health, local government is mandated to provide basic services of water, sanitation, electricity and refuse removal. In terms of community demand and expenditure patterns, these are arguably the most important of all the various powers and functions assigned to local government under Schedules 4B and 5B of the Constitution.

Service provision within local government is undertaken in an asymmetrical manner. The metros are responsible for all four major services, whereas either local municipalities or their district municipalities undertake or share these services. These provisions are set out in Chapter 5 of the Municipal Structures Act, No. 117 of 1998 as amended. In general, all local and metropolitan municipalities are authorised to provide electricity and, in most cases, refuse removal services. In terms of the water and sanitation function, either the district municipality or the local municipality is authorised to provide these services. This service delivery arrangement results in differences in the nature and level of municipal expenditures. Municipalities responsible for all major basic services produce a higher level of service output than those with less service delivery responsibilities; by theoretically using more inputs i.e. resources. This arrangement makes it difficult to measure the efficiency of all municipalities in the country equally and is discussed in more detail in the methodology and FDH analysis sections.

Revenue powers of local government

To fund these expenditure responsibilities, local government commands an array of fiscal instruments devolved to them in terms of Section 229 of the Constitution. These include property rates, user charges for municipal services rendered, surcharges on user charges and other local taxes. On average, local government in South Africa is largely self-financing: in the 2008/09 financial year, 77% of their total operating revenues was derived from own-revenue collection (75% on average from the 2005/06 to 2008/09 financial years). However, because horizontal fiscal and economic inequities exist across municipalities, intergovernmental fiscal transfers play a significant role in ensuring that all municipalities are well funded to fulfil their service delivery mandates. Intergovernmental transfers take the form of unconditional and conditional grants, with unconditional grants usually funding operating expenditure. The LES is the largest unconditional grant to local government and is pivotal for the operating expenses of municipalities: most rural or poorer municipalities are almost fully funded by this grant.

8.2.2 Analysis of municipal expenditure

Municipalities are responsible for providing the basic services of water, sanitation, electricity and refuse removal. Given their importance, it is not surprising that these services dominate local government expenditures. Municipal budgets are comprised of two types of expenditure: operating or current expenditure and capital expenditure. Table 8.1 shows the respective size of the operating and capital budget and outcomes for each type of municipality for the 2008/09 financial year.

107 This is valid as of 2010. The latest re-demarcation process by the Municipal Demarcation Board (the institution established to authorise the boundaries of municipalities) decreased the number of municipalities to 278 and established two additional metropolitan municipalities after the 2011 local government elections (8 metros and 226 local municipalities that fall under 44 district municipalities).

Table 8.1 Budgeted and actual expenditure per municipality type, 2008/09

Municipality category	Total budgeted operating expenditure (revised)	Total actual operating expenditure	Variation	Expenditure shares across categories	Total budgeted capital expenditure (revised)	Total actual capital expenditure	Variation	Expenditure shares across categories	Actual opex	Actual capex
Metropolitan Municipalities (6)	77,352,465	77,326,714	100%	56%	24,989,691	24,837,450	99%	62%	76%	24%
Local Municipalities (237)	46,107,580	49,444,327	107%	36%	17,282,041	11,799,552	68%	29%	81%	19%
District Municipalities (46)	7,520,330	10,903,647	145%	8%	6,623,650	3,501,466	53%	9%	76%	24%
Total	**130,980,375**	**137,674,688**	**105%**	**100%**	**48,895,382**	**40,138,468**	**82%**	**100%**	**77%**	**23%**
Sampled Municipalities (129)	114,422,768	117,150,075	102%	85%	38,265,536	34,577,670	90%	86%	77%	23%

Note: Figures in R'000.
Source: National Treasury Local Government Budget Database

Municipalities spent most of their funds (77% on average) on operating expenditures in 2008/09. The largest expenditure occurs in the six metros, which account for over 56% total operating expenditure and over 60% of total capital expenditure by local government. This is attributed to the greater demand for services in these areas due to the higher numbers of people, households and businesses.

The asymmetrical service delivery arrangements complicate analysis of local and district municipalities' expenditure patterns. On average, district municipalities spend 24% of their budgets on capital expenditure, which is higher than the national average. This trend is probably due to the role of service-authorised district municipalities, where the general practice is for the district municipality to invest in infrastructure while delegating the operational provision of the service to the local municipality through service delivery agreements. Table 8.1 also illustrates the variations in planned (budgeted)[108] and actual expenditure patterns, which differ for each type of municipality and show that metros are better able to plan and budget. The lack of planning and spending capacity for district municipalities is evident, as in 2008/09 they overspent on their operating budget (by 45%) and underspent by almost 50% on the capital budgets.

Composition of operating and capital budgets

Operating expenditure consists largely of the current costs, such as labour and other input costs, necessary for the delivery of services. Figure 8.1 shows the general composition of operating expenditure of all municipalities for 2008/09. The three largest components of municipal operating budgets are 'Other'[109] operating expenditure (31%), employee costs (27%) and material and bulk costs (22%).

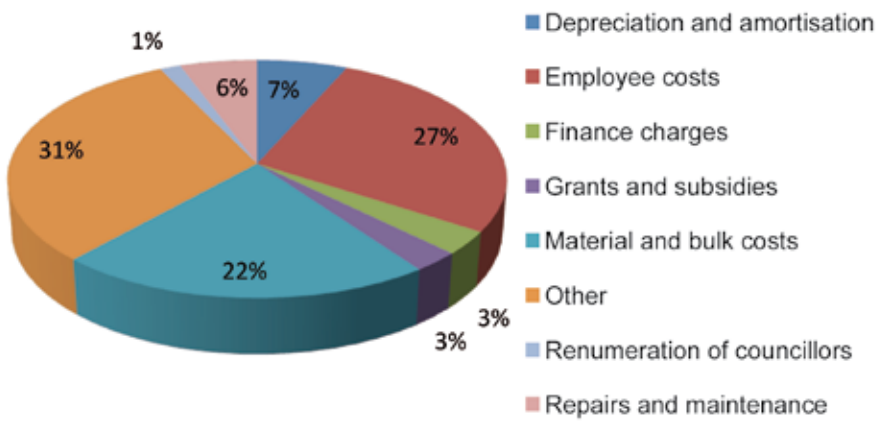
The largest portion of operating expenditure is for input costs (labour, material and capital costs) that are necessary to provide services. Figure 8.1 includes expenditure of municipalities that do not provide all four major services, which means that related input costs (personnel and bulk purchases) would probably be higher for municipalities that provide all four services.

The line item 'Other' is likely to include municipal administration costs, which are also important for providing services. The level of expenditure on repairs and maintenance is significant, as it is related to maintaining the infrastructure that provides services to communities. However, the 6% shown in Figure 8.1 is below the best practice figure of 8–12%, as recommended by the National Treasury. Lastly, the grants and subsidies category depicts transfers from district municipalities to local municipalities, which provide services on their behalf in areas that have service delivery agreements in place.

108 Mid-year revised budgets.

109 The new municipal reporting requirements will ensure that municipalities disaggregate all types of expenditure ensuring the removal of the 'Other' expenditure line item.

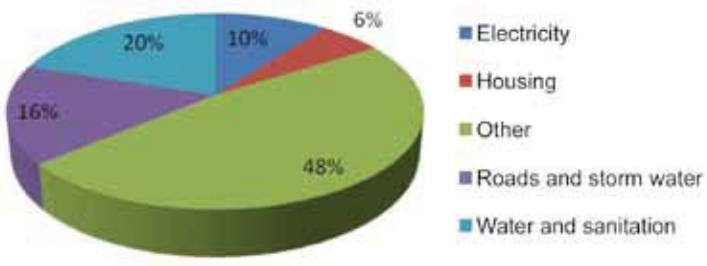
Figure 8.1 Composition of municipal operating expenditure (%), 2008/09



Source: National Treasury Local Government Database

Figure 8.2 gives a breakdown of municipal capital expenditure in 2008/09. Capital expenditure usually entails longer-term investments in social and economic infrastructure. This is necessary to eradicate service delivery backlogs that exist in a developing country such as South Africa (social infrastructure), and provide additional economic infrastructure to promote local economic development.

