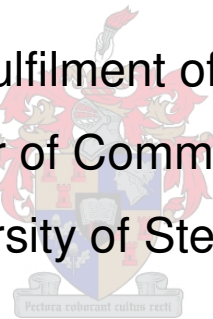


# **Key Efficiency and Equity Aspects of Providing Basic Local Services in South Africa**

**Len Verwey**

Thesis presented in fulfilment of the requirements for  
the degree of Master of Commerce (Economics) at  
the University of Stellenbosch



Supervisor: Professor Estian Calitz

March 2008

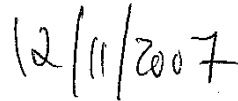
## Declaration

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature:

A handwritten signature in black ink, appearing to be 'WJ' with a long vertical stroke extending downwards.

Date:

A handwritten date in black ink, '12/11/2007', written in a slightly slanted, casual style.

## **Abstract**

In enquiring after the best means of financing basic local services in South Africa, the thesis begins by reviewing the arguments for fiscal decentralisation and the efficiency criteria for expenditure and revenue assignment. The role of local government within South Africa's system of intergovernmental fiscal relations is then evaluated. A chapter is devoted to the efficient pricing of infrastructure for household services. However, it is emphasised that such pricing is unlikely to meet equity criteria for access and affordability. The equity aspect of providing basic local services is explored further from the perspective of South African Constitutional obligations and the current basic services policy framework. A concluding chapter discusses issues arising out of the work and provides some recommendations.

## **Abstrak**

Die tesis vra na die beste manier om basiese plaaslike dienste in Suid Afrika te finansier. Dit begin met 'n oorsig van die argumente vir fiskale desentralisasie en die effektiwiteitskriteria vir uitgawe en inkomste toewysing. Die rol van plaaslike regering binne Suid Afrika se fiskale sisteem word dan ge-evalueer. 'n Hoofstuk word toegewy aan die effektiwiteitskriteria vir prysbepaling van infrastruktuur vir huishoudelike dienste. Pryse wat bepaal word slegs met effektiwiteit in gedagte sal onwaarskynlik huishoudlike dienste toeganklik en bekostigbaar maak vir 'n aanvaarbare persentasie van totale huishoude. Die regverdigheidsaspek van die voorsiening van basiese plaaslik dienste word dus verder bestudeer van die perspektief van Suid Afrika se konstitusionele verpligtinge en die huidige basiese dienste beleidsraamwerk. Die finale hoofstuk identifiseer kwessies vir verdere navorsing en maak ook 'n paar aanbevelings.

## **Acknowledgements**

My sincere thanks to my supervisor, Professor Estian Calitz, for sharing his time and experience so generously, and for his untiring willingness to offer suggestions which have made this thesis far better than would otherwise have been the case. Errors that remain are of course my own.

This thesis is dedicated to my parents.

## Table of Contents

<b>Chapter</b>		<b>Page</b>
	Table of Contents	<b>v</b>
	List of Figures	<b>vi</b>
	List of Tables	<b>ix</b>
<b>1.</b>	Introduction	<b>1</b>
<b>2.</b>	Fundamental Concepts of Fiscal Decentralisation	<b>4</b>
<b>3.</b>	Local Government and Intergovernmental Fiscal Relations in South Africa	<b>27</b>
<b>4.</b>	Efficiency and the Pricing of Infrastructure Services	<b>44</b>
<b>5.</b>	Access and Affordability of Household Infrastructure Services in South Africa	<b>80</b>
<b>6.</b>	Equity in Infrastructure Services: Current Policy and Financing	<b>103</b>
<b>7.</b>	Conclusion: Issues and Recommendations	<b>148</b>
	Bibliography	<b>160</b>

## List of Figures

	<b>Page</b>
Figure 3.1: Grants and Subsidies as Percentage of Capital and Operating Budgets, 2005/2006	<b>31</b>
Figure 3.2: Real Trends in Local Transfer Types (2006/2007 Rands)	<b>33</b>
Figure 3.3: Percentage of Nationally Raised Non-Interest Expenditure	<b>34</b>
Figure 3.4: Sources of Operating Income for Metros and Local Municipalities, 2004/2005	<b>36</b>
Figure 3.5: Trends in Sources of Capital Income for Metros and Local Municipalities, 2004/2005	<b>37</b>
Figure 3.6: Non-Performing Debt as Percentage of Total Debt, 2005/2006	<b>39</b>
Figure 3.7: Adequacy of Municipal Infrastructure for Local Services, 2005: Provincial Averages	<b>42</b>
Figure 4.1: Price and Quantity Curves for the Monopolist	<b>55</b>
Figure 4.2: A Basic Representation of Externality	<b>59</b>
Figure 4.3: Percentage Distribution of Total Electricity Sales, 2003	<b>75</b>

Figure 5.1:	Standard Utility-Possibility Curve	<b>85</b>
Figure 5.2:	Percentages of Households Using Public Tap as Main Water Source, by Province	<b>92</b>
Figure 6.1:	Municipal Estimates of Infrastructure Backlog Costs for FBS	<b>114</b>
Figure 6.2:	Percentage of Municipalities in each Province that Provide Free Basic Water and Electricity	<b>116</b>
Figure 6.3:	Free Basic Services: Percentage of Households Receiving in 2003	<b>117</b>
Figure 6.4:	Free Basic Services: Percentage of Households Receiving in 2005	<b>118</b>
Figure 6.5:	From GDP to Municipal LES	<b>127</b>
Figure 6.6	Budgeted Grants and Subsidies as Percentage of Total Budgeted Municipal Operating and Capital Revenue, 2005/2006	<b>133</b>
Figure 6.7:	Non-Grant and Non-Subsidy Capital Budget Revenue, 2005/2006	<b>135</b>
Figure 6.8:	A Representation of Cross-Subsidisation Structure Within a Single Service	<b>139</b>

Figure 6.9: A Representation of Cross-Subsidisation Structure Within a Single Service, Including Infrastructure Costs	<b>142</b>
Figure 6.10: Property Tax as Share of Small and Large Household Monthly Accounts	<b>146</b>



## List of Tables

Table 3.1:	Adequate Infrastructure for Household/Local Services per Province, 2005	<b>41</b>
Table 4.1:	Costs of Infrastructure Services	<b>51</b>
Table 4.2:	Costs Associated with Backlog Eradication	<b>52</b>
Table 4.3:	Financing Options in a Two-Level Fiscal System	<b>63</b>
Table 4.4:	Rivalry and Exclusion Characteristics of Different Kinds of Goods	<b>69</b>
Table 5.1:	Access to Piped Water in Dwelling or On-Site, 2005	<b>92</b>
Table 5.2:	Reasons for Non-Payment for Water Services	<b>94</b>
Table 5.3:	Provincial Percentage Distribution of Monthly Payments for Water, 2004	<b>98</b>
Table 5.4	Selected Water Supply Data for Three Western Cape Municipalities	<b>99</b>
Table 5.5:	Weighted Average Annual Cost of Expenditure on Water Services (2002 Prices)	<b>100</b>
Table 6.1:	Matrix of Local Revenue Possibilities	<b>104</b>

Table 6.2:	LES Basic Service Component Amounts	<b>128</b>
Table 6.3:	Budgeted and indicative MIG and LES amounts for 2007/08-2009/10 in current Rands	<b>132</b>

# Chapter 1

## Introduction

The aim of this study is to identify and investigate key issues in the provision of basic local services in South Africa from the perspectives of both efficiency and equity. Local governments are responsible for a number of services which are central to the alleviation of poverty and the attainment of adequate household living standards. In many cases, however, they are also characterised by infrastructural backlogs and small revenue bases. Although local revenue efforts could probably be sharpened in many cases, it also cannot be doubted that many of them quite reasonably require and will continue to require transfers from nationally raised revenue if they are to maintain themselves and extend the delivery of services. The need for an effective system of transfers is intensified if one considers the national policy requirement that all households be provided with a 'basic' amount of local utility services such as water and electricity, and that it should be provided free to households who cannot afford to pay for it.

The questions this study grapples with concern the way in which basic local service provision is to be financed and the relative extent to which efficiency and equity considerations should prevail. The efficiency approach to local service provision would entail charging identifiable users for local services in a way which adequately reflects the opportunity costs of their consumption. Strict application of any economic cost-recovery approach, however, would in all likelihood render such services unaffordable to many households. Efficiency considerations need to be balanced by equity-orientated attempts to provide at least basic service amounts to households free or at subsidised rates.

In engaging with these issues the study grapples with some of the theoretical aspects of efficient pricing and equitability, and describes and evaluates the

current policy context in South Africa. An effort is also made to provide a useful sketch of the current state of municipal finances and municipal service delivery.

Conclusions, where they are drawn, are tentative. There is no one-size-fits-all trade-off between efficiency and equity, nor is there one particular mix of transfers, user charges and local taxes which might best realise social objectives in all municipalities. Municipalities find themselves in widely differing circumstances and do not face identical challenges.

This study nevertheless attempts to contribute to the rigour with which questions of local service provision and financing are discussed and makes some suggestions for further research.

Chapter 2 establishes a benchmark model for the assignment of infrastructural expenditure functions and financing powers amongst levels of government under fiscal decentralisation. The model stems largely from a review of the fiscal decentralisation literature and from the criteria developed in that literature for determining which levels of government should be responsible for what, and on what grounds.

In Chapter 3 the system of intergovernmental fiscal relations in South Africa and the particular role of local government within this system is profiled with reference to the benchmark model. The chapter also discusses the current state of municipal finances and some aspects of non-financial performance.

In Chapter 4 the concept of efficiency is defined more explicitly. The conditions under which efficiency prevails in a market characterised by voluntary private transactions are reviewed, as are the arguments for government intervention where these conditions are not met. The particular economic aspects of infrastructure in general and household infrastructure particularly are then discussed. It is noted that two aspects of infrastructural services render them

susceptible to market failure and establish grounds for government provision or regulation, namely economies of scale and externalities. A discussion is provided of the arguments for pricing based on efficiency criteria, as well as the difficulties of such pricing. A final section concretises the preceding ones by presenting a mini 'case study' of electricity provision in South Africa.

Chapter 5 is concerned with the equity aspects of financing and pricing infrastructure services. Equity is defined and linked to the obligations imposed on government by the South African Constitution's Bill of Rights. Access information and the efficient pricing and financing discussion of the previous chapter are used to ask after the needed scope of government intervention in local service provision in South Africa. As in Chapter 4, the analysis is concretised through a mini 'case study': here water services are referred to.

From the base established in Chapters 4 and 5, Chapter 6 examines three specific equity-orientated measures currently being utilised in South Africa, namely the local equitable share (LES), the policy of free basic services (FBS) and the use of cross-subsidisation in particular instances.

Chapter 7 reviews and further elaborates on some of the key issues which arise from the discussions of the preceding chapters. Where possible, recommendations are provided. Issues discussed relate both to policy questions and to concerns around the quality of some of the information currently available to researchers.

## Chapter 2

# Fundamental Concepts of Fiscal Decentralisation

### a. Introduction

This chapter establishes a benchmark model for the assignment of infrastructural expenditure functions and financing powers amongst levels of government under fiscal decentralisation. It is regarded as a benchmark model because it stems largely from a review of the fiscal decentralisation literature and from the criteria developed in that literature for determining which levels of government should be responsible for what, and on what grounds.

In practice, assignment of expenditure functions may depart markedly from what theoretical efficiency-criteria might advise. A benchmark model certainly cannot account for the historical accidents and idiosyncrasies which may have led to a particular distribution of functions between levels of government in a particular country. Further, the reasons countries embark on fiscal decentralisation reforms are often *political* rather than economic. The fiscal system which results may consequently exhibit numerous inefficiencies in the generation or allocation of public resources.

However, the benchmark model clearly remains a necessary point of departure for considering and evaluating actual fiscal systems and the extent to which they are or are not likely to lead to enhanced welfare. Furthermore, the term 'local', (and by extension that of 'local infrastructure', 'local services' and the like) can only derive meaning from its position within a multi-levelled system of government. Thus part of the aim of developing a benchmark model of fiscal decentralisation is to develop a benchmark understanding of local government, and from that in turn to address more specific challenges around local services and especially their infrastructural component in South Africa.

## **b. Fiscal Decentralisation and the Objectives of Public Finance**

Fiscal decentralisation refers to a system of fiscal governance in which public expenditure and revenue decision-making authority is distributed across more than one sphere of government. Wallace Oates's (1973) *Fiscal Federalism* remains one of the clearest and most influential articulations of fiscal decentralisation theory, and follows Richard Musgrave's *The Theory of Public Finance* in citing three objectives which government involvement in the economy should aim to achieve, objectives which private market transactions are unlikely to achieve to a socially desirable extent. This assumption of 'market failure' constitutes the basic grounds for government intervention, though such intervention is justifiable only when plausible reasons exist to assume that the net welfare effect of government involvement will be positive, that is to say that government *failure* in the intervention will not eclipse government successes. Musgrave's three objectives are improvements in stabilisation, allocation and distribution (Musgrave 1959).<sup>1</sup>

The central question then, which Oates asks in providing a theoretical case for the benefits of decentralising public spending and decision-making, is which of these objectives are likely to be promoted by decentralisation, and why.<sup>2</sup>

The function of macroeconomic stabilisation may be associated with Keynesian counter-cyclical demand-side management of the economy through fiscal and

---

<sup>1</sup> In this study the terms stability, efficiency and equity will be preferred, though denoting the same concepts.

<sup>2</sup> The approach followed by Oates is largely consistent with the broader governance principle of 'subsidiarity', according to which public authority for a function should reside at the lowest level of political organisation capable of using it effectively. This principle represents a presumption in favour of non-centralisation and essentially proceeds by assigning all public functions to local governments unless theory or experience makes it clear that this is not feasible. Local assignment, in other words, is the 'default assignment'. This approach is also termed the 'layer cake' approach.

monetary policy, as well as with more modest attempts to ensure that key variables remain within pre-determined ranges, as in an inflation-targeting monetary regime. Oates has little doubt that the stabilisation function should be assigned to central government, rather than to sub-national government. Firstly, control of the money supply should be a centralised function so as to counter the temptation sub-national governments would have to print money rather than raise taxes as a means of raising revenue: failing this, each local unit would strive to be a free rider on other localities' prudent monetary policy. If it is the only one to print money rather than raise taxes, then there will be negligible increases in the rate of inflation and its members have become better off. But of course, as with any situation where the free rider temptation exists, many if not most municipalities will attempt this strategy: "It would clearly be in the interest of each municipality to finance its expenditures by creating money rather than by burdening its own constituents with taxation. The likely outcome would be rampant price inflation; for this reason, some form of centralized monetary control is imperative" (Oates 1973: 4).

In the case of whether or not to assign the stabilisation objective of fiscal policy to sub-national governments, the chief argument against it is that local economies are small and open, that is "their constituents typically purchase a large portion of the goods and services they consume from other localities" (Oates 1973: 4). It follows that the multiplier effects of deficit manipulation will be dispersed across the borders of the authority and are, in most cases, likely to have a fairly small impact on aggregate demand within the sub-national jurisdiction itself. A related objection to decentralisation of the stabilisation function concerns the fact that a cost of deficit-management which is not a feature of locating this function at the national level does come into play when it is assigned to a sub-national level. National debt-servicing transfers of funds are transfers between residents of a shared jurisdiction, whereas sub-national debt-servicing transfers would include a sizeable share of transfers from residents to non-residents of the sub-national jurisdiction. Lastly, although this point is not made by Oates, it may well be the



case that, in order for a central government to retain the ability to 'leverage' the economy (if the basic Keynesian premises in this regard are accepted), it needs to retain control of a fairly large percentage, that is a critical mass, of aggregate public spending and revenue.

Oates's work was published in 1973, shortly before the first OPEC oil-crisis and other political and economic factors which contributed to the 'stagflation'<sup>3</sup> of the mid- and late 1970s. The work not surprisingly reflects a general confidence in Keynesian fiscal policy prescriptions, a confidence which is no longer as widespread and which was partially eroded due to the inability of orthodox Keynesian prescriptions to deal with stagflation.<sup>4</sup> Stagflation led to increased scepticism regarding the extent to which inflation and unemployment could be 'traded off' against each other to achieve a socially optimal outcome. Both monetarists and the new classical school have questioned the ability of *any* level of government to effectively manipulate real economic variables such as output and employment to achieve policy ends.<sup>5</sup> This argument, if anything, strengthens the case against any association of sub-national governments with a stabilisation function.

As far as the distribution function and the objective of greater equity is concerned, the term 'equity' in the context of economic analysis refers to the normative question of a socially acceptable distribution of income. Those who advocate for government economic involvement on equity grounds assume that market transactions are unlikely to generate an acceptably equitable distribution of income and that public intervention in the form of a combination of targeted

---

<sup>3</sup> That is to say the simultaneous occurrence of high unemployment and high inflation rates in the industrialised Western economies.

<sup>4</sup> Current Keynesian and post-Keynesian economists would however argue that no other set of policy interventions, let alone a *laissez-faire* approach, would have fared any better given such large-scale exogenous shocks.

<sup>5</sup> *Monetarist, Keynesian and New Classical Economics* by JL Stein (1982) contains good discussions of the issues.

expenditure and progressive revenue policies can exhibit a redistributionary incidence.<sup>6</sup>

The 'benchmark' argument, in considering whether redistributive expenditure and revenue policies should be actively pursued at a sub-national level, asserts that such policies are likely to have a self-negating effect. For example, a progressive tax system accompanied by expenditure which targets the poor is regarded as one of the main policy options for pursuing redistributive objectives. If a local government were to pursue such objectives there would tend to be an inflow of the poor and an outflow of the rich, since we can assume a far greater mobility of production factors between local regions than between countries.<sup>7</sup> The demographic shift would erode the tax base of the local government whilst simultaneously increasing its expenditure needs, thus making it more difficult for the local government to achieve its objectives.<sup>8</sup> We would in fact see a clustering of the poor in certain areas and the rich in others and this would tend to return us to the need for a supra-local redistributionary authority.

It is worth emphasising that this benchmark model concerning redistribution is largely 'classical' in the sense that it rests on, amongst others, the key simplifying assumptions of perfect information and perfect mobility, as well as the assumption that sub-national fiscal policy *matters* as a potential determinant of locational decisions made by households. It assumes, in other words, that individuals know the direction in which and the extent to which their post-budget income diverges from their pre-budget one as a result of sub-national government spending and taxation, that they are able to compare such a

---

<sup>6</sup> This does not, however, mean that either expenditure or revenue incidence *in isolation* from the other needs to be redistributionary, only that their net effect is.

<sup>7</sup> Tiebout's (1956) seminal 'A Pure Theory of Local Expenditures' emphasises mobility as a key aspect of understanding local level equilibria and what determines them.