Figure 8.2 Composition of municipal capital expenditure (%), 2008/09



Source: National Treasury Local Government Database

The bulk (48%) of capital expenditure of municipalities falls under the category 'Other'. Unfortunately, this category is not disaggregated further but would include expenditure on infrastructure projects, such as street lighting and community halls. Expenditure on roads and storm water constitutes 16% of total capital expenditure, while investments in water and sanitation (20%) and electricity (10%) infrastructure are also significant.

Housing is the lowest of the total capital expenditure, at 6%. The lower share for housing expenditure, and to a certain extent electricity infrastructure, is because the housing function is a provincial competency, while the capital roll-out of infrastructure for electricity is shared between local government and Eskom.[110] Investment in housing infrastructure is usually undertaken by metros, on the authority of the respective provinces.

The focus here is on municipal operating expenditure because operating expenditure results in the immediate delivery of a service or output of municipality for public consumption. Therefore, operating expenditure can serve as a good proxy for a municipality's input costs, as it captures items such as personnel, bulk charges and other administrative costs. In contrast, capital expenditure looks at longer-term infrastructure projects and, although these projects are an output of municipalities, they are not linked to the immediate provision of a service but rather extend the level of a service at a later stage. Municipal operating expenditure is therefore a better proxy to link with immediate municipal outputs.

110 Eskom is a state-owned entity responsible for the generation and transmission of electricity. It also plays a large part in the distribution industry, although this is legally a local government competency.

8.3 Defining Economic Efficiency

Economics deals with the economic problem of the unlimited wants of human nature and limited available resources. Economic efficiency can be defined as how consumers, firms and society choose to use and allocate these scarce resources optimally. Table 8.2 offers a list of the different types of economic efficiency.

Table 8.2 Types of economic efficiency

Type of efficiency	Definition
Productive efficiency	All resources are used efficiently, so that the maximum output can be produced with the minimum level of input.
Technical efficiency	This is similar to productive efficiency.
X inefficiency	This occurs when certain decision-making units (DMUs) willingly use more inputs than required, which results in higher average costs than necessary. This usually occurs in a monopoly where the company has no incentive to cut unnecessary costs because it is already making supernormal profits.
Allocation/distribution efficiency	Distributive efficiency ensures that goods and services are allocated to individuals or communities where the need is greatest or where the utility (advantage or fulfillment from using a good or service) is maximised. DMUs can be productively efficient but might not be distributive efficient.
Dynamic efficiency	This is the ability to adapt new technologies, to maintain an optimum level of production that the latest available technology enables. For example, a DMU might be efficient using an available technology, but could become dynamically inefficient if it does not upgrade to a new innovation that might increase output with the same amount of inputs.
Efficiencies of scale	This is a DMU's long-term incentive of developing economies of scale by producing on the lowest point of the long-run average cost curve.
Social efficiency	This concerns limiting the negative externalities that the production of a good or service might have on society, i.e. social marginal benefits are equal to social marginal costs.

Source: http://www.economicshelp.org/microessays/costs/efficiency.html - Assessed July 2010

Most of the different types of efficiencies would apply to, and should be measured for, municipal performance. When producing an output or service, municipalities should be socially efficient in limiting pollution and other negative externalities in their communities. They should also try to cut input costs by developing economies of scale. Municipalities should be dynamically efficient in keeping up with the latest technologies. This can also apply to human capital, where in some municipalities, officials are not appropriately qualified to perform optimally. Therefore, municipalities need to actively train and develop staff and invest in research and development.[111]

However, this chapter focuses on the technical efficiency of municipalities, and whether they use their inputs (expenditure) effectively to produce the optimal level of outputs (service delivery). A municipality is considered technically inefficient if (1) it produces the same or fewer outputs than another municipality with a higher amount of inputs or (2) a municipality produces fewer outputs relative to another municipality using the same level of inputs.

8.4 Overview of Methodology and Data Sources

The non-parametric FDH technique is used to compute input efficiency scores for a sample of 129 South African municipalities in terms of the technical efficiency (i.e. the optimal use of resources/inputs to maximise services/outputs) of operating expenditure from 2005/06 to 2008/09. Techniques such as FDH rank municipalities relative to the most-efficient municipalities within the sample that fall on the production possibility frontier (PPF). Then, using a cross-sectional tobit regression model, the factors that have an impact on the computed efficiency scores for 2008/09 are identified and quantified in the context of local government in South Africa.

8.4.1 Methods of estimating technical efficiency and relation to local government

The three most-popular methods used to measure technical efficiency are stochastic frontier analysis (SFA), a parametric technique, and two non-parametric techniques, data envelopment analysis (DEA) and FDH. Deprins *et al*. (1984) provide a comprehensive mathematical depiction and practical application of the three methods. The following section gives a simple

111 It is sometimes argued that there is no incentive for government agents and government in general to invest in research and development and to be dynamically efficient, as they are not profit-driven economic actors.

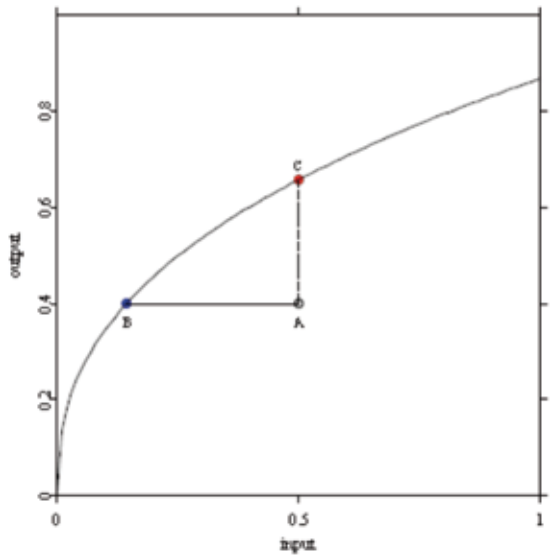
description of one interpretation of these methods. Efficiency analysis is closely related to productivity analysis "In conformity with microeconomic theory, production processes are technical relations of employed inputs to maximum attainable outputs" (Behr and Tente, 2008). This relationship can be viewed in terms of a production set, Ψ, i.e. the combination of a set of inputs that are used to produce a set of outputs. This relationship can be depicted as:

$$\Psi = \{(x,y)|\, x \in \mathbb{R}_+^p, y \in \mathbb{R}_+^q, (x,y) is\ feasible$$

Where Ψ = production set
x = vector of inputs
y = vector of outputs

With p>0 and q>0, using $x_i, \ldots, x_p$ input quantities, it is feasible to produce $y_i, \ldots y_q$ outputs. The SFA, DEA and FDH are methods used to estimate the production set or Ψ. As Park *et al*. (2000) note, "estimators for the production set also induce estimators for the [production] frontier and for the efficiencies". The PPF can be defined as the boundary of a production set where output is maximised and is illustrated in Figure 8.3.

Figure 8.3 Production possibility frontier and technical efficiency



Source: Adapted from Hardle and Jeong, 2005

Each of these methods differs mainly by the assumed functional form of the PPF. With SFA, the functional form needs to be specified, usually in the form of a production function (such as Cobb-Douglas). DEA and FDH analysis simply envelop a given dataset forming the frontier. However, DEA assumes convexity in the form of the PPF, whereas FDH analysis assumes that the inputs and outputs are freely disposable.

These approaches are methods used to evaluate how well a decision-making unit transforms spending inputs into outputs (Afonso and Fernandes, 2003) relative to the PPF. Therefore, "technical efficiency is measured as the distance between an observed production unit and the postulated boundary", as determined by the choice of technique (De Borger *et al*., 1994). Observations or DMUs below the PPF are considered inefficient relative to the PPF.

In Figure 8.3, DMUs B and C lie on the specified PPF, which means that these DMUs are maximising their outputs with their given inputs. DMU A is considered inefficient, as it is producing below the PPF. Relative to point B, A is using more inputs to produce the same level of outputs. The distance BA is therefore the computed input-orientated efficiency measure i.e. A is inefficient as it can produce the same amount of outputs as B with less input. The distance AC depicts the computed output-orientated efficiency measure. Relative to DMU C, A is inefficient because it produces less outputs using the same amount of inputs.

Input-orientated efficiency will be measured over a four-year period for the 129 municipalities. While profit-maximising agents tend to have control over both input and output decisions, municipalities tend to have more control over their inputs than their outputs because of the social nature of services provided by local government in South Africa. Crawford *et al*. (2003) point out that, although outputs and inputs are generally easy to define and measure for a private firm or institution, the

measurement of publicly provided goods, such as education and health care, are complicated. Appropriate proxies to capture such measurements can be developed for local government. For example, municipal services can be equated to outputs from the production processes, i.e. the transformation of inputs (labour and capital) into outputs, with municipal recurrent expenditures as the given input. In the context of South Africa, this would include personnel costs (labour inputs), bulk purchases (material inputs) and other operating expenditure, such as repairs, maintenance and depreciation of existing assets (proxy for capital inputs). Afonso and Fernandes (2003) note that "outcomes may not necessarily reflect the service desired by local residents". The "ultimate outcome [level demanded by communities] of these services is measured by effect indicators, which reflect the degree to which direct outputs of municipal activities translate into welfare improvements" (De Borger and Kerstens, 2000 cited in Afonso and Fernandes, 2003). Approximate proxies for the demand for services by communities are needed, as individual and community utilities or welfare are difficult to measure.

8.4.2 Overview of the FDH efficiency estimation technique

The FDH technique will be used to estimate efficiency scores. This technique was first explored by Deprins *et al*. (1984)[112] in their estimation of labour efficiency in post offices and was developed "on the sole assumption of input and output disposability" (Deprins *et al*., 1984), i.e. the inputs and outputs are freely disposable. This section briefly describes the mechanics, advantages and disadvantages of using the FDH technique compared with the alternatives in the context of local government. Deprins *et al*. (1984), De Borger *et al*. (1994), Park *et al*. (2000), Leleu (2006) and Hardle and Jeong (2005) provide a comprehensive overview of the FDH efficiency estimation technique and non-parametric productivity analysis, in terms of its mathematical fundamentals and related technicalities.

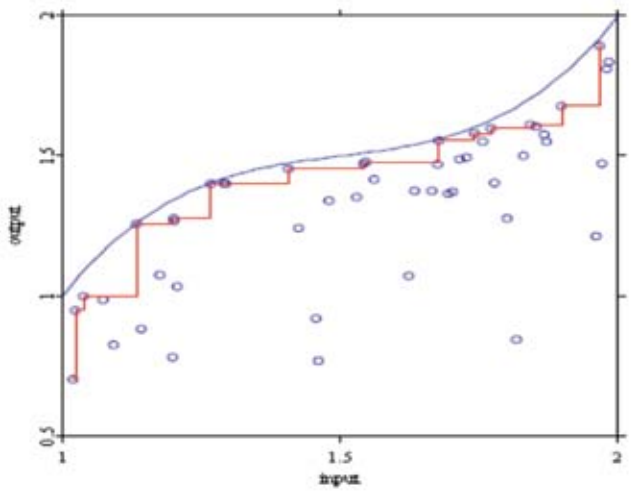
Brief description

The FDH approach is one method of computing a production set, PPF, and the distance between the PPF and an inefficient observation. FDH can be defined as "the smallest free disposable set containing all observations in a sample of production units" (Park *et al*., 2000). Building on the description of the production set (Ψ) above, the FDH production set estimator can be depicted as:

$$\theta^{OUT}(x_1, y_1) = \max_{i|x \geq x_i} \quad \min_{1 \leq k \leq q} \quad y_i^k / y_1^k$$

Here $X\{(x_i, y_i), i = 1, \ldots n)$ is a given sample of observations or, in this case, a sample of municipalities. The FDH technique relies on the free disposable assumption that $x \geq x_i$ and $y \leq y_i$ belong to the production set, Ψ, and is defined as the FDH of the sample X. This is illustrated, as an example, in Figure 8.4.

Figure 8.4 FDH of a given sample



Source: Adapted from Hardel and Jeong, 2005

Given a specific municipality with a point on the Euclidean space, i.e. $\{x_1, y_1\} \in \Psi$ the FDH efficiency score, θ, is the distance

112 Confirmed in Park et al. (2000) and Leleu (2006).

between this point and the disposable hull of the Ψ as estimated. As mentioned, this distance can be output orientated or input orientated. The output efficiency estimate θ^{OUT} can be calculated by:

$$\theta^{OUT}(x_1, y_1) = \max_{i|x \geq x_i} \quad \min_{1 \leq k \leq q} \quad y_i^k / y_1^k$$

As explained above, only the input-orientated efficiency scores will be estimated for the sample of municipalities. The input efficiency estimate θ^{IN} can be calculated by:

$$\theta^{IN}(x_1, y_1) = \min_{i|y \leq y_i} \quad \max_{1 \leq j \leq p} \quad x_i^j / x_1^j$$

The FDH efficiency score for a given municipality is a measure relative to a best-practising or most-efficient municipality that lies on the PPF. Therefore, in an FDH analysis, it is more appropriate to refer to the PPF as the best practising frontier (BPF), as the PPF is determined by the 'efficient' municipalities relative to the given sample and not to a hypothetical DMU, as with SFA and DEA. The technique simply identifies the best practising DMU within the given sample based on the efficiency in transforming inputs into outputs. It is entirely possible that a DMU considered efficient in one sample might not be efficient in another sample.

Advantages of FDH

In general, the FDH technique makes fewer assumptions relative to other techniques. Unlike the SFA technique, the production function of a production set does not need to be specified, as FDH is a non-parametric technique. Compared to other non-parametric techniques such as DEA, FDH does not have many assumptions on the form or structure of the production set, especially for the assumption of convexity, which is paramount in DEA. The removal of these assumptions makes it possible to estimate efficiency relative to the performance of other DMUs within a sample, rather than a hypothetical DMU generated by the production function (in SFA) or a fitted value of the convex production (DEA). Therefore, the FDH approach appears the most practical for the purposes of policy analysis. This is supported by Deprins *et al*. (1984), who conclude that the FDH analysis is the most robust in terms of practical application, as efficiency is measured relative to other productivity units (such as provinces, or post offices in their case). This enables managers and policy-makers to compare performance and to share and learn from best practising entities. Well-performing municipalities can be identified, and less-efficient municipalities can improve performance by mirroring these best practices. Deprins *et al*. (1984) also affirm the relative simplicity of the FHD approach as "simple comparisons of the input and output components of the vectors in a data set" as opposed to a "single large linear programme" for parametric techniques and "many linear programmes" for the DEA technique.

Not requiring the convexity assumption can prove advantageous when analysing production efficiency for municipalities. Generally, the convexity assumptions refer to the theory of firms and the production function of firms, but public agents have no incentive to maximise profit or minimise costs. As noted by De Borger *et al*. (1994), since there is "no generally accepted model of governmental behaviour, the minimal technical and behavioural assumptions underlying the FDH make it a particular useful tool for analysing public sector efficiency".

Disadvantages of FDH

Like other non-parametric approaches to productivity efficiency analysis, FDH is determined by the quality and distribution of the data used and is influenced significantly by the number of observations and number of outliers in the dataset. In the case of outliers, the FDH technique will consider these DMUs efficient and envelop the FDH production curve along them. Several endogenous factors, which have an impact on the quantity of inputs or outputs, may result in possible outliers that can be more frequent in cross-sectional studies. It is therefore not possible to view this outlier as fully efficient, which means that the FDH approach is biased with regards to the nature of the dataset. De Borger *et al*. (1994) confirm that "increasing the sample size increases the possibility of dominance for any given observation, and therefore the probability of being denoted inefficient". Although referring specifically to the DEA technique, Crawford *et al*. (2003) point out that it is possible for a production unit to be deemed efficient simply because it produces more output than any other unit in the sample. They explain that "in (a) single-output example the organisation that produces the most will find itself on the efficiency frontier simply because there is no larger organisation with which to compare it" (Crawford *et al*., 2003). Furthermore, the efficiency scores are highly dependent on the number of inputs and outputs used in the analysis (Crawford *et al*., 2003). The basic fundamentals of FDH are to identify the most-efficient DMUs within a sample, based on using fewer inputs to produce an amount of output (in

terms of input efficiency). The FDH efficiency score is the distance between the efficient DMU and the dominated DMU (the one considered inefficient relative to the same). However, the nature of the method can result in the possibility of undominated DMUs, which Afonso and Fernandes (2003) define as "weak vector dominance limitation". These DMUs are efficient by default, as they have the lowest levels of inputs and outputs and are neither dominated by, or are dominating another DMU. De Borger *et al*. (1994) point out that the FDH techniques cannot be used to determine factor productivity and economies of scale and scope owing to their characteristic of being a multidimensional step function. Da Conceicao Sampaio de Sousa and Battaglin Schwengher [sa] estimate the technical efficiency for courts in the Rio Grande do Sul Province in Brazil using FDH and order-m techniques. In terms of the methodology, they conclude that the efficiency scores computed using the order-m technique appears more reliable than scores computed using the FDH technique.