<sup>8</sup> It is not always noted, however (though Oates does) that this change would in fact be desirable if a municipality's objective were exclusively to reduce intra-jurisdictional *inequality*; however, if the objective were to reduce inequality whilst maintaining or increasing per capita income, it would not be achieved.

circumstance to that in other jurisdictions they may be considering as being more favourable, that they attribute significance to such divergences (i.e. they believe it truly does affect their wallets and/or quality of life) and that they are able to 'pack up and go' relatively easily, if not costlessly. The fact that, in the real world, it may be exceedingly difficult to determine the incidence of sub-national finances, that these are unlikely to differ significantly between jurisdictions and may or may not therefore have an impact on location decisions, and that households tend to be fairly immobile, should go some way to explaining why more sub-national redistribution may take place in practice than indicated by the benchmark model.

The strongest economic argument for fiscal decentralisation is the possibility that it can enhance Pareto-efficiency, that is to say that it can potentially increase the utility derived from public expenditure by some without having to reduce the utility enjoyed by others. Pareto-efficiency as a result of fiscal decentralisation is enhanced both through greater efficiency in resource allocation (allocative efficiency) and (arguably) greater efficiency in production (productive or technical efficiency). A case can arguably also be made for gains in dynamic efficiency. Allocative efficiency is discussed separately below and productive and dynamic efficiency in a subsequent section.

### **c. Fiscal Decentralisation and Allocative Efficiency**

The classical case for fiscal decentralisation is generally made with reference primarily to its ability to enhance the allocative efficiency of public expenditure, that is to say the extent to which the public goods and services offered and the manner in which they are financed are commensurate with consumer preferences, i.e. are those most desired by citizens within the relevant jurisdiction. In making the allocative efficiency argument, the classical case further assumes that a central government is only able to provide a standardised mix of goods and services to its citizens. If we assume that preferences may vary systematically by jurisdiction, then it follows that the amount of utility derived from this mix will also vary by jurisdiction (in some the mix will more closely

correspond to the ideal mix than in others) and therefore that utility could be enhanced by a closer, more customised fit between provision and preference. The assumption of sub-optimal social welfare stemming from centralised provision assumes that centralised provision will be characterised by “probable insensitivity to varying preferences among the residents of the different communities” (Oates 1973: 11). Similarly, Ebel & Yilmaz (2001: 12) note that “(d)ecentralization will increase efficiency because local governments have better information about their residents’ needs than the central government. Decisions about public expenditure that are made by a level of government that is closer and more responsive to a local constituency are more likely to reflect people’s choices than decisions made by a remote central government.” Enhancing welfare entails, then, that government be sufficiently close to communities to accurately identify such preferences and such locational differences in preferences: “A decentralized form of government therefore offers the promise of increasing economic efficiency by providing a range of outputs of certain public goods that corresponds more closely to the differing tastes of groups of consumers” (Oates 1973: 12). The greater the heterogeneity of inter-group preferences, the greater the potential gains in allocative efficiency from fiscal decentralisation.

The allocative argument also, and more subtly, rests on the assumption of imperfect information, or, what amounts to the same thing, of non-negligible transaction costs in the acquiring of information. In the absence of perfect information, and given large costs in acquiring the relevant information pertaining to differences in preferences between jurisdictions, a central government is better off devolving authority to a sub-national body which is able, at less cost, to determine preferences and supply conditions and customise its mix of goods and services accordingly. This perspective acknowledges the informational advantage sub-national governments have over central government in their greater familiarity with the cost-conditions and cost-drivers of service-provision within their own jurisdictions.

However, the allocative efficiency argument applies only if the population sub-sets for which the public mix is customised differ *systematically* from each other. That is to say, if sub-sets are representative of preferences for the population as a whole, then no allocative gains from decentralisation are possible. Also, a country characterised by extreme heterogeneity of preferences but which are distributed randomly across jurisdictions, would not see welfare enhancement from fiscal decentralisation due to greater allocative efficiency. Imagine, for example, a country divided into two jurisdictions. Imagine further that the population speaks one of two languages, M and N, and that the policy issue that is being considered is that of mother tongue secondary education. Allocative gains associated with decentralisation would firstly depend on the extent to which residents have strong preferences: obviously if residents don't really care in what language their children receive tuition then no costs should be incurred to offer schooling in two languages. When residents have strong preferences, this does not automatically mean that allocative efficiency gains are to be derived from decentralisation. If equal numbers of M-speaking and N-speaking residents live in each of the two jurisdictions, then any service customisation which benefits one sub-set of the population will be to the detriment of the other sub-set living in that jurisdiction. Gains only become possible if the preferences are concentrated jurisdictionally, that is for example if M-speakers significantly outnumber N-speakers in the one jurisdiction and vice versa in the other jurisdiction. The minority group will lose in each jurisdiction, but the larger group of residents benefiting in each jurisdiction means total welfare is enhanced.

#### **d. Fiscal Decentralisation, Productive Efficiency and Dynamic Efficiency**

Although the allocative efficiency argument is probably the most often-cited economic<sup>9</sup> argument for fiscal decentralisation, two other efficiency-related

---

<sup>9</sup> There may of course also be political arguments.

arguments should be mentioned, namely that fiscal decentralisation may enhance productive efficiency<sup>10</sup> and that it may enhance dynamic efficiency.

Productive efficiency is understood here as the minimisation of waste, that is the attainment of the maximum possible output with the least input. It may be argued that fiscal decentralisation strengthens democratic accountability and that heightened accountability makes it more difficult for public authorities to use resources inefficiently. Fiscal decentralisation may enhance accountability because, in theory, it allows more participatory opportunities to citizens and communities, and because decision-makers are located closer to the communities they are supposed to serve.<sup>10</sup>

Fiscal decentralisation may also enhance productive efficiency in the delivery of services because it can be assumed that local governments have a better sense of the environmental constraints and cost-structures of delivery in their jurisdiction than a central government.<sup>11</sup> They may also be better placed to develop *innovative* solutions to problems of service-delivery and financing which emanate from that environment. A local government may proceed to some extent by trial and error, knowing that failure would be unlikely to have catastrophic consequences, whereas failure on a national level might well. In this sense public fiscal decentralisation may have some of the merits associated with private, decentralised decision-making and individual entrepreneurship which render a mixed economy more innovative than a state-owned one. The successful

---

<sup>10</sup> Barankay & Lockwood (2006) is a useful analysis of the extent to which fiscal decentralisation may lead to maximum outputs for a given set of inputs or, what amounts to the same thing, ensuring that a given output is attained with a minimum of inputs.

<sup>10</sup> See Shah (1999).

<sup>11</sup> For example, assume a national campaign is launched requiring the phase-in of renewable energy sources in government operations. A local government would possibly have a more nuanced sense of the extent to which solar power, wind power, biofuels would be feasible options within its jurisdiction, assuming it had an appropriate degree of general technical knowledge and expertise.

innovations of one local jurisdiction become available to other, similar jurisdictions and can enhance the cost-efficiency of the system as a whole.

Care should be taken, however, not to present an overly rosy account of this alleged efficiency benefit. Prud'Homme (1995) demystifies some of the optimism relating to participatory governance and improved public performance, particularly under conditions which are prevalent in many developing countries. He points out, for example, that local governments are vulnerable to capture by local elites and may be less transparent than central governments. Local governments may also have a shortage of administrative capacity and the local budgeting process may not be particularly responsive to the demands of citizens. Prud'Homme points out, in further questioning the allocative efficiency grounds for decentralisation, that voters may not vote according to their individual preferences for goods and services, but according to loyalties. Further, "the platforms on which local elections are fought (when they exist) are often vague and unrealistic" (Prud'Homme 1995: 208).

Fiscal decentralisation has at times been advocated as a means of countering the 'leviathan' tendencies of a state which taxes and spends beyond its optimal size. The classical articulation of this argument is probably that of Brennan and Buchanan (1980), who argue that interjurisdictional competition under fiscal decentralisation will compel local governments to use resources efficiently and set tax rates and bases in ways that minimize distortions and create incentives for industrial expansion. If it is assumed that governments tend inexorably and inevitably towards over-expansion at the expense of the private sector, unless checked by a range of formal mechanisms, then this argument may have validity. Constrained by the need to consider what other local governments are doing in order to attract business, local authorities are more likely to engage in strategic behaviour which would presumably include lower tax rates and more incentives (that is negative taxes) than would otherwise be the case.

Bahl and Linn (1992) provide a further argument for fiscal decentralisation, which they in fact regard as the *strongest* argument in a developing-country context, namely that overall resource mobilisation will be increased. They argue that “local governments can tax the fast-growing parts of their economic base more easily than can the central government” (Bahl and Linn 1992: 386). Because this argument emphasises the extent to which public resource-growth is able to keep pace with a fast-growing (and typically structurally transforming) developing-country economy, it is referred to as a ‘dynamic efficiency’ argument. The authors argue that “(a)s the economies of rural areas and secondary cities develop, their taxable capacity and willingness to purchase public services will also develop. It will be very difficult for central governments to capture much of this fiscal surplus because neither central government income nor consumption taxes typically reach small firms, workers in small firms or outside the large cities” (Bahl and Linn 1992: 386, 387).

The dynamic efficiency argument may, in practice, be constrained by various factors. Firstly, there is the extent to which local governments have taxation powers and administrative capacity at their disposal which allow them to capture such surpluses. A centralised system of revenue collection which collects *some* of this fast-growing revenue remains preferable to the more ambitious assignment of revenue-generating instruments to local governments which are unable or unwilling to use them efficiently and fairly. In fact, care needs to be taken that such tax opportunities do not provide local governments with little more than a *carte blanche* to strike at anything that moves. A fine line exists between local resource generation which keeps pace with entrepreneurial spurts and one that stifles such growth by taxing away the profits of emerging small businesses whilst not putting such additional resources to particularly productive use.<sup>12</sup>

---

<sup>12</sup> The risk of a local proliferation of ‘nuisance taxes’ is intensified where national fiscal tightening or other developments exogenous to the local government have reduced the transfer amounts it can count on and it therefore turns to the local ‘captive’ community. In an analysis of intergovernmental fiscal relations in the



### **e. Assigning Powers and Functions on Efficiency Grounds**

Under fiscal decentralisation, the assignment problem is essentially that of who should do what. Given the many functions that modern governments are called on to fulfill, what is needed is a clear theoretical basis for assigning expenditure functions and revenue powers to the various levels. In theory, the ideal sequence of functional assignment should be to first assign expenditure functions and then assign revenue powers and establish a system of grants to ensure adequate resources for the carrying out of functions assigned to the various levels of government. McLure and Martinez-Vazquez (2003), reviewing the Latin American experience, argue that this has not necessarily been the case in practice:

“Designing the other important pieces of a system of decentralized finances, revenue assignments and transfers, in the absence of a clear expenditure assignment is to put the cart before the horse. The decentralization movement in many countries of Latin America over the past decade made this fundamental mistake. Revenues were assigned to sub-national governments and transfers put into place before it was decided what functional competencies would be transferred from the central government to sub-national governments.” (McLure and Martinez-Vasquez 2003: 2). Countering this, they express an efficiency notion based on what may be called the spatial or jurisdictional incidence of costs and benefits: “Responsibility for the provision of services should be at the lowest

---

developing world, Fjeldstad devoted some time to the Tanzanian local tax system. He cites a previous study which found that “(i)n one council...the by-law on hawking and street-trading specifies in detail 38 different components (including license for bicycle repairs, tyre puncture repairs, shoe shiners, car washes, carpenters, firewood, potato chips sellers, etc” (Fjeldstad in Levy & Tapscott (eds) 2001:150). He concludes that, at any rate in Tanzania at the time (though one assumes it is a more prevalent issue), “(l)ocal governments seem to raise whatever taxes they are capable of raising” (Fjeldstad in Levy & Tapscott (eds) 2001: 159) and that basic economic criteria are frequently left out. The Kibaha district, for example, levies a high tax rate on cashew nuts, which are an export crop (Fjeldstad in Levy & Tapscott (eds) 2001: 159).

level of government compatible with the size of the 'benefit area'...Efficiency...is enhanced if consumption benefits are linked to costs of provision via fees, service charges, or local taxes" (McLure and Martinez-Vasquez 2003:5).

The concept of *perfect correspondence* refers to a situation where "there exists a level of government for each subset of the population over which the consumption of a public good is defined (Oates 1973: 34). This conception implies that the distinction between 'public' goods and 'private' goods is more one of differentiation along a benefit-spectrum than absolute: a 'private' good is simply one for which the benefits are consumed by a jurisdiction consisting of only one person; the benefits of a pure public good in the traditional sense on the other hand are distributed to all citizens in a country. In fact, as Stiglitz (1982) usefully suggests in his review of the Tiebout hypothesis,<sup>13</sup> what we have are goods that may be characterised as 'public' or 'private' on the basis of the spatial framework from which we regard them: thus, for example, local public goods are goods "the benefits of which accrue only to those who belong to a particular group...and not to those who belong to other groups...within the society. There is thus an element of "privateness" in local public goods; while *within* the community the good is a pure public good, "between" communities it acts like a private good" (Stiglitz 1982: 3,4). Oates's perfect correspondence similarly implies an infinite number of differently-sized communities who are regarded as communities for the purposes of determining optimal size for a *particular* service. They in fact internalise all the costs and benefits: there are no externality effects to their actions.

Regarded in another way, this argument implies that for every service there is a potentially different optimal trade-off between the gains associated with small jurisdictions and forgone benefits (or opportunity costs) associated with economies of scale and with obtaining information on the preferences of citizens.

---

<sup>13</sup> Discussed below.

Thus a different *optimal* jurisdiction size exists for each service when it is regarded in isolation from other services.<sup>14</sup>

## **f. Preference Revelation: Mobility and Service Charges**

Allocative efficiency is only increased under fiscal decentralisation if resident preferences do indeed vary systematically by jurisdiction. A key requirement of the argument in favour of decentralisation is therefore that means exist to *determine* the preferences of residents as far as the quantity and mix of local public or semi-public goods and services are concerned.<sup>15</sup> If local preferences cannot be determined then no customisation of goods and services is possible. Two sets of arguments have developed in the fiscal decentralisation literature regarding preference revelation, those associated with the Tiebout hypothesis, and those associated with the efficient use of service charges<sup>16</sup> within a jurisdiction to align preference and provision within a balanced-budget local system. Tiebout focuses on local *public* goods, while service charges are best discussed in the context of semi-public goods.

Tiebout, in his 'A Pure Theory of Local Expenditures', begins by noting that a core problem of public economics concerns "the mechanism by which consumer-voters register their preferences for public goods" (Tiebout 1956: 417). Where there is no rivalry in the consumption of a good, in other words where one person's consumption does not reduce the quantity available to another, and therefore where no 'competition' exists between consumers for a good, the price-mechanism cannot equilibrate supply and demand at an efficient equilibrium. The problem is fundamentally that of the 'free rider'. One can imagine the market for a private good, that is a good which is rival in consumption and from which one can be excluded, as approximating an auction: individuals need to bid, that is reveal

---

<sup>14</sup> The ongoing debate and controversy in South Africa around regional electricity distributors (RED's) is partly about the recognition that for some services existing assignments are simply too inefficient.

<sup>15</sup> Semi-public goods are public goods that are either rival but not exclusionary or exclusionary but not rival.

<sup>16</sup> I use the term 'service charges' to cover tariffs, levies, fees and the like.

the monetary value they assign to the good, if they are to stand a chance of deriving utility from consuming the good. Once they own this good, furthermore, they are legally entitled to exclude others from deriving utility from it. One can further imagine an 'auction' for a good that everyone will benefit from but where only those who formally assert that they will in fact benefit are required to pay up for it. Even if social pressure worked against consumer-voters denying that they in fact get *any* benefit whatsoever from a particular public good or service, it is clear that an incentive would exist to understate the extent of preferences for a good or service in the hope that others will reveal theirs accurately and one can enjoy the benefits without contributing proportionately to the costs of provision. The free rider problem applies wherever public goods are funded according to the benefit principle: the free rider hopes that, by understating her or his benefit, (s)he can shift the funding burden.

In a centralised fiscal system, according to Tiebout, no satisfactory mechanism exists to firstly "force the voter to reveal his preferences, [then] satisfy them in the same sense that a private goods market does; and [finally] tax him accordingly" (Tiebout 1956: 417, 418). It is Tiebout's fundamental hypothesis that, for a significant number of public goods and services, decentralisation enables preference revelation and therefore efficient provision. Tiebout argues that the existence of multiple governments in a given sub-national sphere, when coupled with the assumptions of perfect mobility and perfect information, constitute a quasi-market mechanism which effects preference revelation since in fact households can 'shop around' for the jurisdiction whose mix of good and services they prefer.<sup>17</sup> They 'vote with their feet', an option which is more readily available in selecting a local jurisdiction than, for example, in relocating between countries: "The act of moving or failing to move is crucial. Moving or failing to move

---

<sup>17</sup> The three primary assumptions discussed here are those of sufficient choice, mobility and information. Tiebout in fact lists and discusses seven necessary assumptions as well as the implications of relaxing these.

replaces the usual market test of willingness to buy a good and reveals the consumer-voter's demand for public goods" (Tiebout 1956: 420).

It need not even be the case that local government finances are a significant determinant of location decisions, only that they *matter* at the margin. The Tiebout hypothesis would still hold where households are indifferent between the public financial circumstances of a number of jurisdictions, in the same way that Pareto-optimality is achieved in private exchanges as long as exchanges take place up the point where no more welfare-enhancing ones are available. For Tiebout, mobility of households, and competition between jurisdictions for households, performs the same function as the existence of competition in the private market in ensuring Pareto-optimality and 'solves', at the local level, one aspect of the problem of efficient public provision. The more jurisdictions, and the greater their preference variety, the better the chances that each household or individual will live in their 'ideal' community as far as public goods and services are concerned.

The Tiebout hypothesis can to some extent be extended to semi-public<sup>18</sup> local goods also. It can, for example, be assumed that households base their location decision on the combination of public goods and semi-public goods available in a given jurisdiction. However, it is not even necessary, in the case of many semi-public local goods, to evoke the Tiebout mechanism. The very fact that they are not pure public goods means that exclusion becomes possible and that price can, to some extent, be set on the basis of marginal cost pricing. This introduces the second argument on preference revelation mentioned earlier, a matter to be discussed in further detail in subsequent chapters.

## **g. Constraints to Fiscal Decentralisation**

The gains from decentralisation are not costless, but involve opportunity costs which serve as a constraint to the assignment of all expenditure functions and

---

<sup>18</sup> That is goods characterised by *either* rivalry in consumption or the possibility of exclusion.

revenue powers to sub-national governments in the name of allocative efficiency. Key opportunity costs are firstly the earlier mentioned benefits foregone as a result of losing the economies of scale associated with uniform provision. This includes administrative economies of scale, since additional public resources are required for the administration of a multi-level system of governance. Secondly, many public and semi-public goods are characterised by externalities in the non perfect-correspondence scenario, which may require a complex combination of multi-sphere government with a system of transfers for the sake of efficiency as well as equity.

In addition to what may be regarded as *inherent* constraints to the extent of decentralisation, there are also challenges associated with particular external environments and institutional arrangements. Particularly in the context of developing country decentralisation reforms, a lack of adequate administrative capacity in sub-national governments and a lack of responsiveness of budgeting to the differing preferences of jurisdictions seem common challenges.<sup>19</sup>

In general, once these general constraints and challenges are acknowledged, a context is established for debate on the optimal extent of decentralisation in a particular country-context, as well as the *sequencing* of such reforms. In this section we look in more detail at externalities, economies of scale, inadequate administrative capacity and budgetary inflexibility.

Opportunity costs associated with foregone economies of scale stemming from centralised provision mean that the trade-offs between allocative efficiency gains and economy of scale losses need to be considered explicitly. An example is that of a country with a number of mother tongues, varying by region, where residents of each region would prefer for their children to receive primary school mother tongue instruction. The allocative efficiency gains associated with having sub-national governments which can correctly identify and respond to these

---

<sup>19</sup> Momoniat (2001) is a good review of the South African context in this regard.

preferences is clear.<sup>20</sup> On the other hand, foregone economies of scale will almost certainly raise the cost of education, which means less additional public resources are available to promote welfare by other means. In this example there would likely be economies of scale in, amongst other things, the centralised development and provision of learner materials, training of teachers, and the setting and marking of exams. It should also be clear that a national department of education will now need to incur costs merely to acquire and maintain information concerning what the various regional departments are doing, and also that money will have to be spent in transmitting nationally formulated education policy to all the regions in a manner which ensures implementation.<sup>21</sup>

The real issue, consequently, is not to argue in all instances for decentralised assignment because of improved allocative efficiency, but requires a case-by-case consideration of the trade-offs between the welfare gains associated with a closer preference-to-provision fit and the potential welfare losses associated with increased administrative costs and decreased economies of scale. In fact, since one can regard the reduced administrative costs associated with centralisation as an economy of scale benefit too, the trade-off is really between the closer fit and the decreased economies of scale.

Externalities are another aspect of the assignment question which needs to be considered. Externalities are the costs or benefits arising from voluntary, mutually beneficial transactions which are not considered and factored in by the units engaging in such transactions because they affect third parties rather than the transacting parties themselves. If one assumes that the price mechanism regulates transactions in the private sector by indicating the scarcity of some goods relative to others, and the desirability of some goods relative to others in the eyes of rational utility-maximising consumers, then externalities are those

---

<sup>20</sup> See also the earlier discussed example regarding extreme preferences in Section 1c.

<sup>21</sup> The arguments developed here are *economic* only; many political arguments for decentralisation would in many cases convincingly tip the scales where the economic arguments may not be as conclusive.

aspects of preference and scarcity which are not reflected in the price and which, if they were reflected, would in fact cause the price to be different from what it is. They can consequently be regarded as a measure of the price adjustment required to bring private costs and benefits in line with social costs and benefits.<sup>22</sup> In the case of negative externalities, social costs exceed private costs and social welfare would be increased by measures which reduce consumption or production. In the case of positive externalities social benefits exceed private benefits and social welfare would be increased by measures which increase consumption or production.

Although externalities are often associated with private production and consumption, and governments are typically called upon to address externalities since their role is generally conceived as that of representing those interests which are social or collective rather than individual, two qualifications to this view should be kept in mind. Firstly, governments are not necessarily in a position or able to accurately determine the extent of an externality and take the measures required to bring social and private calculations into line. Secondly, and especially pertinently for the discussion that follows, governments are not only called on to address externalities (that is internalise them so actors face true social curves in making their production and consumption decisions), but their actions also generate further externality effects which may be positive or negative, intentional or inadvertent.

Within a fiscally decentralised system, 'interjurisdictional externalities' are costs and benefits of public expenditure and revenue in a particular jurisdiction which affect the utility functions of residents of other jurisdictions positively or negatively. They therefore represent inefficiency in public spending since production and consumption decisions within a particular jurisdiction are made based on 'false' community supply and demand curves. Externalities suggest that a system of fiscal decentralisation should be 'impure' and combine decentralised

---

<sup>22</sup> Hyman (2005), chapter 3, is a good review of the theory of externalities.



revenue powers with a system of intergovernmental transfers. That is, on efficiency grounds sub-national governments should not be required to be fiscally autonomous and fund all their expenditures through self-generated revenues, but should receive transfers from higher levels of government to compensate for the positive externalities of their expenditure.

## **h. Intergovernmental Fiscal Relations**

Fiscal decentralisation generally requires a system of resource transfers between levels of government.<sup>23</sup> There are at least three key reasons for this. Firstly, virtually all fiscally decentralised systems are characterised by a revenue-expenditure mismatch: the expenditure functions assigned to a sub-national government on efficiency grounds cannot be funded by the revenue powers assigned to that government on the same efficiency grounds, since revenue powers tend to be centralised for a mix of political and economy of scale reasons. Secondly, as discussed previously, much sub-national spending generates externalities, usually positive ones. This means that such spending needs to be subsidised by higher levels of government if a socially optimal amount is to be spent, that is if the marginal social benefit rather than the marginal 'private' (jurisdictional) benefit is to determine how much is produced.<sup>24</sup> In the case of revenue-expenditure mismatches, this also creates problems of *horizontal* (that is inter-jurisdictional) equity. The degree of mismatch will typically vary between richer and poorer local governments, and transfers may be used as a form of interjurisdictional redistribution aimed at attaining a greater degree of convergence in sub-national financial circumstances. In some cases greater

---

<sup>23</sup> Fundamentally a choice exists between local *transfers* and the alternative assignment of greater local taxation powers. For a combination of political (i.e. the desire to retain central control) and economic reasons the preference in most cases is for the use of transfers, with most major taxes retained at the national level. The question of what percentage should be shared is structurally identical to the question of how the transfer amount should be determined, which is discussed in the main text.

<sup>24</sup> There is very little reference in the literature to the converse case of intergovernmental fiscal relations and *negative* externalities. Presumably the political challenges of withholding transfers from a jurisdiction whose operations were generating such effects would be very difficult to overcome.

equitability may explicitly be aimed at, as in South Africa's 'local equitable shares', whilst in others this is a necessary consequence of transfers which aim at ensuring minimum living levels or the like in all jurisdictions. In the discussion below this case is regarded as one aspect of the broader use of transfers to address fiscal gaps.

Although in practice the distinction is one of degree rather than kind, it is conceptually useful to distinguish clearly between conditional transfers and unconditional transfers. Conditional transfers<sup>25</sup> are transfers where "the transferring government specifies the purpose, conditions, or both, under which the recipient government should use the grants" (FFC 2006: 4). Unconditional grants, on the other hand, come with 'no strings attached' and the recipient government can "employ the grants according to its own set of priorities" (Oates 1972: 65). Conditional grants arguably represent a vehicle for retaining *de facto* control over public expenditure at the center, since they are used to promote local spending on national or at least supra-local rather than purely local objectives. The risk always exists, therefore, that they can come to gradually eclipse local fiscal autonomy if the share of conditional rather than unconditional transfers in total transfers increases over time.<sup>26</sup> The Financial and Fiscal Commission (FFC) alludes to this, and also identifies a further problem associated with conditional grants, namely the administrative burden which may accompany them: "the increased use of conditional grants may be interpreted by sub-national governments as an intrusion by national government in their areas of competence. This is especially true when conditional grants are relatively large in proportion to total sub-national revenues. The reporting requirements on conditional grants can also be strict and tend to impose an extra burden on sub-national governments" (FFC 2006: 6).

---

<sup>25</sup> Also referred to as conditional, specific purpose, categorical or earmarked transfers and as grants rather than transfers.

<sup>26</sup> A similar risk is associated with circumstances where the ratio of transfers to local revenue increases over time.

As mentioned above, transfers can be used to address a revenue-expenditure mismatch as well as to compensate municipalities for positive externalities. An expenditure-revenue mismatch typically characterises the financial circumstances of sub-national governments. This is also known as the problem of vertical imbalance. In theory transfers should ‘close the fiscal gap’, that is they should represent the difference between expenditure needs and revenue capacity, constrained by the overall resources available to the economy:  $\text{Fiscal Gap} = \text{Expenditure Needs} - \text{Revenue Capacity}$ . A simple specification of total transfers for this purpose might look like this:  $\text{Total Transfer Amount} = (\text{Fiscal Gap}) * \text{Jurisdictions} * (\text{local share of revenue}/\text{total revenue})$ . Such a specification, in constraining transfers to local governments by total revenue, takes into consideration total fiscal capacity.

The problem with the use of transfers to close the fiscal gap is essentially that it may be difficult to quantify both expenditure needs and revenue capacity. This requires establishing a “minimum service level”, cost the expenditure need on this basis, in a manner informed as far as is efficiently possible with a knowledge of local circumstances, perhaps using an indicator such as the tax base, and calculate the transfer based on this. Such an approach would of course entail a degree of interjurisdictional redistribution since the fiscal gap would be larger, all else being equal, in poorer municipalities with less revenue and greater demand for publicly provided services. The indicators used to assess the fiscal gap also need to be beyond manipulation by the local government. Where this is not done so-called perverse incentives are created for inefficient local spending and sub-optimal tax effort. However, this does not mean that, in most developing and developed countries, available data readily allows an *accurate* determination of the fiscal gap.

Conditional transfers are probably best suited for encouraging socially optimal expenditure by local governments on goods with positive externality effects.

Conditional grants can be either matching or lump-sum grants. A matching grant requires “the recipient government to match the contribution by the transferring government” (FFC 2006: 4). The relative share of expenditure, that is the extent to which expenditure is funded by the transfer relative to local own revenue, should reflect the value of the externalities relative to the benefits accruing to residents of the jurisdiction. In practice, this is of course difficult to determine. As the FFC study emphasises, in addition to information problems, matching grants may also create perverse incentives under certain circumstances: “(T)he implementation of the optimal matching requirement for horizontal spillovers is difficult to measure. Matching grants justified on the grounds of spillovers can introduce adverse effects into the transfer system with provinces chasing ‘cheap money’” (FFC 2006: 5).

In the case of a conditional lump sum transfer, no expenditure matching requirement is imposed on the sub-national government, with the only requirement consequently being that the grant is used for the agreed purpose, including the reporting requirements associated with this. The FFC argues that such grants can achieve most of the objectives of matching grants, though without the adverse effects of these. However, Oates argues that lump sum grants in fact amount only to an income substitution effect, unless such a grant is for spending on items which would not have been purchased at all in the absence of the grant. In other words, the jurisdiction will simply allocate its own funds elsewhere and there is no reason to assume that a socially optimal amount of expenditure on the externality good will result. For this reason Oates argues that there is little real difference between a lump-sum conditional grant and an unconditional grant (1972: 77).

# **Chapter 3**

## **Local Government and Intergovernmental Fiscal Relations in South Africa**

### **a. Introduction**

The preceding chapter sketched out a benchmark model of expenditure and revenue assignment under fiscal decentralisation. Various challenges associated with assignment under the benchmark model were also discussed. In this chapter the system of intergovernmental fiscal relations in South Africa and the particular role of local government within this system is profiled with reference to the benchmark model. The chapter also discusses the current state of municipal finances and some aspects of non-financial performance.

### **b. The Constitutional Context of Local Government in South Africa**

The South African Constitution provides for three spheres of government, the national, provincial and local, and describes them as being “distinctive, interdependent and interrelated” (Constitution of South Africa, Act 108 of 1996: Section 3.40 (1)). The term ‘spheres’ is intended to denote a cooperative rather than more strictly hierarchical relationship between national, provincial and local governments. Chapter 3 of the Constitution is entitled ‘Co-operative Government’ and sets down a range of guidelines and principles concerning cooperation between the spheres: they are required to “co-operate with one another in mutual trust and good faith” in order to, amongst others, “secure the well-being of the people of the Republic” and “provide effective, transparent, accountable and coherent government for the Republic as a whole”. Public functions are classified as being either ‘concurrent’, or ‘exclusive’, with exclusive ones being assigned to one sphere of government only and concurrent ones being a shared responsibility of more than one sphere of government. Concurrency can

potentially enhance efficiency since different aspects of a broad function such as health can be assigned to different spheres on the basis of the various considerations outlined in the benchmark model. Conversely, such an arrangement can also be inefficient if it does not generate a clear assignment of functions.

The role of national and provincial government in the functioning of local government is set out quite clearly in section 155 (6) (a) and (7) of the Constitution as being to “provide for the monitoring and support of local government in the province” and to “promote the development of local government capacity to enable municipalities to perform their functions and manage their own affairs”. In addition, both national and provincial government are granted the legislative and executive authority to “see to the effective performance by municipalities of their functions...by regulating the exercise by municipalities of their executive authority”.<sup>27</sup> This conception, in other words, differs from a simple ‘layer cake’ approach which assigns functions simply on an efficient-jurisdiction basis. One way of conceiving the difference is that in the Constitutional articulation it is firstly recognised that local governments may be at an administrative disadvantage and may therefore require support and even monitoring by national and provincial government, and secondly that the responsibility of national and provincial governments requires that they do ensure that local government policies and administration do not in fact hamper or jeopardise national or provincial objectives. If anything, the a priori trend appears to be towards the unionist rather than the federalist state.

Although the sharing of functions (such as between national and provincial government) can enhance efficiency, it can also create problems of

---

<sup>27</sup> Schedule 4 part B and Schedule 5 Part B of the Constitution set out the local government functions which are exclusive to it, subject to the above requirement that national or provincial government may intervene where performance is ineffective. Schedule 4 Part B is for those functions which would be regarded as concurrently national and provincial responsibilities in the case of ineffective performance. Schedule 5 lists those which would be exclusive to the provincial government.

intergovernmental coordination in service delivery. Potential problems include wasteful duplication, unfunded mandates<sup>28</sup> and the shifting of accountability where responsibility is not assigned clearly enough and in sufficient detail.

### **c. Developmental Local Government in South Africa**

The *White Paper on Local Government* (hereafter referred to as the *White Paper*), was prepared by the Department of Justice and Constitutional Development (DoJCD) and concretises and further develops the vision of local government articulated in the Constitution. The *White Paper* places a great deal of emphasis on the notion of 'developmental local government' in South Africa and is in significant respects aligned with the emphasis of the Reconstruction and Development Programme (RDP) (RSA 1994) on ensuring the basic needs of all citizens are met and that the requisite infrastructure is provided.<sup>29</sup> The *White Paper* sees the role of local government as a far-reaching one. Developmental local government is defined in this document as "local government committed to working with citizens and groups within the community to find sustainable ways to meet their social, economic and material needs and improve the quality of their lives" (DoJCD 1998: Sec. B. 1.1).<sup>30</sup> The promotion of social development in a high economic growth environment is regarded as a key aspect of the activities of such local government, though the emphasis is not on a highly interventionist approach but more on a government sphere which establishes an *enabling* local environment: "Local government is not directly responsible for creating jobs. Rather, it is responsible for taking active steps to ensure that the overall economic and social conditions of the locality are conducive to the creation of employment opportunities" (DoJCD 1998: Sec. B.1.1). Specifically, in the *White Paper* it is argued that the primary contribution local government can make to

---

<sup>28</sup>That is a circumstance where a local government fulfils or is required to or decides to fulfil a provincial function without receiving additional funds or the additional powers to raise its own funds.

<sup>29</sup>Cashdan (2001) gives a useful critical assessment of the role of local government in South Africa.

<sup>30</sup> Section references are given as no page numbers are provided in the electronically accessed version used for this study.

social and economic development is in the provision of basic household infrastructure, that is the infrastructure required to deliver those services whose benefits are consumed in the household, rather than individually or socially, such as water, electricity, and sanitation. Thus it is argued that “(t)he starting point must be to prioritise the delivery of at least a basic level of services to those who currently enjoy little or no access to services” (DoJCD 1998: Sec.B.1.1).

#### **d. Municipal Finances and Intergovernmental Fiscal Relations**

It is difficult to generalise about the financial context of local government in South Africa both because of unevenness in the quality of reporting and because of the widely differing circumstances and capacities of municipalities. As the *White Paper* puts it: “Urban and rural municipalities, and even those in different metropolitan areas, are in very different financial circumstances, with very different prospects for providing adequate services at reasonable costs. Some municipalities, particularly those in rural areas, do not have adequate tax bases to fund the delivery of even a minimum level of basis services” (DoJCD 1998: Sec G.1.1). The large metropolitan municipalities (metros), for example, are largely financially self-sufficient and are able to provide a range of services from revenue raised through the property tax and various service charges. On the other hand, many municipalities, and certainly most rural ones, have very little potential at present to generate their own revenue and are entirely reliant on equitable share allocations and conditional grants they receive from national and provincial government.<sup>31</sup>

Even given these concerns, however, local government in South Africa, having some revenue sources at its disposal, is far more financially autonomous than provincial government, which is virtually entirely reliant on conditional grants and the provincial equitable share (PES). The figure below shows, by province, the

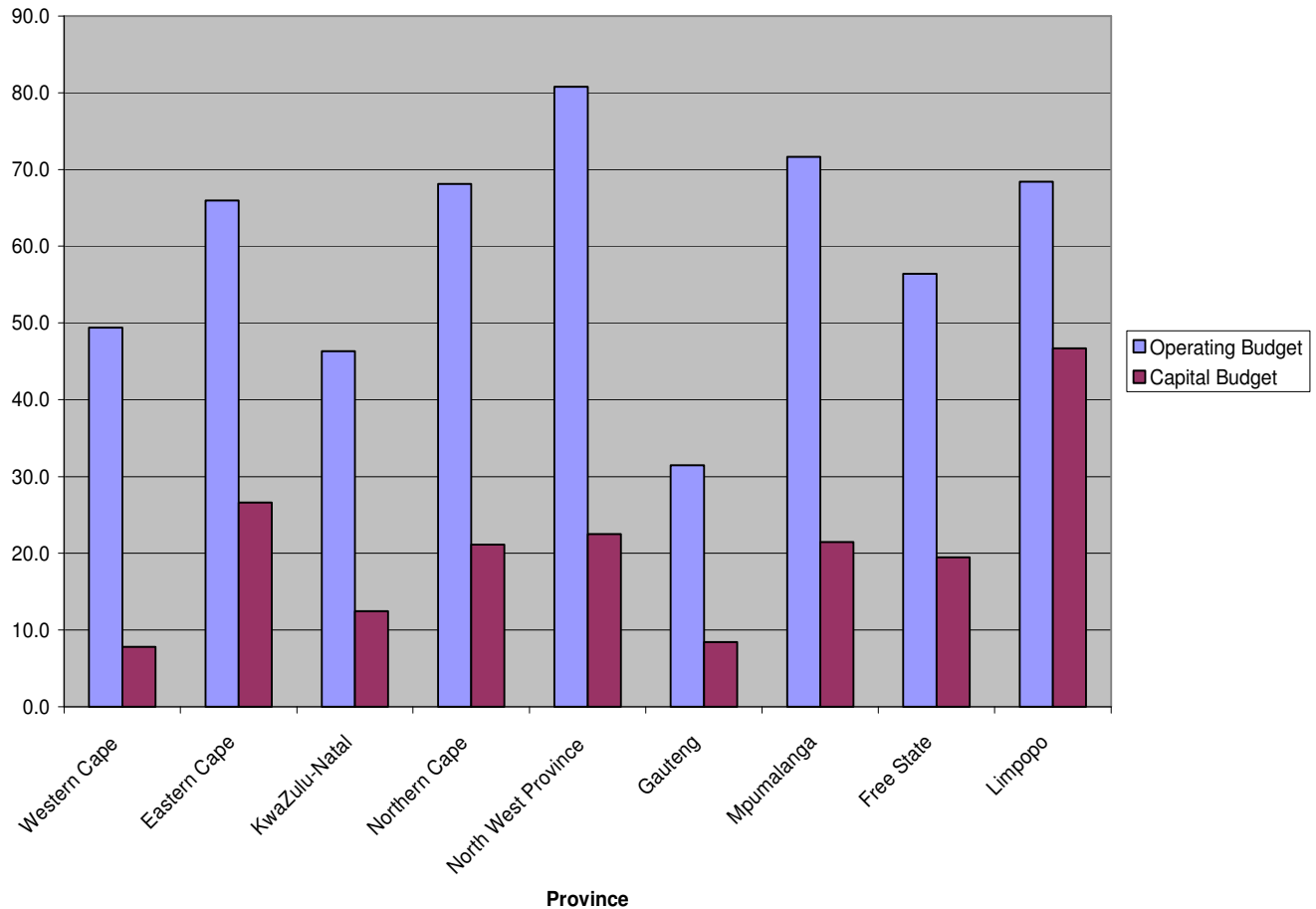
---

<sup>31</sup>Ajam (2001) in Levy & Tapscott (eds), (2001) is a good review of intergovernmental fiscal relations in South Africa.



percentage of total municipal budgeted revenue<sup>32</sup> which was received as subsidies and grants for 2005/2006 for both the operating and capital budgets.