8.4.3 Overview of the regression model

A tobit regression model is used to estimate the factors that influence the efficiency scores generated by the FDH analysis primarily because of the nature of the distribution of the FDH efficiency variable (Boetti *et al*., 2009). Several municipalities from the sample are likely to be deemed efficient, i.e. situated on the PPF, and will therefore have an efficiency score of 100% or 1. As the distribution of the variable will probably be clustered at the upper level, ordinary least squares (OLS) estimates will probably be biased.

8.4.4 Data sources and rationale for sample

The primary data used for estimating efficiency is budget data, which is sourced from the National Treasury Local Government Database for 2005/06 to 2008/09. Service-level data for water, sanitation, electricity and refuse removal is sourced from the Non-financial census (NFC) of municipalities, an annual survey undertaken by Statistics South Africa (StatsSA), over the same period. To estimate the population per municipality over this period, the most comprehensive available population data was used, which was from the 2001 Census and the 2007 Community Survey.

The asymmetrical approach to service delivery (whereby a municipality is either authorised to provide a service or not) makes it difficult to compare the spending patterns across all municipalities. Municipalities that are authorised to provide services would have higher levels of expenditure and so it is inappropriate to assess the expenditure levels of each and every municipality. Therefore, in order to have a level playing field, the analysis only considered municipalities authorised to provide the four major services of water, sanitation, electricity and refuse removal. This narrowed the sample size to 129 municipalities, including all the metros and 'urban' local municipalities.[113] It excluded all district municipalities, as they are generally only authorised to provide water and sanitation services. As Table 8.1 shows, the sample of 129 municipalities constitutes 85% and 86% of total operating and capital expenditure respectively. Their expenditure patterns mirror their provision of the four basic services and the higher demand for services; thus this sample captures the bulk of total local government operating expenditure for 2008/09.

The data used in the regression model is mainly sourced from the 2009 NFC of municipalities. This data includes information on service delivery arrangements (i.e. whether services are outsourced), management and general vacancy rates and the arrangement with regards to the provision of FBS within each municipality. Information on the qualification and experience of the municipal manager and chief financial officer (CFO) of each municipality was sourced from the Municipal Demarcation Board capacity assessment database. The Independent Electoral Commission provided detailed information on the 2006 municipal election results per municipality to obtain the political composition of local government.

8.5 Review of Previous Empirical Studies

Literature widely uses frontier production techniques to calculate efficiency scores for a range of DMUs, with studies including efficiency analysis of firms, banks, post offices, social services and countries. Such analysis can be applied to any context where the calculation of a productivity set is possible. The literature surveyed applies parametric and various non-parametric techniques to measure the efficiency of local government expenditure. Afonso and Fernandes (2003) use the non-parametric FDH technique to measure the efficiency of a sample of Portuguese municipalities, estimating both input and output-oriented efficiency scores using a one input/one output FDH model. The research finds that, on average, municipalities in the sample can produce the same amount of outputs with 39% less inputs. The method they use to measure outputs is an index of services, which is followed in this study to simplify the analysis. Previous studies that identify and estimate the efficiency of municipal expenditure first use one of the methods available to compute efficiency scores for each municipality. These scores are then regressed against various factors

113 Termed as the secondary cities or top 21 local municipalities with the largest budgets.

that would theoretically have an impact on the efficiency of local government. To compute efficiency scores, most studies use the parametric SFA method (rather than non-parametric techniques) because of its statistic fundamentals. Da Motta and Moreira (2009) include a range of political, economic and institutional variables in their SFA estimation of efficiency scores for municipal expenditure in Brazil. The results show that large costs incurred in computerising or improving the administrative technology actually have a negative impact on efficiency. Outsourcing a service produces the same result, which led to Da Motta and Moreira (2009) concluding that these practices do not necessarily improve transparency levels. The results also show that a re-elected mayor correlates with greater expenditure efficiency of municipalities, which indicates that political stability is important to maintain efficient local governments.

De Borger and Kerstens (1996) apply radial and non-radial FDH techniques to estimate the efficiency for Belgian municipalities. The study assesses the impacts and quality of FDH efficiency scores dependent on whether radial or non-radial distance measures were used. Firstly, they conclude that, when calculating efficiency scores, the non-radial methods are more robust than radial distance measures, and that individual municipality scores were very sensitive to the method used, more so for the radial distance measures. While acknowledging this finding, the radial distance measure will be used in this research for the FDH analysis. Secondly, using a tobit regression model, De Borger and Kerstens (1996) conclude that fiscal revenue capacity and block grants positively influence municipal efficiencies; whereas the nature of how services are financed (own revenue or grants) and municipal political dynamics negatively influence efficiency.

Boetti *et al*., (2009) assess the impacts of increased fiscal decentralisation policies on the expenditure efficiency of a sample of Italian municipalities, using both SFA and DEA techniques to calculate the efficiency scores. Commenting on the De Borger and Kerstens (1996) study, Boetti *et al*., (2009) note "the use of more than one methodology to measure efficiency stems from the attempt to check the robustness of the results obtained through different measurement techniques". They conclude that more autonomous local jurisdictions (in terms of fiscal powers and expenditure decisions) are more efficient in their expenditure. They also find a link between expenditure patterns and the electoral cycle. This suggests that expenditure increases closer to election periods, as politicians attempt to carry favour with the electorate. One is also encouraged to refer to the Boetti *et al*., (2009) study for a comprehensive list of previous studies that estimate expenditure efficiency scores for municipalities and the specific technique used in computing the scores (found in annexure A of their paper) and thereafter a list of studies that identified and quantified the determinants of the efficiency of local government expenditure (found in annexure B of their paper). De Borger and Kerstens (1996) and Boetti *et al*., (2009) use a methodology similar to the one proposed here to estimate the factors influencing the efficiency scores generated. Both use the tobit regression method because of the distribution of the efficiency score variable.

8.6 Research Limitations

Local government data, when available, is plagued with problems and inconsistencies. Firstly, for a project of this nature, one would have preferred a more robust measure of municipal outputs as it relates to service delivery and related administrative outputs. Instead, output proxies are used, in the form of population and access to services of consumer unit indices. However, the problem with such proxies is that they do not measure the quantum of services consumed by households and businesses. In other words, the implicit assumption is that consumer units consume the same amount of a given service. This is a very strong assumption, as a large business consumer unit is likely to consume more services (such as water and electricity) than a household consumer unit. Yet, the output measure used loses these dynamics, which are likely to have an impact on the level of services for certain municipalities. It is also highly likely that poor reporting, especially by the poorer and less-capacitated, municipalities have affected the results of the efficiency scores. Under-reporting on expenditure indicates lower usage of input costs, which incorrectly suggests that these municipalities are relatively efficient.

8.7 FDH Analysis

8.7.1 Data compilation and calculations

Inputs x (factors of productions) and outputs y (goods and services consumed) need to be defined in order to assess the input-orientated efficiency of the sample of municipalities. To simplify the analysis into a single dimensional analysis of the input-output production set, various inputs and outputs were computed into single input and output indices.[114]

114 Future work will explore multivariate inputs and outputs.

The first step was to use municipal operating expenditure per capita as the input variable in an attempt to estimate the efficiency of these recurrent expenditures. Municipal operating expenditure includes personnel costs (labour inputs), bulk charges (materials) and other necessary expenditure (capital related) required in the delivery of a service and proved to be the most feasible proxy for municipal inputs. This variable was produced by dividing the operating expenditure from 2005/06 to 2008/09 by the total population (as generated) within each sampled municipality for the period under review.

The second step required an appropriate measure of municipal outputs. As water, sanitation, electricity and refuse removal are the primary service mandates of local government, and hence the largest composition of their budgets, these are the 'outputs' local government produces using its inputs. A measure of the level of services (outputs) provided by a municipality, or alternatively the demand for such services, was obtained from the NFC of municipalities as the number of consumer units with access to these services. These service levels are defined as:

- access to piped water (for water provision);
- grid electricity (for electricity);
- a ventilated pit latrine and a flushable toilet (for sanitation); and
- removal of solid waste at least one a week (for refuse).