**Figure 3.1: Grants and Subsidies as Percentage of Capital and Operating Budgets, 2005/2006**



Source: National Treasury (2006: pp 86-104), own calculations

The funding arrangements aimed at ensuring an adequate degree of support from national government to local government consists of a combination of

<sup>32</sup> Since this is budgeted revenue rather than revenue outcome, these figures assume, perhaps optimistically, that municipalities are able to realistically estimate future revenue streams from own sources. The sources of own revenue listed are 'Regional Levies', 'Property Rates', 'Electricity', 'Water', 'Sanitation', 'Refuse Removal' and 'Other Income'.

conditional transfers (e.g. the municipal infrastructure grant for basic service infrastructure) and equitable share funding. Basic service provision to those who are unable to pay for it is to be funded from the equitable share: “The Consolidated Municipal Infrastructure Programme provides grants for bulk and connector infrastructure, to enable municipalities to cover the capital costs of household infrastructure up to a basic level for low-income households. The *equitable share of nationally raised revenue* to which local government is entitled will enable municipalities to subsidise the operating costs of providing basic services to poor households” (DoJCD 1998 Sec. G. 2.2). Alternative or additional means of extending services to those who may not be able to afford user charges set at market-related cost-recovery rates include cross-subsidisation between poorer and richer households. In this case different rates are charged to different users: generally the rate ‘pivots’ around the cost-recovery rate in such a way that the surplus generated through units consumed at a higher than cost-recovery rate can be used to fund basic service levels at a lower than cost-recovery rate to targeted households. Property tax revenue may also be allocated to subsidised basic service provision and expansion.<sup>33</sup>

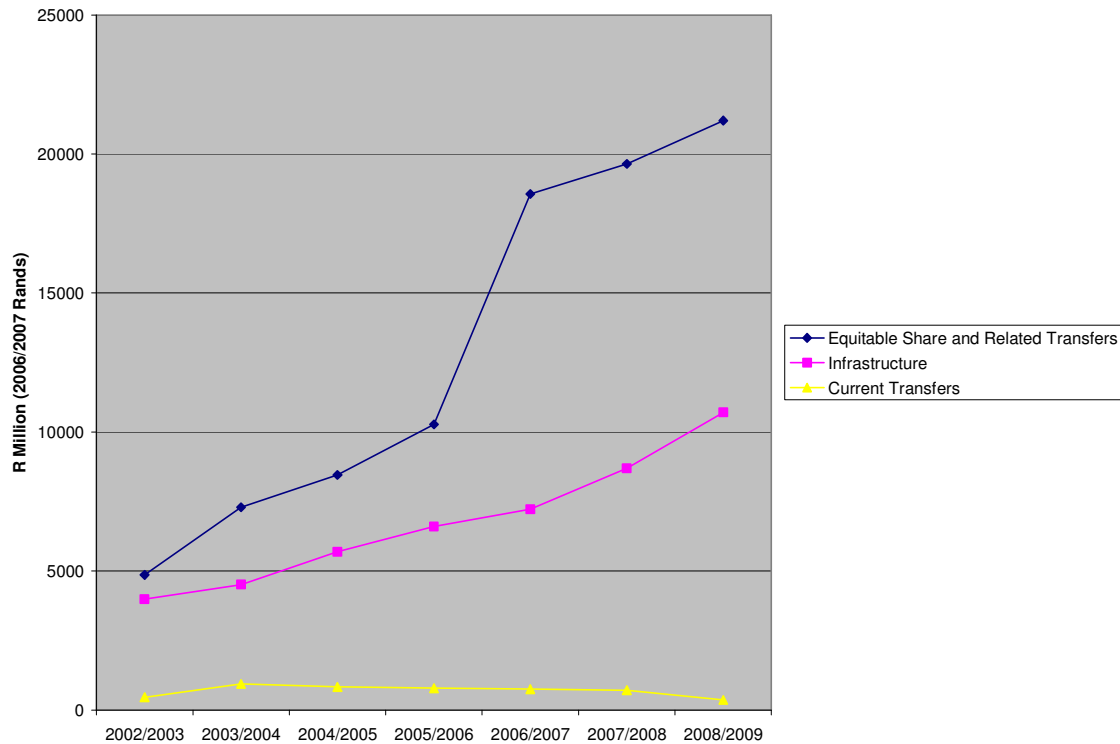
The local equitable share is an unconditional transfer to which all municipalities are constitutionally entitled. The amount going to each municipality is determined by a formula on which the FFC offers recommendations to the Minister of Finance. Trends in the total LES and conditional grant amounts in recent years are summarised in Annexure A of the *Local Budgets and Expenditure Review 2001/2002 – 2007/2008*<sup>34</sup> (National Treasury 2006: 105). Direct transfers from national government to local government are categorised as either equitable share and related transfers, infrastructure transfers or current transfers, that is to say transfers intended to support municipal operational expenditure.

---

<sup>33</sup> These equity-orientated options are discussed in detail in Chapter six.

<sup>34</sup> Henceforth referred to as the *Local Review*.

**Figure 3.2: Real Trends in Local Transfer Types (2006/2007 Rands)<sup>35</sup>**



Source: National Treasury (2006: 105), own calculations.

The large LES increase from 2005/2006 to 2006/2007<sup>36</sup> to compensate for RSC abolition is clear here, as is the steady nominal increase in infrastructure grants over the period, the bulk of which is transferred through the Municipal Infrastructure Grant (MIG). For example, in 2004/2005 the MIG transfer

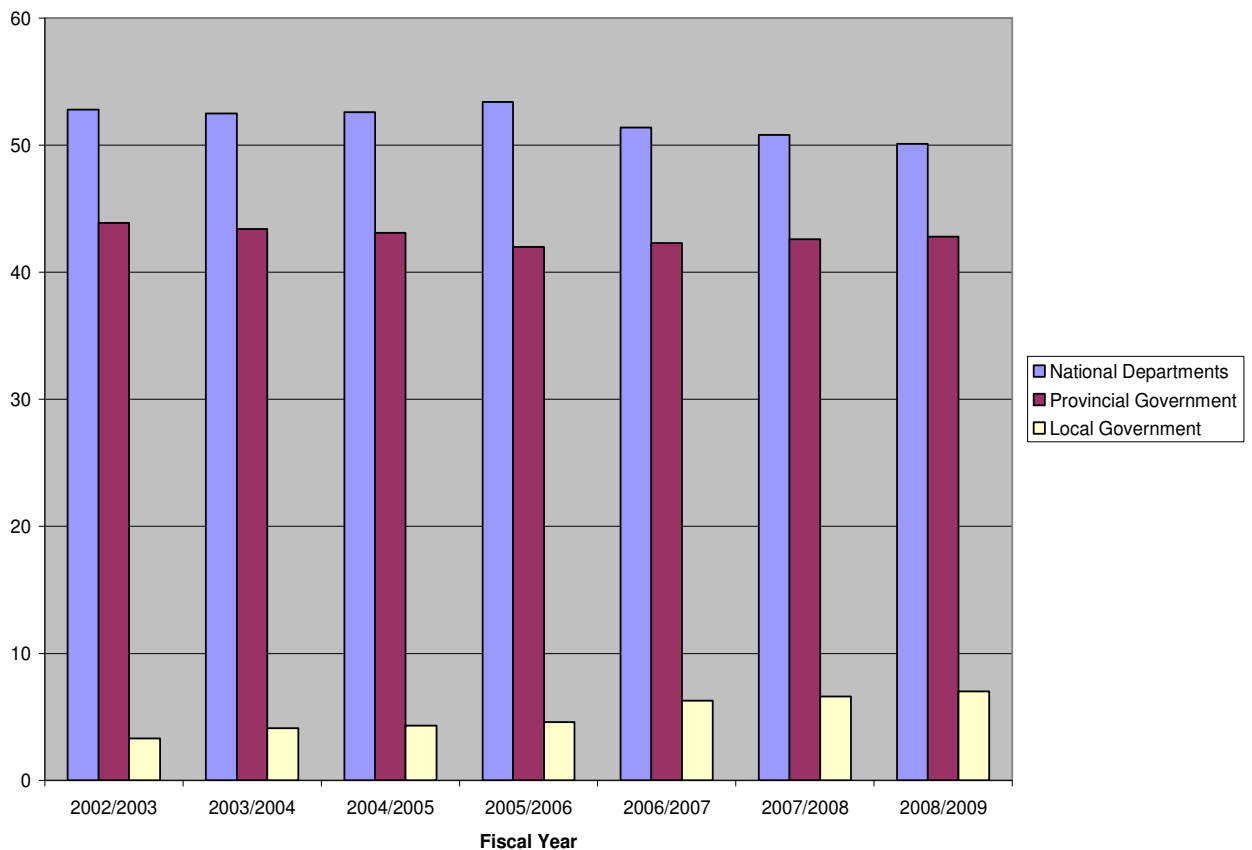
<sup>35</sup> Nominal values provided in the Review were deflated using average metropolitan all items CPI values. In this and subsequent graphs data for 2006/07 are budgeted amounts and for subsequent years are planned amounts.

<sup>36</sup> 2005/2006 is a revised estimate and the allocations from 2006/2007 to 2008/2009 are medium-term estimates. The rest are expenditure outcomes.

constituted 84% of total infrastructure transfers to the local sphere, the rest being transferred through mechanisms such as National Electrification Programme funding, funding for the implementation of Water Services Projects and Disaster Relief and Poverty Relief funds.

Figure 3.3 below indicates trends in the percentage of nationally raised revenue that went into non-interest expenditure by the three spheres of general government. The data excludes expenditure financed from provincial and local taxes and other revenue sources.

**Figure 3.3: Percentage of Nationally Raised Revenue Devoted to Non-Interest Expenditure**



Source: National Treasury (2006: 123), own calculations.

This visual representation suggests that the provincial share has remained more or less unchanged over this period, but that the local share has increased at the expense of the national share. However, it would be an oversimplification to conclude from these figures that we are in the midst of a shift where more national funds are likely to be going to the local sphere to reflect a *reprioritisation* of local functions. In fact, the shift (of local government share in revenue generated nationally from 4.6% of revenue in 2005/2006 to 6.3% of revenue in 2006/2007) primarily compensates local governments for the loss of revenue due to the discontinuation of the RSC levy. The *Budget Review 2007* emphasises that some R 24 billion of the additional medium-term allocations<sup>37</sup> is regarded as RSC compensation.

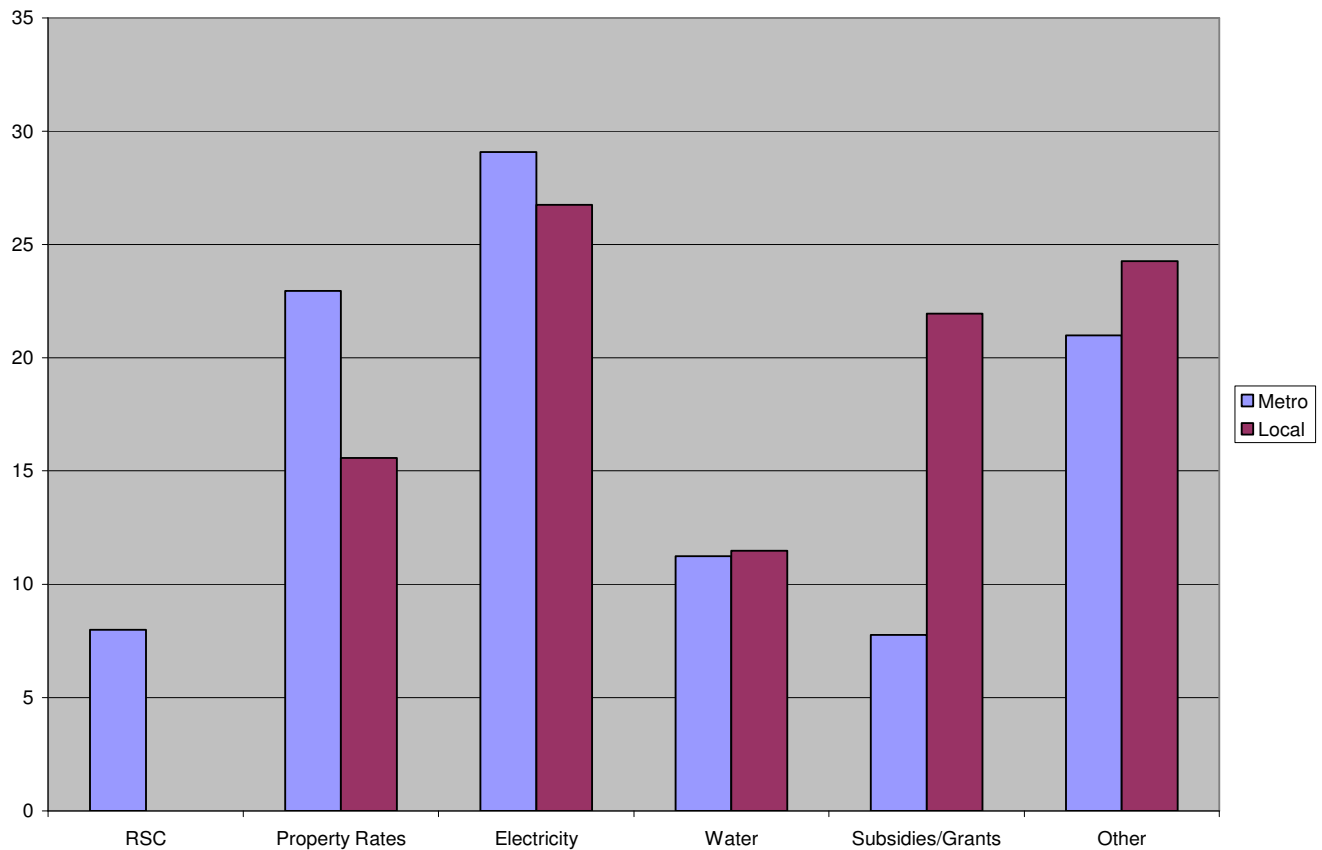
The problem with trying to gain a bird's eye perspective on local finances is that much of the financial data, whilst reflecting the extent of national government's contribution to the local sphere, says nothing about a municipality's own circumstances and, equally importantly, its own revenue *effort*.<sup>38</sup> Thus, for example, the Division of Revenue (DoR) allocation to local government provides little indication in itself of the 'resource envelope' available to specific local governments. The *Local Review* provides some assistance in this regard as it summarises the main sources of own revenue for the various categories of municipality. The main sources of revenue for municipalities are property rates, user charges (water and electricity), and 'other income'. The *Review* lists some examples of such 'other' municipal income such as "traffic fines, rental on housing stock, interest on investments, recovery of outstanding debt and the use of previous years' surplus funds" (National Treasury 2006: 13).

---

<sup>37</sup> That is additional to that proposed for 2006/2007 in Budget 2006.

<sup>38</sup> The phrase 'revenue effort' rather than 'tax effort' is preferred here as it also applies to effort in collecting fees from services provided by the municipality.

**Figure 3.4: Sources of Operating Income for Metros and Local Municipalities, 2004/2005, by Percentage**

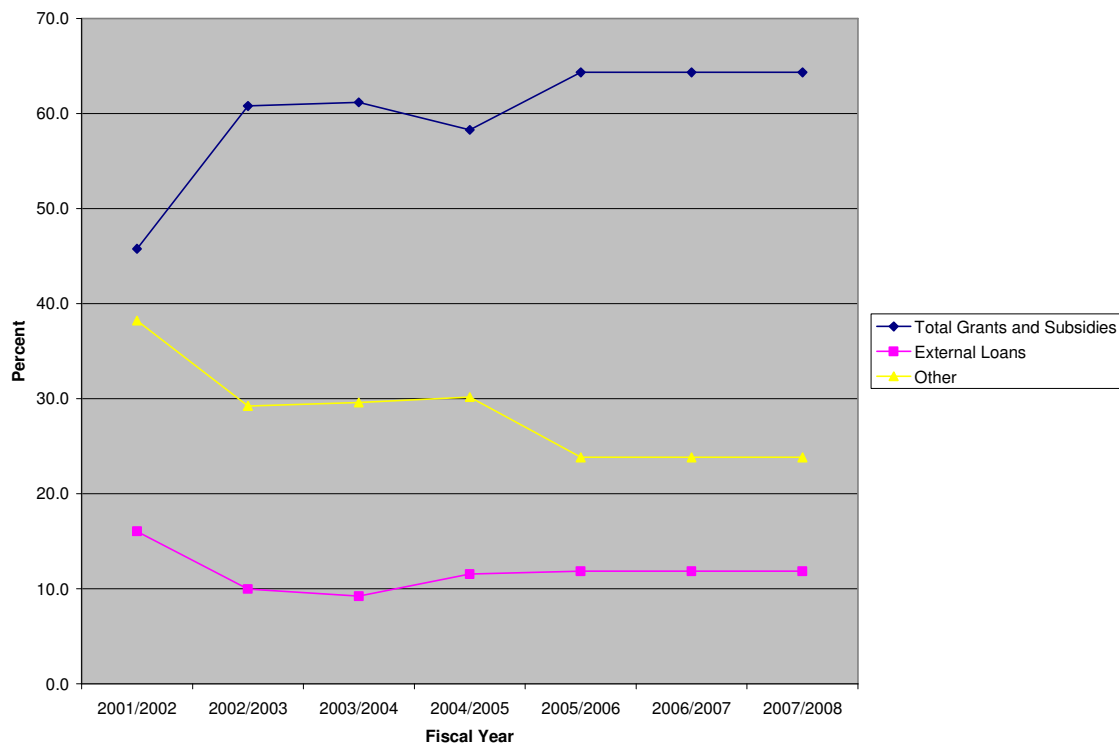


Source: National Treasury (2006: 16), own calculations.