The primary reason for using access to services at a consumer unit level from the NFC, rather than only household data, was because municipalities provide these services to businesses and other sectors of their communities, not just households. Furthermore, the use of consumer units also captures the service to the point of delivery and billing, which makes it an accurate indication of demand for such services.

Municipal outputs are not only dependent on the four basic services they provide. Municipalities also provide a host of other smaller services and administrative outputs to their communities. To measure the level of these 'other' and administrative services, the total population of each municipality was used as an indicator for the demand for these services.

Each output indicator was converted to a ratio of the sample mean for each of the four years. A municipality with a ratio higher than the sample mean indicated a higher demand for these services in the specific municipality relative to other municipalities, and hence a higher required output level. Each ratio was then equally weighted[115] and used to quantify a single output index (i.e. the average of all five service ratios) for each financial year.

8.7.2 Results

Over the period, four municipalities remained constantly efficient: Thembisile in Mpumalanga, Polokwane in Limpopo, Mangaung in the Free State and eThekwini in Kwazulu-Natal. The average efficiency score was 0.30 in 2005/06, peaking at 0.39 in 2007/08, and declining to 0.35 in 2008/09. This suggests that, on average, municipalities in the sample can obtain the same level of output with at least 60–70% less inputs (resources).

Average efficiency increased gradually until 2007/08 and declined in 2008/09. This is possibly due to the initial impacts of the global financial crisis and economic recession, which could have resulted in a decline in resources available to municipalities (driven by a decline in local taxes and slower growth rates in intergovernmental transfers). As the demand for services remained stable, municipalities probably had to become more efficient in using scarcer resources during the recession and were perhaps not successful.

115 There are two reasons behind this relatively strong assumption. Firstly, it was convenient not to complicate the analysis with different weights per service. Secondly, the composition of municipal budgets varies across municipalities depending on the nature of their customer base. Equal weights were used to average out these effects.

Table 8.3 lists the municipalities with 100% input-orientated efficiency scores over the period.

Table 8.3 Efficient municipalities (%), 2005/06–2008/09

Municipality	2005/06 Score	Municipality	2006/07 Score	Municipality	2007/08 Score	Municipality	2008/09 Score
Bushbuckridge	100	City of Cape Town	100	Bushbuckridge	100	Bushbuckridge	100
City of Cape Town	100	City of Johannesburg	100	Dr JS Moroka	100	City of Cape Town	100
City of Johannesburg	100	Emfuleni	100	eThekwini	100	eThekwini	100
Dr JS Moroka	100	eThekwini	100	Mangaung	100	Mangaung	100
Emfuleni	100	Mangaung	100	Mbombela	100	Mbombela	100
eThekwini	100	Msunduzi	100	Nkomazi	100	Polokwane	100
Maluti-a-Phofung	100	Polokwane	100	Polokwane	100	Thembisile	100
Mangaung	100	Rustenburg	100	Thembisile	100		
Moses Kotane	100	Thembisile	100				
Polokwane	100						
Thembisile	100						
Number of efficient municipalities	**11**		**9**		**8**		**7**

Source: Author's calculations

In the 2005/06 financial year, 11 municipalities were computed as efficient, but the number decreased to just seven in 2008/09. Four municipalities (eThekwini, Mangaung, Polokwane and Thembisile) were constantly efficient over the period, while two municipalities (the City of Cape Town and Bushbuckridge in Mpumalanga) were efficient in three of the four years.

Table 8.4 lists the ten municipalities with the highest input-orientated efficiency scores relative to the efficient municipalities depicted in Table 8.3.

Table 8.4 Top ten efficiency scores (%), 2005/06–2008/09

Municipality	2005/06	Municipality	2006/07	Municipality	2007/08	Municipality	2008/09
Ekurhuleni	96.63	Buffalo City	97.04	Nelson Mandela	95.64	Emfuleni	97.68
Madibeng	94.88	Maluti-a-Phofung	95.42	City of Cape Town	94.58	Nelson Mandela	94.60
Rustenburg	94.56	Dr JS Moroka	87.69	Emfuleni	91.06	!Kheis	87.63
Nelson Mandela	93.73	Ekurhuleni	85.01	Buffalo City	90.8	City of Johannesburg	84.22
Mbombela	92.10	Nelson Mandela	84.94	Ekurhuleni	86.03	Rustenburg	84.13
Newcastle	80.29	City of Tshwane	81.17	!Kheis	85.98	City of Tshwane	82.92
City of Tshwane	79.23	Mbombela	66.00	Msunduzi	82.32	Nkomazi	82.00
City of Matlosana	76.57	Nkomazi	65.67	Albert Luthuli	78.09	Msunduzi	81.36
Buffalo City	70.67	Albert Luthuli	64.25	City of Tshwane	75.33	Dr JS Moroka	80.07
Matjhabeng	66.72	City of Matlosana	57.36	City of Johannesburg	74.91	Ekurhuleni	79.01

Source: Author's calculations

The range of the ten most-efficient municipalities differed across the period. In 2008/09, efficiency scores for the ten municipalities ranged from 79.01% to 97.68%, a smaller range than in 2006/07 (57.36% to 97.04%). The municipalities highlighted in red were efficient in the preceding year, but their performance declined in the following year. For example, as Table 8.3 shows, Dr JS Moroka used its resources efficiently in 2005/06, with a 100% efficiency score, only to see its performance drop to 87.69% in 2006/07 (see Table 8.4). In other words, this municipality used 12.21% more resources in 2006/07 than was required based on the demand for services from its communities.

Twice, Dr JS Moroka appeared efficient in one year (2005/06 and 2007/08), only to become relatively inefficient the next year. The City of Cape Town was efficient in three of the four years; in 2007/08 it was inefficient with a score of 94.58%. Table 8.5 lists the ten least-efficient municipalities.

Table 8.5 Ten least-efficient municipalities (%), 2005/06–2008/09

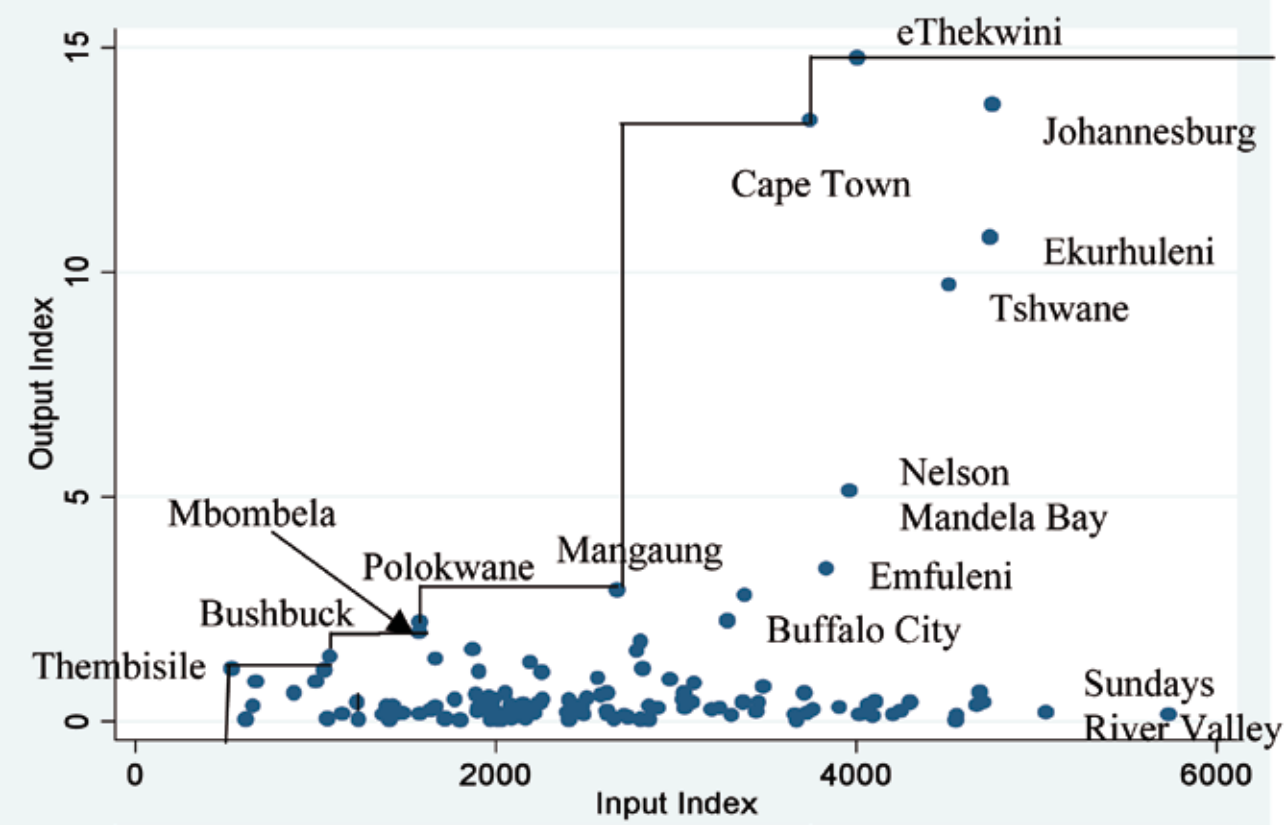
Municipality	2005/06	Municipality	2006/07	Municipality	2007/08	Municipality	2008/09
Mookgopong	8.17	Hessequa	9.51	Hessequa	16.07	Mookgopong	R 12,74
Tlokwe	8.02	Gammagara	9.15	Kungwini	15.33	Hessequa	R 12,63
Oudtshoorn	7.88	Laingsburg	9.04	Midvaal	15.15	Kouga	R 12,47
Laingsburg	7.76	Lesedi	9.03	Kareeberg	14.41	Laingsburg	R 11,77
Saldanha Bay	7.56	Saldanha Bay	8.81	Saldanha Bay	14.13	Blue Crane Route	R 11,75
Hessequa	7.46	Bitou	8.50	George	14.03	Knysna	R 11,48
Knysna	7.05	George	8.39	Overstrand	14.00	George	R 11,44
Bitou	6.38	Midvaal	8.36	Kouga	13.20	Overstrand	R 11,40
Overstrand	6.03	Overstrand	8.18	Bitou	12.52	Bitou	R 10,61
George	5.91	Knysna	6.94	Knysna	12.52	Sundays River Valley	R 9,35

Source: Author's calculations

The efficiency scores of the municipalities in Table 8.5 are extremely low and of concern. They suggest that most of these municipalities are inefficiently spending on average over 90% of their resources. Most of these municipalities are found in the Western Cape. The municipalities highlighted in red are among the ten least efficient municipalities across the four financial years while the municipalities in yellow are among the 10 least efficient municipalities in at least three years.

Figure 8.5 depicts the BPF and the seven municipalities computed as efficient for the 2008/09 municipal financial year (Bushbuckridge, City of Cape Town, eThekwini, Mangaung, Mbombela, Polokwane and Thembisile).

Figure 8.5 Best practice frontier, 2008/09



Source: Author's calculations

The seven efficient municipalities all lie on the BPF, while inefficient municipalities are below the curve. The further away from the curve and along the x-axis (input index), the more inefficient the municipality in terms of input-orientated efficiency. The most inefficient municipality is Sundays River Valley, which is found furthest away from the BPF and the furthest along

the x-axis. Interestingly, Mbombela municipality in the Mpumalanga Province improved its efficiency late in 2008/09, which was around the same time that it received assistance from the Mpumalanga Provincial Government, in terms of Section 139 of the Constitution.[116] The improvement in efficiency is probably related to the intervention.

8.7.3 Analysis of results

The FDH efficiency analysis for input-orientated efficiency identifies the municipality that uses the least inputs to produce the most outputs relative to other municipalities in the sample. The analysis therefore compares similar types of municipalities and identifies the most efficient. Although Bushbuckridge is notorious for poor performance and related problems, it is more efficient in using resources than similar municipalities, such as Moretele. In identifying good performance, the FDH allows similar municipalities to benchmark themselves against better-performing municipalities of a similar nature.

A shortcoming of the analysis is that the output measure is consumer units with access to services, which means that the quantum of service delivery cannot be assessed; the implicit assumption is that all consumer units consume the same levels of outputs. Furthermore, differences in the quality of the services cannot be assessed. Therefore, municipalities, which have relatively few consumer units with access to services but provide a higher quantity and quality of the service, would be unfairly assessed. The quantity of resources used can be impacted by the quality and quantity of services provided. However, it could also indicate that there are differences between the distributions of resources across the various types of municipalities. For example, there is a widely accepted perception that Western Cape municipalities are relatively well resourced in the maximisation of own revenues. Furthermore, Western Cape municipalities have greater fiscal capacity due to their favourable economic and demographic characteristics relative to other parts of the country. The analysis suggests that it is possible that these municipalities are overly endowed with financial resources relative to the demand for their services as many of these municipalities are inefficient. In other words, their resource capacity exceeds their outputs generated, resulting in a relatively inefficient use of given resources.

Alternatively, relatively more efficient municipalities could either be providing a lower level and quality of services, or have fewer resources available to them relative to the demand for services from their communities. These reasons could be related to lower per capita expenditure levels.

Assuming that quantity and quality of services influences the efficiency scores, these results are likely to be influenced by income and economic disparities that exist across municipalities. Municipalities in richer areas of the country have more scope for own revenue collections while other municipalities are more dependent on grants. Therefore, certain municipalities being over endowed with resources results in an inefficient use of these resources relative to other less-resourced municipalities and the demand for services in these jurisdictions.

However, the analysis also suggests that the equalisation nature of intergovernmental transfers, in particular the LES, has been rather unsuccessful. Equalisation grants are meant to address the mismatch between resource distribution and demand for services. Yet, certain municipalities appear to be overly endowed with resources resulting in an inefficient use of resources, while others are resource constrained. Therefore, the distributive nature of thse transfer systems could be improved.

Nevertheless, it is important to reiterate that municipalities are deemed efficient relative to the sample. If a different sample were used, many efficient municipalities identified would probably no longer be efficient. Therefore, these municipalities cannot be declared fully efficient, but merely more efficient than other municipalities in the sample.

Another shortcoming of FDH analysis is the ‘efficiency by default’ nature of observations, where they do not dominate nor are they dominated by other observations. It is possible that a municipality such as Thembisile could be efficient by default across the years.

116 Section 139 of the Constitution mandates the provincial executive to intervene in the affairs of a municipality in cases where the municipality cannot or does not fulfil its constitutional obligations.

8.8 Tobit Regression Analysis

8.8.1 Description

The next step of the analysis is to identify and estimate the specific factors that influence the efficiency scores. Although the analysis calculated FDH scores over a four-year period, this section will only estimate the factors impacting efficiency scores for the 2008/09 financial year because of a lack of frequent data available at local government level. A cross-sectional model will be used for the 129 municipalities. A major disadvantage of using this model for such analysis is that dynamic variables cannot be ideally captured.

8.8.2 Description of variables used and a priori expectations

Models of efficiency for South African municipalities should primarily focus on internal governance and capacity issues, which can include financial management, administrative and political capacity. To assess the ability to hold local government institutions accountable, control variables that capture the nature of the communities served in terms of levels of education and social and economic participation can also be added. In other words, the primary drivers of such efficiency are both municipal internal capacity and administrative arrangements and the general ability of communities to hold local government accountable and monitor performance. For example, fiscal capacity can assess both institutional arrangements, in terms of level of fiscal decentralisation, and the ability of communities to monitor municipal performance. The latter is justified by the assumption that communities are more alert to the spending of monies generated directly from their taxes. Generally, data constraints prevented a more robust look at the profile of communities.

Based on the argument above, four models were used that gradually increased the scope of factors that theoretically or practically had an impact on municipal efficiency levels. The first model considered intergovernmental factors and internal municipal capacity. These variables include the fiscal capacity (measured as local taxes as a percentage of total revenue), institutional capacity (total vacancies) and financial capacity (financial reporting capacity).[117]

For fiscal capacity, a positive effect on efficiency would be expected the more a municipality is funded by local taxes, as municipalities are likely to be more accountable to local constituents paying their taxes. However, the efficiency scores suggest that better-resourced municipalities use these funds inefficiently, which could be because of an over endowment of resources, higher quality of services or a larger community with poor information in how taxes are spent. In this respect, greater fiscal capacity can have a negative effect, which implies that the more resources a municipality has from constituents, the more likely it will spend it inefficiently.