In the case of operating income, the metros are not very reliant on grants and subsidies, which make up quite a small portion (7.8% in fact) of total operating income. For the local municipalities, the RSC was not available as a revenue source and property rates constitute a less substantial source of income. The intergovernmental grants clearly play a much more significant role in their case, and constitute 21.9% of operating income.

In the case of capital income, larger variations may occur from year to year. Figure 3.5 shows the share of total capital funding made up by grants and subsidies, external loans and 'other' from 2001/2002 to 2007/2008, as presented in 2006.

**Figure 3.5: Trends in Percentage Composition of Capital Income by Source for Metros and Local Municipalities, 2004/2005**



Source: National Treasury (2006: 16), own calculations.

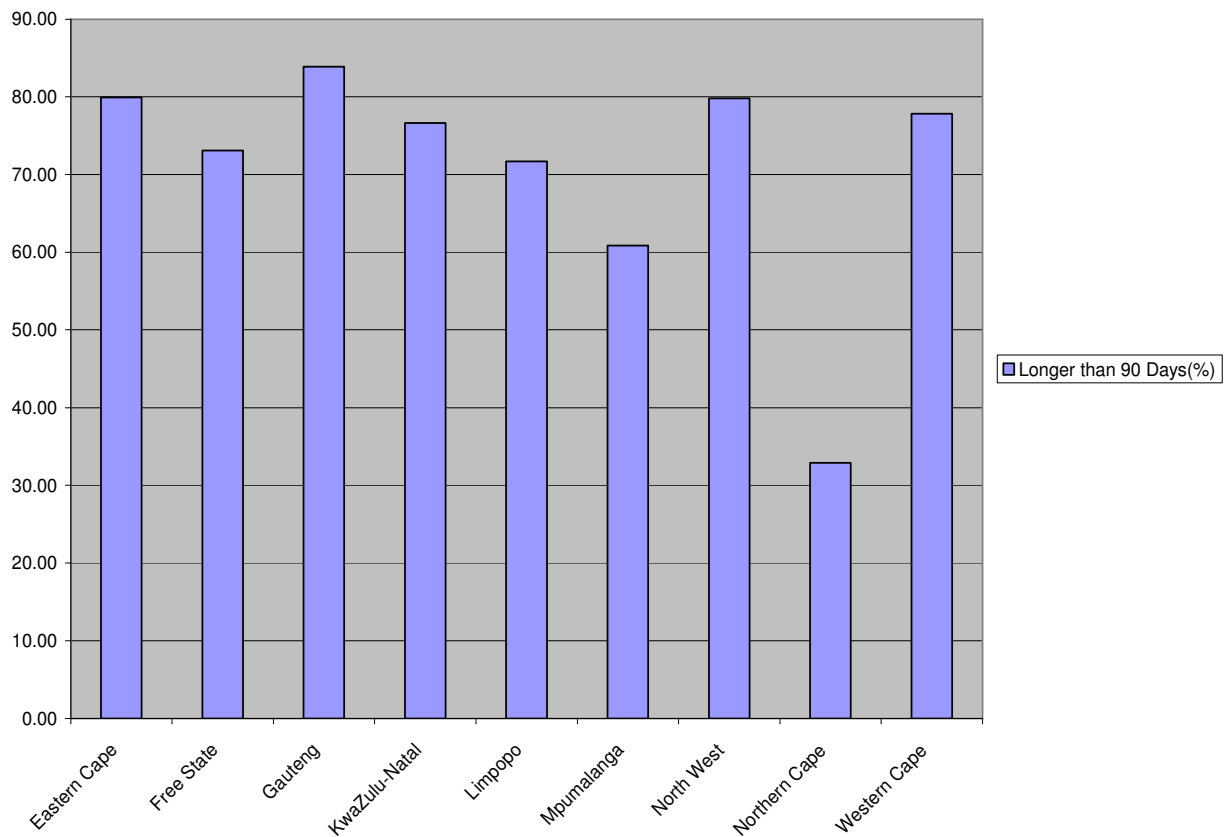
In the case of capital income the metros are in a better position to borrow externally for investment projects: external loans constitute 52.7% of capital income in their case, whilst loans made up only 11.5% of capital income for the local municipalities. The reduced extent to which local municipalities are able to borrow externally of course makes them more reliant on intergovernmental grants, such as the municipal infrastructure grant, for the financing of investment

where costs and benefits are distributed over extended periods of time. Such grants and subsidies total 58.3% of capital income for the local municipalities.

Such average values do of course hide large differences between South African municipalities, not only between metros and local and district municipalities, but also between different local municipalities, different metros and so forth. A further factor which exacerbates some of the difficulties associated with local government financial circumstances is the high percentage of outstanding payments for services provided by the municipality. At the commencement of the 2005/2006 financial year, for example, the total amount owed by municipal debtors stood at R 37.3 billion in current rands, of which approximately half (R 18.7 billion) was due to municipalities in Gauteng (National Treasury 2006: 136). Of further concern is the high percentage of outstanding municipal accounts of more than 90 days. Figure 3.6 shows, by province, the percentage of outstanding accounts which is more than 90 days old. Although debtors are reflected as assets in the balance sheets of municipalities, it is likely that fairly large portions of this service debt will never be recovered and will have to be written off. This has had, and continues to have, a significant impact not only on the current resources available to municipalities, but also on the ability of municipalities to borrow funds for capital expansion in the commercial credit market at favourable rates. The high percentage of non-performing service accounts on the municipal books means that they may be regarded as a credit risk by commercial lenders since their current revenue may not be adequate to cover loan amortisation payments.



**Figure 3.6: Non-Performing Debt as Percentage of Total Debt, 2005/2006**



Source: National Treasury (2006: 136), own calculations.

### **e. Municipal Capacity to Provide Utility Services Sustainably**

The *Non-Financial Census of Municipalities*<sup>39</sup> is an annual questionnaire completed by each municipality which provides a range of useful figures related to local services from the supply-side. That is, questions try to ascertain a

<sup>39</sup> Henceforth referred to as the *Non-Fin*.

municipality's own perception of its performance in providing services.<sup>40</sup> Information from the *Non-Fin* is not directly comparable to *General Household Survey* (GHS) and other survey and census results since the *Non-Fin* uses the concept of a 'Consumer unit' or delivery point as its consumption unit. This would not conform precisely to a 'household' as the GHS uses it, primarily because more than one household could be serviced by a delivery point.

There are some concerns about *Non-Fin* results which relate to the way questions are posed and the degree of leeway municipalities may have to put a positive spin on matters as a result of questions being formulated vaguely. For example, municipalities are required to indicate whether they have a free basic services<sup>41</sup> policy in place. Whilst the attempt to get some information on whether free basic services are a part of the policy framework is laudable, one doesn't know what 'having a policy in place' would actually mean. Does it, for example, require approval by the mayoral council? Is the mere existence of a document, for example one prepared by a consultant on behalf of the municipality, sufficient? Having a free basic services policy in place can mean many different things as can, for that matter, the 'implementation' of such a policy. Thus this kind of information does need to be compared to, amongst others, demand side information, such as that obtainable from household surveys.

Table 3.1 below duplicates a table from the *Non-Fin* and shows, by province, the extent to which municipalities possess the infrastructure required to provide water, electricity, sewerage and sanitation, and solid waste management services.

---

<sup>40</sup> The level of disaggregation is the district or metro, not individual local municipalities.

<sup>41</sup> South Africa's Free Basic Services policy is discussed further in the next two chapters, both descriptively and analytically.

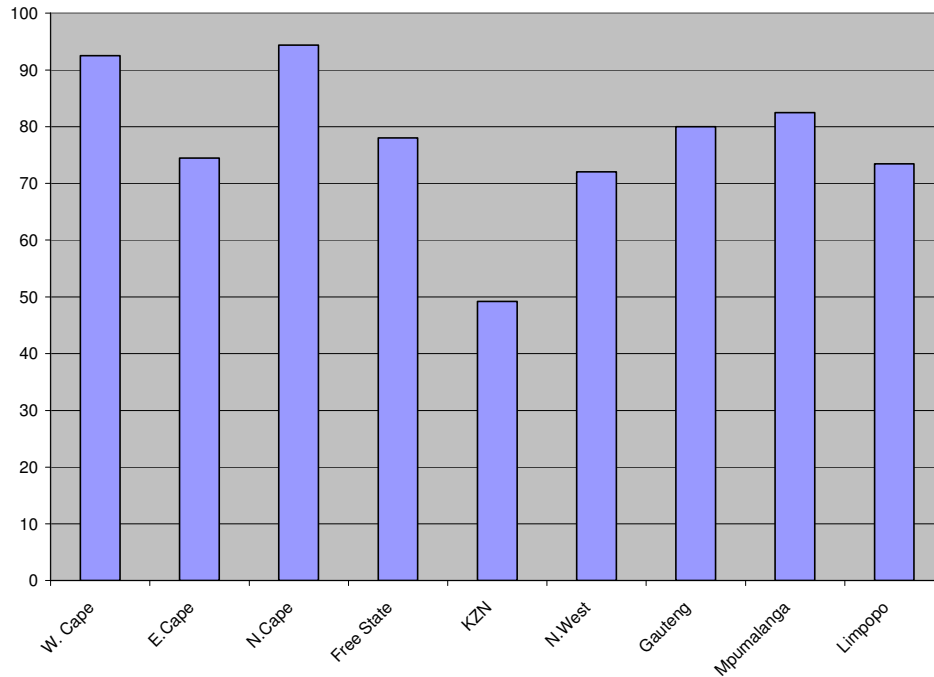
**Table 3.1: Adequacy of Infrastructure for Household/Local Services per Province, 2005**

Province	Total Number of Municipalities	Nr. Of Municipalities having adequate infrastructure for:			
		Water	Electricity	Sewerage and Sanitation	Stormwater Management
Western Cape	30	28	27	28	28
Eastern Cape	45	34	26	34	40
Northern Cape	31	30	27	30	30
Free State	25	20	18	20	20
KwaZulu-Natal	61	26	28	26	40
North West	25	20	15	18	19
Gauteng	15	12	12	12	12
Mpumalanga	20	17	15	17	17
Limpopo	32	26	19	25	24
<b>Total</b>	<b>284</b>	<b>213</b>	<b>187</b>	<b>210</b>	<b>230</b>

Source: StatsSA (2005b: 12).

The table shows that, for the country as a whole, 75% of municipalities indicate that they have the infrastructure required for water, 66% have it for electricity, 74% for sewerage and sanitation, and 81% for solid waste management. However, as with municipal finances, large differences between municipalities are concealed by these average figures. Particularly glaring infrastructural backlogs are present in the Eastern Cape for electricity, with only 58% of municipalities stating they have the required infrastructure for electricity, for example. In KwaZulu-Natal, furthermore, significant infrastructural backlogs appear to exist for all four services, with only 43% of municipalities having required water infrastructure, 46% electricity infrastructure, 43% sewerage and sanitation, and 66% solid waste management infrastructure. Figure 3.7 uses the information from table 3.1 to calculate a very simple and unweighted measure of infrastructural adequacy by province: the average number of municipalities possessing infrastructure for all four services has simply been divided by the total number of municipalities in the province in each case.

**Figure 3.7: Adequacy of Municipal Infrastructure for Local Services, 2005:  
Provincial Average Percentages**



Source: StatsSA (2005b: 12), own calculations.

In addition to variations between *provinces*, the data also allow identification of particular districts<sup>42</sup> where most municipalities do not possess the requisite infrastructure. Not surprisingly, these tend to be rural districts, and they tend to be concentrated in the provinces where backlogs are predominant. However, these municipalities are typically characterised by low populations dispersed over large areas. Thus, from the point of view of social welfare, it may not be the case that allocative efficiency gains would be achieved through channeling funds from urban infrastructure expansion projects to rural ones, since the urban one may well impact positively on a much larger number of people.

<sup>42</sup> The district is the lowest level of disaggregation for this census.

It is difficult, however, to know how to interpret the question of whether a municipality has infrastructure in place. Presumably virtually every municipality has *some* service infrastructure in place, so what does this question actually seek to measure? The questionnaire does not, regrettably, engage in any way with the adequacy of the infrastructural stock to satisfy household demand, that is with municipal backlogs. It appears that the municipality is not required to estimate the percentage of households with access as a percentage of total households. In the absence of more explicit threshold criteria the risk is that the presentation of information in this format will continue to present an overly rosy account of the infrastructural state of things.

## **f. Conclusion**

This chapter has attempted to provide a clear picture of the financial circumstances of South African municipalities and the broader system of intergovernmental fiscal relations within which local government financial matters are situated. As illustrated above, enough information is publicly available on both the expenditure and revenue aspects of municipal finances to draw some conclusions. Unfortunately, however, the manner in which information is presented, and the manner in which questionnaires have been conceptualised, does not always allow conclusions to be drawn on the adequacy of municipal infrastructure stocks. A further challenge for any study which takes a 'bird's eye view' of trends in local finances and delivery is the heterogeneity of local government circumstances. On the one hand are the metro's which engage in large-scale, complex operations funded largely through their own rates and taxes. On the other hand there are numerous local municipalities who generate so little revenue from their own tax bases and services provided that even a skeleton municipal administration would be difficult to fund were it not for transfers from national and provincial government. The next chapter delves into more detail regarding the financing of infrastructure services and develops and refines the appropriate efficiency criteria which needs to be utilised.

## Chapter 4

### Efficiency and the Pricing of Infrastructure Services

#### a. Introduction

This chapter focuses on the question of efficiency. The concept of efficiency, as used in economic analysis, is defined more explicitly. The conditions under which efficiency prevails in a market characterised by voluntary private transactions are reviewed, as are the arguments for government intervention where these conditions are not met. The particular economic aspects of infrastructure in general and household infrastructure particularly are reviewed. It is noted that two aspects of infrastructural services render them susceptible to market failure and establish grounds for government provision or regulation, namely the economies of scale associated with infrastructure service provision, as well as their positive externalities. Whether such services are provided by the public or private sector, they are of such a nature that fees can be charged which recover at least some of the costs of provision. A discussion is thus provided of the arguments for pricing based on efficiency criteria, as well as the difficulties of such pricing. A final section concretises the preceding ones by presenting a mini 'case study' of electricity provision in South Africa.

#### b. General Aspects of Efficiency and Market Failure

Efficiency in economic analysis refers to the optimal use of scarce resources by a society such that no change can make some people better off without simultaneously making others worse off. This efficient outcome, also referred to as *Pareto-efficiency*, is obtained when production and consumption efficiency are attained simultaneously in a market for a given private good or service; in other words, when more of one good cannot be provided without producing less of another good, and when the amount of a good available to one consumer cannot be increased without reducing that available to another. More formally, Pareto-efficiency requires that the marginal rate of transformation in production (MRT)

equals the marginal rate of substitution in consumption (MRS). In their seminal paper 'Existence of an Equilibrium for a Competitive Economy', Arrow and Debreu (1954) demonstrated that a competitive equilibrium must always generate a Pareto-efficient allocation of resources in production and consumption. However, real markets do not necessarily meet the conditions for being competitive, and consequently Pareto-efficiency constitutes an evaluative benchmark rather than an assumption about existing markets.

Under Pareto-efficiency overall social welfare is at its highest obtainable level given the current endowment of resources and assuming that no inter-personal redistribution takes place. In fact, an infinite set of Pareto-efficient outcomes can be generated even from the so-called 2x2x2 model<sup>43</sup> consisting of two factors of production, two products (say X and Y) and two consumers (say L and M), since each different initial income distribution will generate a different efficient outcome.

Pareto efficiency need not, however, be accompanied by equitability, the maximisation of social welfare or related concepts, since different Pareto-efficient outcomes represent different distributions of X and Y between L and M. Theoretically, it is possible through a process of social choice (a collective decision-making mechanism) to determine the particular point of Pareto efficiency that corresponds to maximum social welfare.<sup>44</sup> However, once we allow for the possibility that overall welfare can be determined and enhanced through redistributing income from some to others, we are necessarily also involved in the (methodologically problematic) ordinal comparison of inter-personal utility. It seems reasonable to assume that taking R10 from a rich person and giving it to a poor person would raise overall welfare: the marginal gain in utility of R 10 for the poor probably exceeds the marginal loss in utility to the rich person. But there is no rigorous theoretical basis for this position save

---

<sup>43</sup> Often employed because its insights can be represented visually: this becomes difficult in a three or more variable model.

<sup>44</sup> This matter is attended to further in Chapter 5.

that money, as a kind of commodity, is characterised by declining marginal benefit to the person who possesses it. Since the determinants of demand for money are likely to be more complex than in the case of other commodities this assumption may in fact not hold in numerous circumstances.<sup>45</sup> Put another way: Pareto efficiency only means that no 'painless' reallocation can take place.

For private, voluntary market transactions between individuals to generate a Pareto-efficient outcome, a number of conditions need to prevail. These relate generally to the dual requirements that a market must firstly exist and that price-quantity determination through production and consumption decisions within that market needs to incorporate all the socially relevant information. A typical list of requirements for markets to operate efficiently, that is for an equilibrium price to exist which corresponds to relative scarcities and preferences, is provided by Hyman (2005: 59):

- All productive resources are privately owned
- All transactions take place in markets, and in each separate market many competing sellers offer a standardized product to many competing buyers
- Economic power is dispersed in the sense that no buyers or sellers alone can influence prices
- All relevant information is freely available to buyers and sellers
- Resources are mobile and may be freely employed in any enterprise

Conversely, of course, market failure occurs when one or more of these conditions is not met. The existence of market failure establishes a necessary, though certainly not a sufficient, condition for the use by society of a collective

---

<sup>45</sup> One attempt to adjust the Pareto-criterion to allow for redistributory policies without delving into the realm of values is the Kaldor-Hicks criterion which regards a policy as justified if the gains are distributed in such a manner that the winners could in principle compensate the losers and still be better off. (reference). This amounts to allowing redistribution if it is clear that the gains to the winners outweigh the losses to the losers. It can also be regarded as a sanctioning of redistribution where it is a positive-sum game rather than a zero-sum game.



decision-making mechanism ('government') in order to enhance its overall welfare or the welfare of certain of its member who are regarded, according to some explicit or implicit criteria, as being especially deserving. There are, however, a number of reasons why 'government failure' might occur in the attempt to address market failure, and in so doing make things worse than in the non-intervention scenario. The primary reason is probably that government itself does not have all the relevant information at its disposal; related to this is the difficulties policy-makers face in interpreting both the direction of trends and the impact of intervention, whose effects may be delayed or transmitted along complex chains through the economy. Additionally, even assuming for argument's sake that policy-makers are omniscient, this does not mean they necessarily have adequate control over the departments and other agencies, as well as the maximising behaviour of the bureaucrats tasked with implementing policy. A good policy badly implemented and administered may often do more damage than a laissez faire choice would have done. Lastly, it is also erroneous to think of 'government' as a homogenous entity pursuing one set of objectives. Diverging objectives and incentives may mean that a policy which appeals to politicians may be poorly implemented by bureaucrats who feel they would be disadvantaged by its effects.