The higher number of vacancies a municipality has implies administrative capacity constraints that are likely to make the municipality inefficient in its utilisation of resources in the delivery of services. One would expect a negative relationship in this regard. On financial capacity, a high capacity municipality in terms of its financial management[118] would likely have a positive effect on efficiency. It must also be pointed out that the use of this categorisation correlates well with financial planning, audit outcomes and the general capacity (in terms of quality and quantity of skills) of the finance department within the municipality. This variable captures most of these dynamics.[119]

The second model included political factors: the percentage of the council controlled by the majority party and the employment status of the councillors. The first variable can be interpreted in two ways. Firstly, a healthy political opposition improves the accountability and monitoring of the municipality's performance. Secondly, a municipality with a considerable majority may experience less friction or arguments in implementing policies and procedures, which could mean greater efficiency. It is also likely that a non-linear relationship exists here, with positive effects as the dominant party obtains more seats at a diminishing, and thereafter negative, rate. The nature of councillors within a municipality can also have an impact on municipal efficiency. The municipal council monitors and oversees the performance of the municipality and ensures accountability to its constituents. Full-time councillors are probably more effective in this role, which in turn ensures that the municipality is more efficient.

The model was extended to include variables of the municipal manager's profile; the municipal manager is the chief account-

117 As part of the Municipal Finance Management Act, No. 56 of 2003, municipalities are categorised based on their capacity to meet reporting requirements in terms of high, medium and low capacity. In this exercise, a dummy variable is used to capture high capacity municipalities in terms of their financial reporting with the low capacity as the reference variable.

118 See footnote 117.

119 Several indicators of financial capacity were used to derive the relevant categorisations. This included audit outcomes and capacity of the CFO.

ing officer of a municipality. The variables, which include experience, time in the post and educational qualifications, should have a positive impact on municipal efficiencies. A competent person leading a municipality improves performance. It is assumed that financial and other administrative capacity (for example the profiles of the CFO or other managers) are captured in the high and total vacancy variables described above.

The fully specified model included all the above variables and introduced service-specific variables, such as whether services were outsourced and the municipality's approach to the provision of FBS. The latter variable is unique to South Africa, as municipalities give a certain portion of free services to indigent households per month. It is important to assess the impacts of these different approaches on the provision of these services. These factors were included in the model as dummy variables. A municipality was given 1 if a specific service was outsourced and 0 if not. In terms of the different methods used to provide FBS, the technical approach was used as the reference base in determining the dummy variable relative to the other approaches.

8.8.3 Results

Table 8.6 illustrates the results of the four regressions while Annexure 8A provides the details and descriptive statistics of the variables used.

Table 8.6 Tobit regression results

Dependent variable: Municipal efficiency scores 2008/09				
Independent variables	Model 1	Model 2	Model 3	Model 4
Decentralisation	-0.752***	-0.600***	-0.642***	-0.600***
	(-6.61)	(-4.80)	(-5.22)	(-5.13)
High	0.310***	0.305***	0.292***	0.292***
	(6.80)	(6.88)	(6.68)	(6.85)
Vacantpostsgeneral	0.150	0.158	0.165	0.167
	(1.09)	(1.17)	(1.25)	(1.34)
constant	0.729***	0.363**	0.238	0.046
	(9.81)	(2.26)	(1.44)	(0.23)
Majority		0.379**	0.450***	0.462***
		(2.55)	(3.06)	(3.19)
Partimecouncil		0.015	0.031	0.043
		(0.24)	(0.50)	(0.71)
Mmexperience			0.005**	0.004**
			(2.38)	(2.12)
Mmtimeinpost			0.005	0.005
			(0.62)	(0.64)
Mmdiplomamore			0.035	0.014
			(0.87)	(0.36)
Wateroutsourced				-0.057
				(-0.86)
Electricityoutsourced				-0.001
				(-0.02)
Refuseoutsourced				0.278***
				(3.77)
Waterbroad				0.245*
				(1.71)
Watertarget				0.321**
				(2.18)

Source: Authors analysis

Table 8.6 Tobit regression results (continued)

Dependent variable: Municipal efficiency scores 2008/09				
Independent variables	Model 1	Model 2	Model 3	Model 4
Watergeographic				0.069 (0.25)
Electricitybroad				-0.219* (-1.77)
Electricitytarget				-0.241** (-2.02)
Electricitygeographic				0.680 (0.29)
Refusebroad				0.141 (1.06)
Refusetarget				0.095 (0.75)
Refusegeographic				-0.323 (-1.07)

Note: Absolute value of statistic in parentheses.
**Significant at 10% level*
*** Significant at 5% level*
**** Significant at 1% level*
Source: Authors analysis

The fiscal capacity variable is significant in all four models and has a negative impact on efficiency. In other words, the more fiscally capacitated a municipality, the more inefficient its use of resources. This goes against general theory, which expects that the larger the local taxes generated, the more local authorities are accountable to their communities. However, it confirms the analysis and distribution of inefficiencies across the municipalities. In the efficiency analysis, the inefficient use of resources increases the more resources a municipality generates in general. This could indicate the large economic disparities across the country, as certain municipalities generate more resources relative to the demand for services in their areas. This view suggests that the equalisation grants system could be improved to appropriately match the demand for services in municipalities relative to their fiscal capacities. Vacancy rates have no significant impact on efficiency.

Other important factors that have a positive impact on efficiency are the financial capacity and management within the municipality and the experience of the municipal manager. The financial capacity measure is statistically significant in all estimations and confirms that financial and administrative ability does have an impact on municipal performance. The time in the post and educational qualification of the municipal manager are statistically insignificant, whereas experience has a significant and positive effect on efficiency. This may be largely because of historical reasons and the general skills shortages in the country. Municipalities find it difficult to attract the appropriate skills and level of education, and so have to rely on the employee's experience. Historically, people with limited educational levels were given senior positions in municipalities and, with experience, ultimately improved their performance. The analysis confirms that municipalities should adhere to the minimum skills requirements for the employment of senior officials in municipalities, as stipulated in the Municipal Regulations on Minimum Competency Levels.

The employment status of the councillors (i.e. whether they are employed on a part-time or full-time basis) has no significant effect on efficiency. However, efficiency increases with the more seats the majority party has in the council. This suggests that a fragmented council made up of several parties can hinder the implementation of policies and procedures. A nonlinear relationship between the composition of the council and efficiency is possible, but this was not assessed.

Outsourcing or commercialising services can lead to efficiency gains. The analysis finds that efficiency is improved when refuse removal services are outsourced, but when water and electricity are outsourced, the impact on efficiencies is not significant. This suggests that municipalities tend to provide the refuse service inefficiently through overusing resources.

In 2002 government announced the FBS policies, as part of the national social security network. Local government provides a portion of services free every month to indigent households and municipalities have discretion over the methods they use to implement such policies. These methods were assessed against the efficiency scores generated to identify which method was the most efficient relative to the service provided.

It was found that the nature of the service informs the best practice mechanism used to provide FBS. For example, municipalities that used the technical approach for the provision of free basic electricity are the most efficient. The technical approach to free basic electricity includes the use of prepaid meters whereby a portion of the service is provided free of charge and thereafter the household has to buy if it wants to use extra. This proves more efficient as all other dummies have a negative effect on efficiencies or are insignificant. In terms of the targeted approach, the municipality needs to identify a household as indigent and then supply this household with a portion of electricity for free. In the broad based approach, all households receive a portion of electricity free, with households that use more than the stipulated free amount paying more to cross-subsidise poorer households. The geographic approach entails the identification of an area or community as poor and simply providing the entire community with free services. It is practical that the use of the technical approach in providing specifically, free basic electricity, gives the municipality more control and minimises losses in terms of households using the service and not paying for it. In the other methods, households can continue to use the service until it is disconnected, with the amount used being unpaid and thus a 'loss' in electricity and revenues. This analysis supports the use of the technical method in providing free basic electricity.

The targeted approach was identified as the most efficient in terms of providing water and sanitation functions. Unlike electricity, a municipality is legally not allowed to cut water to households, as it is a basic right. Therefore, the technical approach is inefficient in this case. The targeted approach ensures that households qualified to receive free services will receive these. The broad-based approach would result in a greater degree of inefficiency in all households receiving the services, because of a high possibility that richer households will receive free services if they use below the stipulated amounts i.e. there is an implicit and untested assumption that richer households consume more of a service.

The methods used to implement free basic refuse were all statistically insignificant. This suggests that there is little difference between them in terms of efficiency gains.

8.9 Conclusion and Recommendations

8.9.1 Conclusion

The chapter examined the technical efficiency of a sample of South African municipalities, based on their operating expenditure over a four-year period. It assessed how well municipalities transformed their inputs relative to the demand for services from their communities. Efficiency scores increased over the years but decreased in 2008/09, possibly due to the impacts of the recession. In general, the sample of municipalities could use 60–70% less resources and still maintain the current quantum of services provided. The following observations were made:

- Resources across the country are inequitably distributed, which results in certain municipalities having more resources relative to their service output requirements.
- Better-resourced municipalities use these resources inefficiently.
- This suggests that the distributive performance of equalisation grants, specifically the LES, can be improved.
- Quantity and quality of services can also have an impact on municipal inputs and outputs.
- The FDH technique can be used as a powerful tool to benchmark municipal performance relative to a set of similar municipalities. This identification process allows less efficient municipalities to learn from the experience and best practice methods employed by more efficient municipalities.

The chapter also used a tobit regression model to quantify the impacts of several factors on municipal efficiency, using the efficiency scores generated for 2008/09 as the dependent variable. This estimation produced the following results:

- The more resources a municipality generates from local taxes, the more inefficiently it uses its resources. This result highlights the skewed distribution of economic activity and resources across the country.
- The financial capacity of the municipality has a significant and positive impact on efficiency. This suggests that municipalities with the capacity to plan and monitor their finances better are more efficient at using resources.
- The capacity and skills of the municipal manager are also important in improving efficiency in municipalities. The experience of the municipal manager has a significant impact on efficiencies, which is probably due to historical reasons.

- The methods used in providing services and FBS also have a significant impact on municipal efficiency. Outsourcing the refuse function has positive effects on efficiency, while the technical approach for the provision of free basic electricity and the targeted approach for the provision of free basic water and sanitation were the most efficient.

8.9.2 Recommendations

Based on the key findings, the following recommendations are made:

- Municipalities should strive for technical and distributive efficiency to maximise their service level outputs with low average input costs, and to optimise the welfare of their communities.
- Municipalities should ensure that appropriately skilled and experienced people are assigned to senior administrative roles in municipalities. This is particularly important for the municipal manager and chief financial officer posts. The minimum requirements prescribed in Municipal Regulations on Minimum Competency Levels should be fully adhered to.
- Municipalities should consider outsourcing municipal services where efficiency gains are apparent, prioritising the refuse removal function.
- Municipalities should use the most efficient method relative to the nature of the service when providing FBS to households. This would mean using the technical approach for free basic electricity and the targeted approach for free basic water and sanitation.
- Poorly performing municipalities need to learn from the experience and best practice methods employed by municipalities that are able to ensure good performance in spending, efficiency in using resources and maximising outcomes.

Annexure 8A

Variables	Description	Mean	Max	Min	Stan Deviation
Efficiency	FDH efficiency scores generated for the 2008/09 financial year (%)	0.35	1.00	0.09	0.25
Decentralisation	Municipal fiscal capacity measured % of total revenues generated from local taxes	0.65	0.96	0.11	0.18
High	Dummy variable = 1 if municipality classified as HIGH capacity in terms of the MFMA financial capacity classification. Reference category: MED and LOW Categories				
Vacantpostsgeneral	% of vacant posts of total posts in organogram	0.14	0.58	0.00	0.14
Majority	% of council seats held by majority party	0.67	0.92	0.34	0.14
Partimecouncil	% of councillors that are part time	0.77	1.00	0.00	0.29
Mmexperience	The experience of the municipal manager in years	12.35	46.00	0.00	9.70
Mmtimeinpost	Number of years the municipal manager has been in the current post	2.09	10.00	0.00	2.19
Mmdiplomamore	A dummy variable = 1 if the municipal manager has a diploma, degree or post graduate degree with matric education or lower as the reference category				
Wateroutsourced	A dummy variable = 1 if the water and sanitation service is outsourced or commercialised with not outsourced being the reference category				
Electricityoutsourced	A dummy variable = 1 if the electricity service is outsourced or commercialised with not outsourced being the reference category				
Refuseoutsourced	A dummy variable = 1 if the refuse service is outsourced or commercialised with not outsourced being the reference category				
Waterbroad	A dummy variable = 1 if the municipality is using the broad based approach in the provision of free basic water and sanitation with the technical approach the reference category				

Watertarget	A dummy variable = 1 if the municipality is using the targeted approach in the provision of free basic water and sanitation with the technical approach the reference category				
Watergeographic	A dummy variable = 1 if the municipality is using the geographic approach in the provision of free basic water and sanitation with the technical approach the reference category				
Electricitybroad	A dummy variable = 1 if the municipality is using the broad based approach in the provision of free basic electricity with the technical approach the reference category				
Electricitytarget	A dummy variable = 1 if the municipality is using the targeted approach in the provision of free basic electricity with the technical approach the reference category				
Electricitygeographic	A dummy variable = 1 if the municipality is using the geographic approach in the provision of free basic electricity with the technical approach the reference category				
Refusebroad	A dummy variable = 1 if the municipality is using the broad based approach in the provision of free basic refuse with the technical approach the reference category				
Refusetarget	A dummy variable = 1 if the municipality is using the targeted approach in the provision of free basic refuse with the technical approach the reference category				
Refusegeographic	A dummy variable = 1 if the municipality is using the geographic approach in the provision of free basic refuse with the technical approach the reference category				

References

Afonso, A and Fernandes, S. 2003. *Efficiency of Local Government Spending: Evidence for the Lisbon Region*. Working paper 2003/09. Lisbon: Technical University of Lisbon.

Behr, A and Tente, S. 2008. *Stochastic Frontier Analysis by Means of Maximum Likelihood and the Method of Moments*. Deutsche Bundesbank Discussion Paper. No 19. Series 2: Banking and Financial Studies.

Boetti, L, Piacenza, M and Turati, G. 2009. *Fiscal Decentralisation and Spending Efficiency of Local Government – An Empirical Investigation on a Sample of Italian Municipalities*. University of Pavia: Società Italiana di Economia Pubblica.

Crawford, I, Klemm, A and Simpson, H. 2003. *Measuring Public Sector Efficiency*. Green Budget: United Kingdom.

Da Conceicao Sampaio de Sousa, M and Battaglin Schwengher, S. [Sa]. *Efficiency Estimates for Judicial Services in Brazil: Nonparametric FHD (Free Disposable Hull) and the Expected Order-M Scores for Rio Grande Do Sul Courts. Department de Economia:* Universidade de Brasilia.

Da Motta, RS and Moreira, A. 2009. Political factors and the efficiency of municipal expenditure in Brazil. *Economica Aplicada,* Volume 13(4): 353371.

De Borger, B and Kerstens, K. 1996. Radial and nonradial measures of technical efficiency: an empirical illustration for Belgian local governments using an FDH reference technology. *The Journal of Productivity Analysis*, Volume 7: 41–62.

De Borger, B, Kerstens, K, Moesen, W and Vanneste, J. 1994. A non-parametric free disposable hull (FDH) approach to technical efficiency: an illustration of the radial and graph efficiency measures and some sensitivity results. *Swiss Journal of Economics and Statistics*, Volume 130(4): 647–667.

Deprins, D, Simar, L and Tulkens, H. 1984. Measuring labour-efficiency in post offices. *The Performance of Public Enterprises: Concepts and Measurements.* Amsterdam: North-Holland. pp. 243–267.

Hardle, W and Jeong, S. 2005. Nonparametric Productivity Analysis. SFB 649 Discussion Paper 2005-013. Berlin: Universitat zu Berlin.

Leleu, H. 2006. A linear programming framework for free disposable hull technologies and cost functions: primal and dual models. *European Journal of Operational Research*. Volume 168: 340–344.

National Treasury. 2008. *Local Government Budgets and Expenditure Review: 2003/04–2009/10*. Government Printer: Pretoria.

Park, BU, Simar, L and Weiner, CH. 2000. The FDH estimator for productivity efficiency scores. *Econometric Theory*, Volume 16:855–977.

Silva Portela, MCA and Borges, CP. 2001. Non-Radial Efficiency Measurement: A Review and New Insights for Free Disposable Hull Technologies. Porto: Univerdidade Catolica Portuguesa.