It appears as though government failure has at times been particularly pervasive in the infrastructural sphere. The World Bank's 1994 *World Development Report* firstly emphasises the pivotal role government has played in relation to infrastructure: "The record of success and failure in infrastructure is largely a story of government's performance" (World Bank 1994: 4). Much time is, however, devoted to an analysis of government failures in relation to infrastructure investment and maintenance. Noting that the growth of the global infrastructure stock, as well as related welfare-improvements for millions, has in some respects been spectacular, the *Report* nonetheless also argues that in a number of significant respects performance has been disappointing. Specifically, the following infrastructure-related failures are listed (World Bank 1994: 4-7):

- Allocative inefficiency with insufficient prioritisation of essential services and a lack of responsiveness to the preferences of users
- Overemphasis on devoting resources to further infrastructure investment rather than good maintenance of existing infrastructure
- Technical inefficiency and waste as a result of which many developing countries experience a wide divergence between infrastructural capacity and actual outputs
- Inadequate integration of environmental sustainability awareness into infrastructural decision-making

The *Report* accordingly recommends various ownership and service administration options along the continuum from maximum government involvement to wholly privately owned and provided services.

In order better to understand the role of government with regard to infrastructure, a brief review some of the pertinent economics aspects of infrastructure is provided in the next sections.

### **c. Some Economic Aspects of Infrastructure**

Infrastructure is a form of fixed capital characterised by bulkiness, high sunk costs, and significant economies of scale stemming from its network attributes. It is typically associated with the physical networks required to distribute resources such as water and electricity to households and businesses, with the buildings and related physical assets necessary to provide social services such as health and education, and with transport and communications networks such as road, rail, port and fixed-line telephone systems. The expansion of some aspects of infrastructure necessarily must occur 'en masse' rather than marginally, in large units rather than piecemeal. Fourie (2006: 12) provides an example: "Although the electricity grid can be expanded on the margin, electricity generation capacity necessarily will expand in large quantities, as new power plants are built". He goes on to mention one significant consequence of this bulkiness: "infrastructure

is usually either underprovided (not enough capacity to satisfy demand) or overprovided (production not at full capacity).”.

As capital, infrastructure is an economically useful income-generating factor of production. It is not consumed in use but remains available to production for an extended period of time, during which it typically depreciates in value. From this definition it follows that the costs associated with providing infrastructure are two-fold: the initial investment or ‘gross fixed capital formation’, and the subsequent maintenance and operation costs of infrastructure, which are a form of recurrent expenditure typically financed from the general revenue stream available to a firm or public authority. The fact that a fairly clear separation of the investment and maintenance aspects of infrastructure is possible also enables funding for each aspect to be considered separately and possibly allocated separately if efficiency or other criteria seem to warrant it. The relation between investment expenditure and maintenance and operations expenditure also implies that, from a local budgeting perspective, an initial investment, captured on the capital budget, will subsequently need to be maintained through increased allocations on the operating budget.

A further significant economic aspect of the bulkiness of much infrastructure is its immobility. In the words of an OECD conference report on investment for development: ‘Once built, a road or hydroelectric dam cannot sensibly be dismantled and moved elsewhere’ (OECD 2005: 3). To say that infrastructure is largely immobile amounts to saying that a large proportion of its start-up costs are in fact *sunk* costs. A dam has no opportunity costs, no ‘next best use’, once the infrastructural outlay has taken place. High sunk costs also mean high risks for the provider, who may not be able to pull out of the investment if circumstances change or if initial assumptions and projections turn out to have been erroneous. The long time-frames for recouping on an infrastructural investment mean that uncertainty and therefore potential risk are amplified: information on the global and country economic environment a decade hence, for

example, will necessarily be based on the scantiest of evidence and mostly on conjecture.<sup>46</sup> Many private firms would therefore not be interested in infrastructural investment in the absence of some forms of assurance or insurance against unforeseen developments. Private firms may shy away from the economic and country risks of large-scale infrastructural investment under circumstances where they do not have access to financial markets which allow them to insure effectively against the associated risks, or where these risks are regarded as so high that insurance is hard to come by even in developed financial markets. The OECD report cited earlier provides an example of this kind of risk: “Investors in infrastructure are often vulnerable...to changes in government regulations...Before they invest, the government may promise them prices high enough to cover the costs of investment, but afterward the government will be tempted to please customers and voters by keeping prices low” (OECD 2005: 2).

Economies of scale, externalities and equity are probably the strongest grounds for advocating for government ownership or regulation of infrastructural investment. The first three are discussed further in the rest of this chapter. Chapter 6 grapples exclusively and in detail with questions relating to equity.<sup>47</sup>

Before discussing economies of scale and externalities further, the next section deals in more detail with some of the specifics of household infrastructure, which narrows the discussion to infrastructure pertaining to local government.

---

<sup>46</sup> This line of thinking famously of course led Keynes to advocate for a public investment body charged with ensuring socially optimal levels of investment; it also informs the various development financing banks, such as the World Bank in the global context and the DBSA in South Africa.

<sup>47</sup> A fourth one would perhaps be that of risk associated with large-scale infrastructural investment, which would possibly deter private investors in the absence of adequate assurances and guarantees by government.

#### **d. Components of Household Infrastructure Services**

At times terms such as ‘basic infrastructure’, ‘local services’, ‘household services’ and ‘access to basic services’ appear to be used more or less interchangeably in the South African context. The *White Paper*, for example, contains as a glossary entry the term ‘basic household infrastructure’, which it defines as “(e)ssential municipal services such as water, sanitation, electricity, roads, stormwater drainage and street lighting needed to sustain a healthy and safe standard of living” (DoJCD 1998). Infrastructure is here simply equated with services.

Distinctions need to be drawn, however, between the various cost components of providing a service, since persuasive economic arguments may exist for financing different components from different sources. Table 4.1 provides a very simple ‘taxonomy’ of potential costs, organised in terms of the budget they would be attributed to in the municipal case.

**Table 4.1: Costs of Infrastructure Services**

<b>Capital Costs (Capital Budget)</b>	<b>Recurrent Costs (Operating Budget)</b>
Expanding infrastructure	Bulk purchasing of the resource
Replacing Infrastructure	Administering delivery (‘Operations’)
	Maintaining infrastructure (‘Maintenance’)

Explicitising costs like this also enables costs associated with eradication of backlogs to be distinguished from costs associated with continuing to provide a service. In fact, for purposes of estimating the future costs of various policy options household circumstances should be distinguished as shown in table 4.2.

**Table 4.2: Costs Associated with Backlog Eradication**

Household Type	Capital Costs	Recurrent Costs
Already Connected (A)	No	Yes
Connection to Occur (B)	Yes	Yes
Not connected and no connection plans (C)	No	No

It is not that difficult, for planning purposes, to conduct a simple costing of scenarios for each group over a given period for different rates of backlog eradication and service standard. The total cost which needs to be budgeted for consists of course of recurrent costs plus capital costs. The key relationship between them is further that the more expenditure is incurred on the capital budget, the more needs to be budgeted for maintenance and operations (M&O) subsequently in the recurring budget. Similarly, if more households are provided with access then the demand for the service will of course increase, which also means an increase in the bulk resource the municipality needs to purchase. A municipality will therefore have to ensure that it possesses reliable estimates, across a range of service conditions and over a medium-term planning framework, of the value of *a*, *b* and *c* in the following, where they essentially represent measures of the existing and additional costs of maintaining infrastructure and expanding infrastructure.

Existing M&O Cost =  $a * \text{Existing Infrastructure Stock Value}$

New M&O Cost =  $b * \text{New Infrastructure Stock Value}$

New Bulk Resource =  $c * \text{New Infrastructure Stock Value}$

Though the efficient and effective delivery of household services is not a function only of infrastructure adequacy, the emphasis in South Africa quite naturally falls on infrastructural considerations since backlogs still exist. However, as the above breakdown of functions and costs implies, the efficient and effective *administration* of basic services is equally vital, as is adequate allocations to infrastructure maintenance in the budget. For example, the World Bank's *World*

*Development Report 1994* emphasises efficiency and effectiveness in infrastructural use, that is to say sound management of existing infrastructure assets ('productive efficiency') and allocative efficiency in the services generated from the infrastructure, as much as it emphasises the need for new infrastructure investment. Hulten concurs with this view and cites a suggestive example from Easterly and Levine: "while Chad may have 15 000 telephones, 91 percent of all telephone calls are unsuccessful" (Easterly and Levine (1996) in Hulten 1996: 2). Similarly, the World Bank (1994) reports that it cost Africa almost four times as much to replace/rehabilitate infrastructure as it would have cost to maintain it adequately. A further implication of this perspective is that pricing and budgeting arrangements need to be successful for all the components of the service.

### **e. Economies of Scale**

Infrastructural networks enable economies of scale in the extension of a service to households (or businesses), assuming that these households are located within the infrastructural service area. The large economies of scale associated with infrastructure argue against infrastructural duplication<sup>48</sup> between alternative providers: this 'natural monopoly' aspect often constitutes an argument for government provision or regulation of new infrastructure formation (that is infrastructure investment, a flow variable) and maintenance of an existing stock of infrastructure. In the apt formulation of the earlier OECD report: "The problem of infrastructure provision has its roots in the potential for market power that results from economies of scale. It rarely makes sense to have two competing roads between two points – or competing electricity grids" (OECD 2005: 2).

Where economies of scale are present, the average and marginal unit costs of service extension decline, and would decline over any range of output needed to satisfy demand in the service area. Since start-up costs tend to be substantial and the marginal costs associated with connecting new users negligible as a

---

<sup>48</sup> But technological developments do yield close substitutes which may have smaller sunk costs, as for example the challenge to fixed line telecommunications services by cellular telephones.

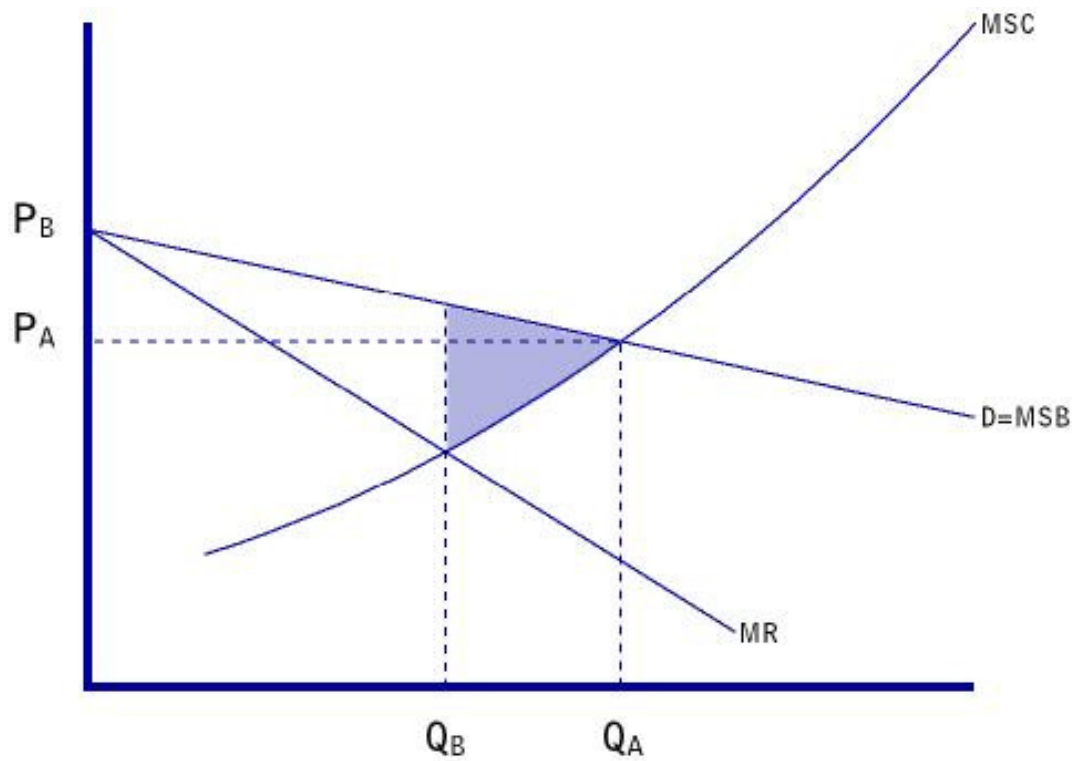
percentage of this, a clear first-mover advantage exists in the case of infrastructural services. Once a firm has incurred such initial costs, which are furthermore typically *sunk* costs given the physicality and immobility of infrastructure, the firm is able to provide the service to additional users at a lower cost than a potential entrant to the market who still needs to incur such initial costs. The firm has also credibly signalled its intention to *remain* in the market, which will also dissuade potential other entrants from actively contesting the market. Within the framework of the structure-conduct-performance or SCP<sup>49</sup> approach to market structure and industry efficiency, the large start-up costs and likely economies of scale associated with infrastructure thus constitute an entry barrier to other firms, who are unlikely to contest the market, and therefore reduces the likelihood that the 'market' will in any way be a competitive one. Without competition, or the credible threat of competition, market power arises, which is the ability of a monopolistic, price-making firm (or oligopolistic firms) to conduct business so as to earn abnormal profits by producing less than the socially efficient amount. Figure 4.1 represents the conventional depiction of this circumstance. Efficiency entails production where marginal social cost (MSC) and marginal social benefit are equalised. This occurs at quantity QA and price PA. However, market power means that the monopolist is able to produce a quantity where marginal revenue MR equals MSC, but at a price PB on the demand curve. In essence the monopolist is able to produce less and ask a higher price for the product. The shaded area represents the net loss to society as a result of monopolistic pricing in this example.

---

<sup>49</sup> The SCP approach argues that the structure of a market plays a large role in determining the conduct of firms which in turn largely determines the performance of the market in the sense of its degree of efficiency.



**Figure 4.1: Price and Quantity Curves for the Monopolist**



Constraints to unlimited extension at declining unit rates would consist mainly of congestion-related increases in costs: such increases would be associated with either increased network inefficiencies as demand peaks, or with increased costs of bulk resources as demand for them increases. The latter case would not have any impact on the validity or scope of the economy of scale argument. In the former case, growing network inefficiencies would probably have to be addressed by further infrastructural investment, a circumstance which would be identical to new infrastructure investment.

The conventional structure-conduct-performance model links the *structure* of a market, that is the number and size of firms producing a closely substitutable product, to the conduct of firms when it comes to setting prices, innovating, keeping costs down and investing. This, in turn, has implications for the

performance of the market and the economy as a whole in relation to the attainment of allocative and productive efficiency. The presumption is that in perfectly competitive markets, optimal performance is the norm, whilst markets that are not competitive, such as monopolistic, oligopolistic, and monopolistic competition markets, may be inferior: “Although firms may be productively efficient, the level of output is unlikely to meet the requirements of allocative efficiency....firms ...posses a degree of *market (or monopoly) power*: that is, they have some discretion in determining the price at which they sell their output...they are able to raise price above the level of marginal cost” (Ferguson & Ferguson 1998: 15).

Ferguson and Ferguson (1998:18) suggest that recent refinements in the theory of market power have stressed that what may determine the conduct of firms is not so much the existing market structure, but the *contestability* of the market, that is the extent to which a firm earning long-term abnormal profits could find itself under challenge by new entrants and needs to consider this when setting prices and the like. The less significant the barriers to entry of a market, and therefore the more contestable the market, the more conduct will approximate the perfect competition model, even if few firms operate in it at a particular time. Where entry barriers are large, a ‘first mover advantage’ resides with the incumbent firm, especially where entry into the market entails incurring significant sunk costs, that is costs which cannot be recouped on exiting the market.

Even where market structure suggests that non-competitive pricing may prevail, this need not be the case and a dynamic assessment of why the particular market is structured in the way that it is also needs to be made. In some cases it may be due to the exceptionally efficient performance of the dominant firm, which has simply eliminated its rivals by outperforming them over time. Regulation or more drastic intervention is unlikely to create greater efficiency here. In the case of natural monopoly industries, furthermore, efficiency is maximised precisely when only one firm is operant, since these industries are characterised by

negligible marginal costs and therefore decreasing average costs over the entire range of output required to satisfy demand. Duplication of, for example, generation facilities or a transmission network would be extremely wasteful since the likelihood is that both would be utilised at less than full capacity.

Whilst the efficient market structure in the natural monopoly case may consist of a single firm, this does not assure that the incumbent firm, whether privately or publicly owned and managed, will behave as policy-makers concerned with efficiency would like it to: there is no a priori reason to assume that prices charged by a non-regulated state-owned utility, say an electricity generator and distributor, will be efficient and not in fact similar to monopoly price-setting practices in other markets. Intervention can take the form of introducing regulation where it currently does not exist. Such regulation requires regulator capacity to establish a set of plausible benchmark efficient prices against which actual prevailing prices can be tested. It can also take the form of changes in ownership, that is privatisation or nationalisation.

It is increasingly recognised, however, that the mere transfer of ownership from the private sector to the public sector does little to enhance efficiency (Ferguson & Ferguson 1998: 194-228). What matters is the extent to which competition or the possibility of competition (contestability) exists in conjunction with effective regulation. Both a private and a public company may exploit market power and operate inefficiently or set price so as to earn abnormal profits if given the opportunity to do so. In the case of natural monopolies, no competition can be introduced to enhance efficiency, since a one-firm market is the efficient outcome. Effective regulation of price is the best means to ensure efficiency in this case. The fundamental aim of regulation, whatever technical approach is employed, must be to determine what the equilibrium price would be in the market, *if it were competitive*. One method of achieving this is to clearly articulate the means by which price should be determined, and comparing the results of such an exercise with prices obtained in practice. In essence the means of

assessing price will derive from estimates of the *cost* of service provision to the provider.

These issues are clearly relevant to the provision of infrastructural services, and especially the infrastructure itself, which is a bulky network characterised by economies of scale. The potential for inefficient price-setting characterises two sets of relations, firstly in a utility's provision of bulk resources to a municipality, and secondly where the municipality, as a retailer, supplies services to the households and businesses which are its customers. Both utility and municipality can be regarded as potential price-makers rather than price-takers, by virtue of their privileged access to requisite infrastructure and, in some cases, their legislated monopoly power.

The natural monopoly problem is not limited to circumstances where an entire country's demand can and should be satisfied by one producer. It potentially occurs whenever economy of scale considerations render particular markets uncontested, difficult to contest, or legally incontestable. Thus, *regional* monopolies may engage in inefficient pricing as much as national ones, as may municipalities if their access to municipal distribution infrastructure is exclusive or difficult to contest. In the case of the 'retail' transactions between a municipality and households, a municipality may operate in an uncontested market and function as the only supplier of water and electricity to households. This market is likely to be more contestable than that of generation and distribution, however, since close substitutes may be available which bypass the need for access to municipal infrastructure: the substitution of gas or paraffin for electricity and the use of commercial water vendors as alternative providers of water services are two examples. The example of cell phones has in recent years also shown how innovation can erode natural monopoly power deriving from privileged infrastructure access. Furthermore, local elections are likely to compel a municipality to refrain from obvious monopolistic pricing, since 'service delivery'

and perceptions around its affordability will remain a key issue in South African local elections for some time to come.

### f. Local Infrastructure Externalities and Financing Options

Externalities imply that the benefits and/or costs of transactions between the municipality and the household are not limited to the transacting entities, but that the effects of the transaction, whether positive or negative, are also social. In other words, the private supply and demand curves faced by the municipal provider and a household are not identical to the supply and demand curves they would face if the *social* impact of their transaction were factored into the analysis. The transaction *price* for such a privately derived equilibrium is different from what a socially derived equilibrium price would be, and consequently more or less is produced or consumed than socially desirable.

**Figure 4.2: A Basic Representation of Externality**

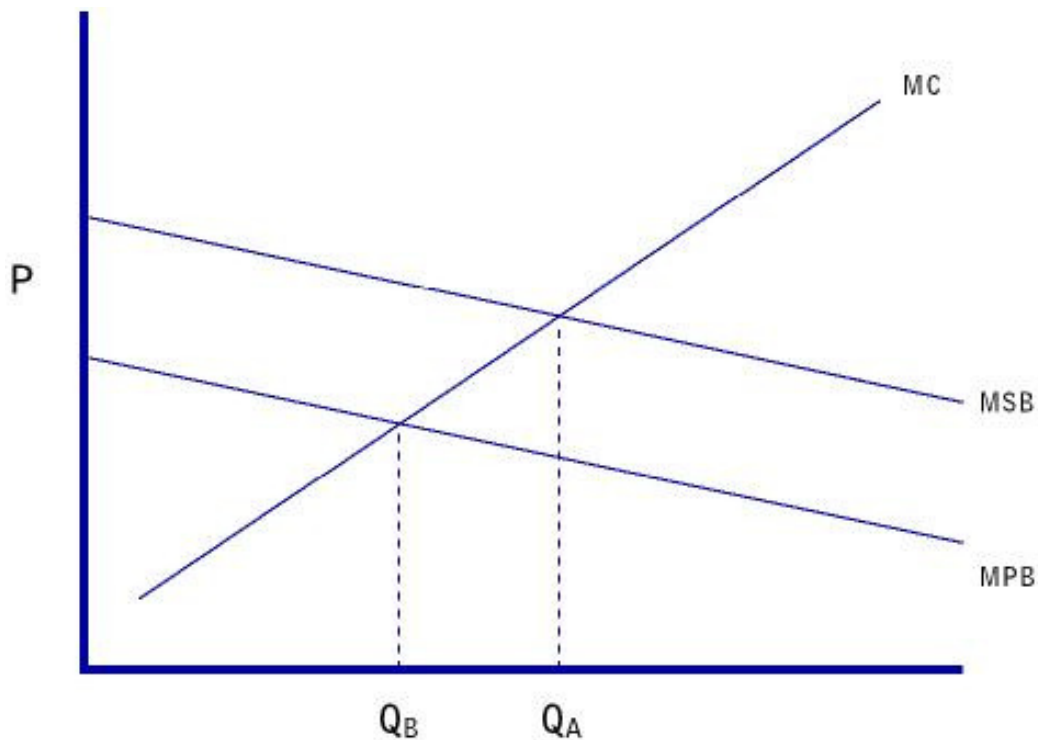


Figure 4.2 represents the concept visually. From a social point of view one would want the producer to produce a quantity  $Q_A$  where the marginal cost curve intersects the *social* benefit curve (MSB). However, the producer only factors in private benefits and produces where the marginal private benefit (MPB) curve intersects the marginal cost curve, that is at a quantity  $Q_B$ . A less than optimal amount is therefore supplied.

Infrastructure provision typically exhibits strong positive externalities, especially in developing countries, which are not always adequately incorporated into economic decision-making. Positive externalities associated with an adequate and well-maintained infrastructure stock include the benefits of reduced production and transaction costs, the faster dissemination of knowledge, skills and 'best practice' in an economy whose parts are closely connected, and more broadly still the increased productivity and therefore competitiveness of the economy, as well as the positive externality effects associated with a higher level of median and mean well-being in the population through improved access to potable water, sanitation facilities and so forth.<sup>50</sup>

A number of policy options are theoretically available for addressing externalities. In practice, however, it is very difficult to distinguish an actual transaction price from a 'socially optimal' one: that is, it is difficult to quantify the extent of a presumed externality. This also means that it is very difficult to enhance efficiency and social outcomes by trying to *compensate* for externalities. The externality argument can quite easily be abused by interest groups, such as advocates for trade protection, who cite all sorts of social benefits if their industries are more protected from international competition. Rosen articulates the cautionary, somewhat sceptical view well: "Many people who have never heard the term *positive externality* nevertheless have a good intuitive grasp of the term and its policy implications. They understand that if they can convince the

---

<sup>50</sup> Cf. Canning & Bennathan (2000).

government their activities create beneficial spillovers, they may be able to dip into the treasury for a subsidy” (Rosen 2005).

Bearing these concerns in mind, the general policy options available to bring the socially inefficient outcomes associated with externalities closer to marginal efficiency and greater social welfare include the following:<sup>51</sup>

- Taxes (e.g. to encourage less production where negative production or consumption externalities occur)
- Subsidies (e.g. to encourage more production where positive production or consumption externalities occur)
- Creation of a market (e.g. for emission fees in the case of pollution)
- Establishing of property rights (e.g. induce self-interested transactions where formerly private opportunity costs of resources were zero, such as common fishing grounds)
- Regulation (e.g. limiting production with negative externalities by threat of a fine)

Pigou is generally credited with developing the argument that subsidies and taxes can be used to align social and private supply and demand curves. Where the social benefit is greater than the private benefit, a subsidy to encourage more production and/or consumption may be considered; where private benefit is greater than social benefit a tax may be considered.

In the case of intergovernmental fiscal relations, some grounds exist for a higher level of government’s attempts to use transfers to subsidise<sup>52</sup> those local services which have positive externality effects benefiting adjacent jurisdictions and which would be underprovided in the absence of such transfers. In principle, the government of a benefiting jurisdiction could also be expected to bear this

---

<sup>51</sup> Derived from Rosen (2005), Ch.6 and Hyman (2005), Ch.3

<sup>52</sup> The case of negative externalities is not discussed further: although these doubtless occur as a result of government actions, they are best regarded as inadvertent rather than as a policy objective.

cost. In a sense, intergovernmental externalities shift the debate from consideration only of *pricing* (that is, for semi-public goods where exclusion is possible, determining an appropriate user fee) to consideration more broadly of *financing*. If the price at which a cost-recovering municipality provides a service to households does not capture social benefits, then an argument is established for a higher level of government, pursuing broader objectives, to subsidise some of the service.<sup>53</sup> In much the same way that governments are called on to intervene where market transactions between private actors generate externalities, a higher level of government can, in theory at any rate, be called on here to finance benefits<sup>54</sup> arising from self-financed municipal service delivery which go to residents of other municipalities. Put another way: it is assumed that a municipality, financing its own services under a balanced-budget constraint, will provide a service up to the point where marginal municipal cost and municipal benefit are equalised (cf.  $Q_B$  in Figure 4.2): this implies under-provision of services characterised by regional or national externalities.

It should be clear that the *primary* source of funding (i.e. the ultimate incidence of the cost) for publicly provided services characterised by the possibility of exclusion, from an efficiency perspective, should still be the households who benefit and who can adjust their behaviour to maximise utility. Externality-motivated subsidisation merely entails that *some* of the burden of service provision is financed by those who do not directly consume the service in question, which induces a larger supply amount. In such cases difficult decisions have to be made concerning the relative proportion of the service to be financed through taxation and through user fees, where the tax component (more specifically national tax revenue transferred to the municipality) represents an attempt to capture externality effects. By definition a tax is a compulsory transfer

---

<sup>53</sup> The argument that a higher level can also punish a lower level for negative externalities is probably correct from an economic perspective but is likely to be much more politically controversial.

<sup>54</sup> In fact, in this study the emphasis is on the case of municipal transactions generating positive externalities.



from an individual to government where no inherent link (that is direct reciprocity) exists between the resource transferred from the individual to government and the service (s) the individual receives. Although the *economic* incidence of a tax can be shifted, its statutory incidence cannot. In the case of a user charge or consumer tariff,<sup>55</sup> a fee is charged for a particular service, and the individual or household can adjust consumption of the service (that is their economic behaviour) in order to maximise utility given a budget constraint. They can usually decide whether to consume or not, and they also have discretion, partial or total, over how much they consume. Gildenhuys (1997:98) notes that user charges and consumer tariffs are both characterised by the absence of *compulsion*: “their payment is voluntary because the user or consumer has a choice in buying the services.”

**Table 4.3: Financing Options in a Two-Level Fiscal System**

	<b>Tax (T)</b>	<b>User Fee (U)</b>
<b>National Government (N)</b>	NT	NU
<b>Local Government (L)</b>	LT	LU

Table 4.3 presents a simple matrix of payment options for a local service in a two-level fiscal system, given this basic dichotomy between taxes and fees. Funding through national taxation (NT) requires the use of transfers from a central revenue fund to a municipality. In South Africa both the local equitable share (LES), an unconditional transfer, and various conditional grants to municipalities, such as the municipal infrastructure grant (MIG), are examples. The local taxation (LT) option refers to the use of tax instruments assigned to local government (in South Africa and many other countries primarily the property tax with its immobile base) to finance the service. In the case of LT, although

---

<sup>55</sup> As Gildenhuys (1997: 98) articulates it, a user charge is a charge for a service which “remains available for use by other users on the payment of the prescribed user charge”, that is, the service is not *consumed* in use, though it must be maintained. Health and fire services are examples. A consumer tariff, on the other hand, is used to recover costs for services requiring resources which “are completely exhaustible and must be continually replenished by new stock as consumption continues” (Gildenhuys 1997: 100).

financing is not linked directly to the service provided, revenue is generated within the service jurisdiction.<sup>56</sup> The local user fee (LU) refers to a municipality charging an individual or household for a particular service. The burden of financing is here imposed directly on beneficiaries of the service. The payment may consist of a 'flat fee' regardless of the amount consumed<sup>57</sup> but may also be linked to the amount consumed, such as the pervasive use of a rate charged per unit consumed. A third, slightly more elaborate option, also quite widespread in practice, is the combination of a once-off service connection fee with a rate per unit consumed.<sup>58</sup>

Given this framework and the externality-argument in favour of higher-level government subsidisation, two questions present themselves. Firstly: what kind of transfer will best achieve this aim? Secondly: how does one determine, or even plausibly estimate, the monetary value of externalities in a given instance? It is, after all, one thing to speak in general terms of positive externalities, and quite another to compensate a given municipality on the basis of their value.<sup>59</sup> Attempting to answer either of these questions is further complicated by the fact that potential transfer beneficiaries may behave strategically in order to receive transfers, which may or may not be desirable in its impact on the efficiency of the intergovernmental system. It is, in other words, very difficult to ensure that the

---

<sup>56</sup> A household may consume a subsidised local service and pay a property tax, such that the net effect approximates a market-related transaction, but conceptually and in an efficiency analysis the tax stands separate from the service.

<sup>57</sup> It may even be efficient to charge a flat fee if the costs (such as metering and trained staff) of implementing and administering a per unit billing system outweigh the likely additional revenue, especially if the social resistance to metering poor households is factored in as a kind of (social) capital loss. A flat fee corresponds to the notion of a lump-sum tax, which is efficient from an allocative point of view but could be highly regressive and regarded as inequitable.

<sup>58</sup> The notion of a national user fee (NU) refers to those local public services that are 'exported', that is consumed and paid for by non-residents, such as a local tourism levy which is used to fund the municipal maintenance of a game reserve. These are not particularly relevant to the rest of the analysis and will not be referred to further.

<sup>59</sup> The arguments and questions would apply equally to benefiting and non directly benefiting municipal residents.

desired changes in municipal budgetary decision-making do indeed take place. The cases of a conditional matching grant and a lump sum transfer are discussed in more detail.

Imagine a municipality which has promised residents that it will build 150 low cost (R 35 000) houses in the coming financial year, that is spend a total of R 5 250 000, representing 15% of its total budget for the year. The figure of R 35 000 roughly corresponds to total costs, including administrative and other overhead costs. The National Department responsible for housing has funds for supporting such initiatives, which it allocates to municipalities in proportion to their own housing programme expenditure (as conditional matching grants in other words), subject of course to the municipality presenting a viable proposal and to the aggregate amount the Department has available. The Department's objective in transferring such funds is to ensure that the municipality now builds *more* houses than the 150<sup>60</sup> it had intended to build given the self-generated resources at its disposal. The Department agrees to transfer funds to cover 10% of the cost of each house built, that is R 3 500 per house, or, if 150 are built, R 525 000, if 165 are built R 577 500, and so on.

From the Department's side, it assumes that its conditional matching grant should induce a budgeting change generating *more* houses than the planned 150, since the municipality has already committed R 5 250 000 to low-cost housing in its budgeting.

*The Department assumes that the original situation:*

Houses Built = Total Municipal Housing Allocation / Cost per house

---

<sup>60</sup> Such subsidies could also aim to reduce the cost to beneficiaries of a given housing commitment, rather than increase the number of houses being constructed. The reasoning behind such a consumer subsidy is something different, however. Depending on how the recipient government uses the transfer, a subsidy intended to increase the supply of houses, may well end up as consumer subsidies.

$$\begin{aligned} \text{Houses Built} &= \text{R } 5\,250\,000 / \text{R } 35\,000 \\ &= 150 \end{aligned}$$

*now changes to:*

$$\begin{aligned} \text{Houses Built} &= (\text{Total Municipal Allocation} + (.1 \times \text{Total Municipal Allocation})) / \text{Cost per House} \\ \text{Houses Built} &= (\text{R } 5\,250\,000 + \text{R } 525\,000) / \text{R } 35\,000 \\ &= \text{R } 5\,775\,000 / \text{R } 35\,000 \\ &= 165 \end{aligned}$$

To the extent that economies of scale generate decreasing unit costs for the construction company and they pass on some of this to the municipality (say R 1000 less per unit if 165 are constructed as opposed to 150), the number of houses built could in fact end up being 170 (R 5 775 000 / R 34 000).

The problem is that the Department assumes that the municipal allocation is fixed, when it is probably more realistic to assume that what is fixed is the political commitment to build 150 houses. Re-allocation of funds available to the municipality is likely to occur, subject to this accountability constraint, since in fact the Departmental intervention has changed the relative prices faced by the municipality for houses and other goods and services. There is no way to ensure that the conditional matching transfer leads to more houses being built, since there is no inherent reason to assume that the municipality values an additional 15 houses to the same extent that the department does. The municipality acts in its own self-interest, and so the problem being grappled with here is really the intergovernmental version of the familiar principal-agent problem. Structurally it does not matter whether the municipality is a disinterested authority with a different schedule of social priorities to the Housing Department or pursues the familiar objectives of larger salaries, more status, perks and the like. The municipality may, based on its honest assessment of local challenges, prefer to spend what is required to build the 150 houses it is politically accountable for, but

to utilise any surplus funds for maintenance of its roads. It may end up reducing local taxes or, less laudably, use the extra funds to increase the salaries of senior municipal staff. The municipality may simply re-allocate some of its *own* housing funds elsewhere in the course of the year. Given asymmetries of information and the related costs of acquiring additional information, there is no real way in practice for a national department to prevent this *even when transfers are conditional*. Economically, what has happened is that the *Department* is now funding some other municipal activity which would have remained under- or unfunded. Facing a less tight budget constraint, and being accountable for an unchanged output (to build 150 houses) the municipality may also, even with disinterested intentions, become less productively efficient in constructing houses. That is, it may well end up building 150 houses at a total cost of R 5 775 000, or R 38 500 per house. There is, again, no real way for a department to determine whether the R 3 500 additional unit cost is a matter of inefficiency or not, since arguments can always be summoned to account for the fact that actual costs exceeded budgeted costs.

This example should not only be taken to refer to *ex post* budgetary adjustment by the municipality once it has received a transfer. The typical scenario is probably better regarded as a fundamentally strategic one where the municipality *anticipates*, as it prepares its original budget, that transfers are available and adjusts its aims accordingly, generating a deadweight loss as resources are allocated differently.

Thus, the use of conditional matching grants potentially generates deadweight losses as a result of changes in the municipal budget, as well as potentially increased productive inefficiency (where no re-allocation takes place) as a result of a less hard budget constraint. For the Department's aims to be realised, in other words, no re-allocation and no increased productive inefficiency should occur.

A lump sum transfer to the municipality, such as the local equitable share (LES) in South Africa, is by definition not ring-fenced or earmarked and can be used by the municipality in accordance with the preferences of its residents. Such a transfer is allocatively efficient as it does not distort the relative prices of different municipal functions and therefore, as with a lump sum transfer to household and individuals, does not induce changes in economic behaviour with accompanying deadweight losses.

Given this difference between intended and actual use of conditional transfers, the technical problem associated with using a lump sum transfer to correct for externalities is of course that there is no reason to assume that such a transfer will be used on services generating externalities that are desirable from the point of view of the national government. Given that additional resources become available to the local government concerned, a benefit nonetheless accrues to the community generating the positive externality. A lump sum transfer, after all, is a form of general budget support rather than a means to influence budgeting through earmarking resources.

Thus, whilst it must be acknowledged that a municipal benchmark cost-recovery fee needs to be adjusted to or could be designed to reflect externalities in theory, in practice this may be very difficult to do. Given that South Africa uses transfers primarily to promote fiscal convergence, this may be less of a concern than in circumstances where transfers are required solely to reflect and compensate for externality effects.

### **g. The Difficulty of Determining Efficient Price for Infrastructure Services**

As discussed, Pareto-efficiency requires that the right amount of the right things be provided to the right people. In the case of private goods this entails that information be available to producers (in this case government) on what to produce and how much of it, and that consumers (households) maximise utility

subject to their budget constraints. That is, marginal costs and individual benefits need to be determinable: for this to obtain exclusion and rivalry in consumption need to be possible.

**Table 4.4: Rivalry and Exclusion Characteristics of Different Kinds of Goods**

	Private	Semi-Public (a)	Semi-Public (b)	Public
Rivalry in Consumption	Yes	Yes	No	No
Exclusion	Yes	No	Yes	No

With reference to the classification in Table 4.4 of different types of goods and services in terms of the criteria of rivalry and exclusion, household infrastructure services, such as water and electricity services, may be regarded as semi-public goods of type (b): exclusion is possible but no rivalry<sup>61</sup> in consumption occurs, that is *marginal* costs are negligible and can't be used as a basis for pricing. From an infrastructural point of view, adding the Tambo household does not entail that the service infrastructure is less able to accommodate the Jones household. The quandary in the case of such semi-public goods is acute: users *can* be charged, but it is difficult to determine, on an economically rational basis, how *much* they should be charged. One possibility, discussed further below, is that average costs be used to determine price where marginal costs are very close to zero.

From an efficiency perspective, the appropriate financing *principle*, however, must surely be the benefit-principle where users of a service should contribute to its financing in proportion to the benefit they derive from it. Those who do not benefit should consequently also not be required to contribute to it – and are excluded. Where, as with household infrastructure services, exclusion is

---

<sup>61</sup> This only holds if congestion is not taken into account. The case of congestion is discussed in more detail below with reference to the problem of pricing services where the average load falls below productive capacity but the peak load breaches this barrier.

possible, specific beneficiaries can be *identified*. Benefit-financing is taken in this study to refer to a conception of benefit along marginal analysis lines. That is, it is assumed that declining marginal benefits in consumption apply and that household decision-making is constrained by its available budget and structured through the interaction of income and substitution effects. Benefit-financing is likely to generate efficiency in *consumption* because economically rational, utility-maximising households will consume quantities of a service until the benefit or utility they derive from the last unit equals the price they pay for that last unit. Where benefit exceeds cost more units will be consumed and where cost exceeds benefit less units will be consumed. Bird (2004: 2) clearly articulates the basic argument for user charges for at least some public services on the basis of efficiency rather than revenue-generation: “The main economic rationale of user charges is thus *not* to produce revenue but to promote economic efficiency. Well-designed charges achieve this goal *both* by providing information to public sector suppliers as to how much clients actually are willing to pay for particular services *and* by ensuring that what the public sector supplies is valued at least at (marginal) cost by citizens.” Social over-consumption is likely to accompany services provided at no cost; the effect of this is particularly damaging where, as in the case of public infrastructure such as roads and airports “the resulting crowding is taken as a signal that there should be even more of the under-priced service provided” (Bird 2004: 2).

However, from the supply side the municipal provider cannot rely on marginal cost determination in setting price, at least not in any simple manner. This is because the marginal approach can be used for some components of the service but not for others. Marginal cost pricing can be applied to the bulk resource purchase and the maintenance and operations components of infrastructure services. M&O responsibilities usually reside with the municipality, which needs to finance these activities out of self-generated resources or equity/convergence orientated transfers from national government. M&O pricing requires that the municipality assess, for a given financial year, the costs of maintaining the



existing infrastructure stock and operations, and the additional operations costs of admitting new users to a system which is running below its full capacity, that is which can admit additional users without incurring additional capital expenditure.

However, the South African context remains characterised by infrastructure backlogs and urgent demand for capital expenditure on infrastructural expansion. Here, the fundamental problem of marginal cost as a basis for setting price is that large infrastructure networks are characterised by decreasing average costs as more users are added and the service output generated by the network increases. This essentially stems from the large *initial* outlays required. Initial outlays remain a large proportion of total costs up to some point before the capacity threshold of a facility is reached. As a result, average costs decline over this entire range of output and, by definition, marginal costs remain less than average costs and, as discussed above, are often negligible from the perspective of the additional demand represented by particular households.

Even if marginal costs are not entirely *negligible*, pricing problems arise whenever each unit generated costs less on average than the unit that preceded it, that is when average cost is declining. For a producer to remain financially viable, total revenue and cost need to be equalised: for a given output, this means average cost and revenue also need to be equal. But if marginal cost lies below average cost, then pricing at marginal cost will mean that a loss per unit is sustained over the entire output, and a firm setting prices on this basis will be losing money.

A number of possible solutions exist where marginal pricing leads to shortfalls due to large economies of scale for services which nonetheless should be provided since they result in net benefits to society. The strongest alternative to marginal cost pricing is to use average cost instead. In essence this approach “estimates the total financial cost of providing a particular service and divides by the number of units currently provided to obtain the appropriate user charge”.

(Bird 2004: 8). The approach can readily be adjusted to also include price-setting for expansion, since an estimate of such costs incurred divided by the number of intended new users would provide a usable average cost projection. The problem with this approach is that it is not economically efficient, since additional resources are not allocated on the basis of opportunity costs. Average costs lie above marginal costs across the entire relevant range: setting price to equal this average cost therefore means that users are being *over-charged* and socially inefficient under-consumption will result.

Pricing on the basis of average instrumental cost (AIC) is an attempt to link the unit price paid by the user to both operating and capital costs incurred in providing services such as water and electricity.<sup>62</sup> Marginal cost is used to determine the fee for the additional maintenance and operations expenditure incurred per additional user admitted to the network. The discounted capital cost per user of expanded output is added to this to get an approximation of the additional capital and operating costs represented by each additional user. The multi-part tariff is a reflection, on the pricing side, of the AIC means of computing cost, since the multi-part tariff typically includes a 'connection' fee intended to cover average capital cost, as well as a per unit rate which covers costs of operations, maintenance and resources consumed. The connection fee need not be a once-off fee but could also be thought of as an annual capital fee to be paid by households.

Capital expenditure of an infrastructural nature clearly requires large initial outlays and confers benefits over an extended period of time. It should therefore be financed through borrowing to ensure that the financing burden is distributed fairly across present and future users. Thus, the amortisation schedule for a particular infrastructural asset represents the cost of capital and can be used to calculate average capital cost by dividing it by expected number of users. Once

---

<sup>62</sup> Kahn, A (1970) remains a clear exposition of the economic aspects of cost and its relation to the regulation of price.

the asset has been paid off, the use of AIC-based pricing is no longer required since only marginal costs remain applicable.<sup>63</sup> This would suggest that prices paid by users should be lower once the capital costs are paid.

A second problem of pricing infrastructure relates to how to deal with problems which occur when the generation and transmission system is operating at close to its peak capacity. Fluctuations in demand mean that a system operating on average at less than peak capacity may nonetheless be strained past the breakdown threshold at times of peak demand, with consequences which can be both politically and economically costly. It has been mentioned that the problem here is that no incremental adjustment in 'plant-size' can be made, since the system is at or near its capacity limit. As Fox and Edmiston point out, efficient allocation of available units during peak hours can be achieved if price is raised above marginal cost during these periods: "In the face of excess demand, the user charge must be raised above marginal cost to the level where excess demand is zero. The resulting price is efficient because it allocates the services to those who place the highest valuation on them (as evidenced by their willingness to pay)" (Fox & Edmiston 2000: 9). This assumes that such a price can be calculated and is politically acceptable. However, continued congestion will eventually generate various inefficiencies and the long-term option will have to be additional infrastructure investment financed in the manner discussed above.

---

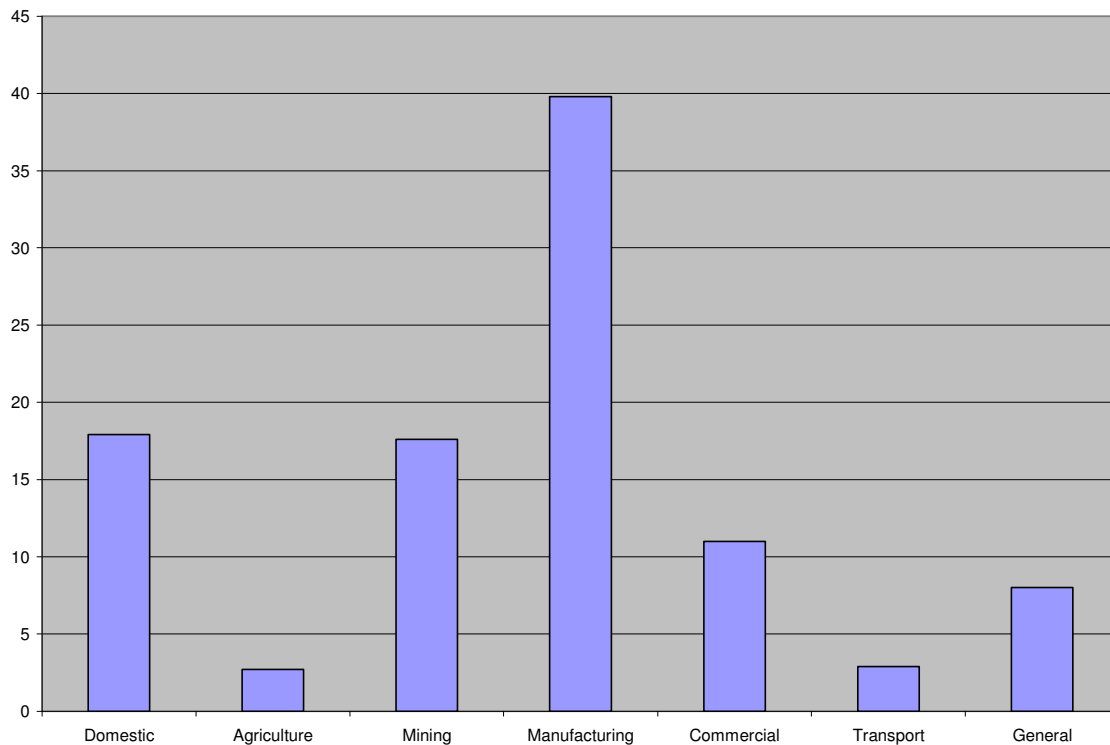
<sup>63</sup> A complicating factor is whether the cost of financing the asset or the monetised benefit to the user should be used to calculate the capital fee. Of course the benefit is inferred from the cost, but some decision needs to be made under circumstances where the asset continues to confer benefits after it is paid off. Imagine an asset which will be paid off after five years but will confer benefits for ten years. User 1 is charged an annual capital fee from year 1 to year 5, calculated on a discounted average cost of capital basis. User 2 is connected in year 5: should user 2 be charged a capital fee for one year, representing the last year of repayment, or should user 2 be charged for five years, to make it equitable with user 1 and relating to the benefit received?

The broad approach to price-setting from an efficiency perspective articulated here clearly also serves as the point of departure in assessing the extent to which the prices charged by utilities to municipalities and the prices charged by municipalities to households are efficient. To the extent that they deviate significantly from prices predicted by this approach a presumption exists that they do in fact constitute an abuse of market power. The next section concretises the discussion by looking in some detail at the case of pricing electricity in South Africa.

#### **h. Electricity Pricing in South Africa: An Application of Preceding Considerations**

Generation and transmission of electricity in South Africa are dominated by Eskom, a limited company which is wholly state-owned. Eskom generates approximately 96% of South Africa's electricity, of which approximately 92% comes from coal. Eskom also owns and operates the national electricity transmission grid. Although households make up the largest component of electricity demand by number of service points, they do not constitute a large percentage of total electricity demand in South Africa, as figure 4.3 shows for 2003.

**Figure 4.3: Percentage Distribution of Total Electricity Sales, 2003**



Source: National Treasury (2006: 22); own calculations.

As regards the distribution of electricity, this is conducted by Eskom as well as by some of the larger municipalities. The *Local Review* points out that 171 local municipalities and 2 district municipalities are currently licensed electricity distributors (National Treasury 2006: 31). The same publication reproduces a table from the National Electricity Regulator Statistics for South Africa, 2003, which indicates that Eskom and municipalities have an approximately equal share of domestic customers: in 2003 Eskom had 3 376 276, or 49.3% of total customers, and municipalities and others had 3 470 054, or 50.7% of total customers (National Treasury 2006: 31). However, the volume of sales provided by each is significantly different. Of the total 34 074 593 MWh sold in 2003, 76.6% was sold by non-Eskom distributors, that is municipalities and others, whilst 23.4% was sold by Eskom. This divergence suggests the extent to which, in the domestic sector, Eskom has been responsible for servicing low-

consumption households, that is poor and, one can assume, rural households, who would fall outside the scope of municipal provision.<sup>64</sup>

From the perspective of assessing efficient pricing of electricity, two aspects need to be considered, namely the price at which the bulk resource is sold to municipalities by Eskom (that is wholesale price determination) and the price at which municipalities sell electricity to households, that is retail price determination. Both Eskom and municipalities do of course possess at least potential market power as suppliers and are in a position to charge socially inefficient rates to households.

Eskom bulk electricity prices are regulated by the National Electricity Regulator of South Africa (NERSA). In its most recent articulations concerning regulatory methodology, NERSA emphasises the use of a 'rate of return' methodology. An electricity utility's pricing should be derived from its revenue requirements, and these in turn "should be equal to the *cost to supply* electricity plus a fair return on the rate base" (NERSA 2007:1). Including a provision for a fair return on the rate base recognises that the utility will have to borrow in order to finance expansion, and that the cost of borrowing should be factored into the determination of revenue requirements: "*Return* means the amount of money to be earned by the suppliers of capital for their investment in the business. ... The return is calculated as the weighted average cost of capital (WACC), where the cost of capital is represented by the cost of debt and the cost of equity invested in the regulated entity" (NERSA 2007: 1). The document provides a useful summary of the advantages and disadvantages of this approach to regulation. Key advantages include the comparative simplicity of the approach and its flexibility under changing economic circumstances. Key disadvantages are mostly associated

---

<sup>64</sup> Thus the tentative conclusion that a limited company has been at the forefront of extending electricity services in commercially marginal contexts! This suggests either the inefficiency of the municipal budgeting process or the paucity of resources going to that sphere or the lack of basic institutional ability in many instances.

with the alignment of price determination and cost and the incentives this potentially creates to firstly not reduce costs and secondly overestimate costs. A secondary disadvantage concerns difficulties and disagreements which might arise around determination of a fair rate of return (NERSA 2007: 2).

In one respect it is quite easy to determine whether Eskom's pricing practices are efficient or not: since Eskom supplies electricity to both its own distribution section and to municipalities, these can be compared on the assumption that Eskom would tend to engage in efficient internal pricing. As the *Local Review* notes: "Eskom's prices for bulk electricity provided to municipalities are generally higher than the prices charged to Eskom distribution. The Eskom retail tariff structure is designed to allow Eskom to subsidise low-income and rural electricity customers on a national basis" (National Treasury 2006: 30). The argument, as summarised here, is that Eskom is overcharging municipalities in order to ensure it has adequate revenue to provide electricity to low-income households and unconnected households in municipalities which would not be in a position to extend and maintain electricity services themselves. By all accounts Eskom has been successful in its programme of electrification. Nonetheless, it is a curious circumstance, perhaps not adequately remarked on, when a state-owned enterprise is tasked with fulfilling an equity function by over-charging the local sphere of government! It should be noted that, in effect, this approach amounts to a form of inter-municipal redistribution: the larger, wealthier municipalities that are in a position to extend distribution and connector infrastructure and provide the service to households are being overcharged for bulk resources in order for Eskom to provide electricity services in poor municipalities. This represents a pragmatic policy solution to the lack of capacity, be it financial, technical or administrative capacity, in many municipalities, but it is of concern nonetheless because the subsequent distortions brought about are difficult to predict, as in the case of any manipulation of price away from market-values.

Bahl (2001: 2) mentions that in South Africa municipal electricity services are often priced and operated so as to generate a surplus which can be used to finance other services. This statement is borne out by the information on recoverable costs for electricity 2005/2006, as well as the information found in the non-financial census of municipalities. In the case of some larger volume providers, income exceeds total cost by a fair margin. Thus, for example, in Nelson Mandela Bay  $(\text{income} - \text{total cost} / \text{total cost}) \times 100$  equals 14.8%. In Mangaung the same ratio is 23.5%. In Gauteng, two of the three metros exhibit a similarly large difference: Ekurhuleni has a ratio of 18.9% and Tshwane of 9.7%. Johannesburg, on the other hand, does not generate a surplus on its electricity account: the ratio here is 1.2%. Similarly, the larger KwaZulu-Natal municipalities also do not generate surpluses on their electricity accounts: eThekweni's ratio, for example, is given as zero. Cape Town also has a zero ratio.

Neither the generation of surpluses, nor a balanced account, nor for that matter deficits on the electricity account, is automatically a positive or a negative phenomenon from a price efficiency perspective, since such figures say very little about the productive and allocative efficiency of the electricity distributor. Similarly, comparing the average price per unit bought and unit sold for municipalities is not particularly revealing because the average rate may mask vastly different pricing schedules. South African municipalities do, however, lose a great amount of electricity units in distribution, which must surely be the largest aspect of municipal productive inefficiency in electricity services. Nationally, twice as many units are lost in distribution as are distributed as free basic services (4 243 701 000 kwh as opposed to 2 012 923 000 kwh).

## **i. Conclusion**

This chapter has attempted to apply the well-developed concepts of efficiency and market failure to the specific topic of household infrastructure services. The likely existence of economies of scale in the infrastructural component of such services, and the fact that they can be assumed to generate positive



externalities, provide grounds for government involvement in the interests of efficiency. Infrastructural services, however, are a form of semi-public goods: exclusion is technically possible whilst rivalry in consumption does not feature unless a 'congestion threshold' is breached. Given the possibility of exclusion, the use of service charges, which align the financing burden with the benefits received by households, becomes the efficient option. Efficient pricing, however, is not that easy to determine for infrastructural services, since their economy of scale characteristics mean that average costs exceed marginal costs across output quantities and the application of a marginal cost approach to pricing would result in losses. Average incremental cost, combined with increased rates during peak hours (in effect a congestion charge), can lead to fairly efficient outcomes. However, as the concluding analysis of South African electricity pricing suggests, each particular context will come with its own pricing challenges in the pursuit of efficiency. There is, furthermore, no reason to assume that efficient pricing of infrastructure services in a developing country such as South Africa will render such services affordable to a socially acceptable proportion of the population. Thus, the efficiency analysis must be complemented with an equity perspective. This is developed in the next chapter.

## **Chapter 5**

# **Access and Affordability of Household Infrastructure Services in South Africa**

### **a. Introduction**

This chapter is concerned with the equity aspects of financing and pricing infrastructure services. The first section clarifies what is meant by equity and links the concept to the nature of the obligations imposed on government by the South African Constitution's Bill of Rights, specifically the extent to which it requires government to take reasonable measures to ensure a lifeline amount of service to all citizens. The Constitutional obligation requires Government to function as a 'provider of last resort', that is to ensure that individuals and households who would be excluded if a market-orientated ability to pay criteria were applied have access to certain services in quantities and quality that are constitutionally acceptable.

Having established that all citizens are constitutionally entitled to a minimum core basket of services and what that entails, the next section uses access information, available income data and the efficient pricing and financing discussion of the previous chapter to ask after the needed scope of government intervention in South Africa. This means, essentially, asking after the extent to which Government's Constitutional obligations would be met if it followed a *laissez faire* approach or intervened only with the aim of enhancing efficiency, rather than with more explicitly redistributive or social justice intentions. A redistributive approach appears to remain essential over the short- and medium-term given the poverty of residents and the large expenditure-revenue mismatches in many municipalities (that is the costs of exercising their functions far exceeds the revenue that is or can be extracted from the revenue sources available to them). From this basis the next chapter examines three specific equity-orientated measures currently being utilised in South Africa, namely the

local equitable share (LES), the policy of free basic services (FBS) and the use of cross-subsidisation in particular instances.

### **b. Equity and Social Welfare**

While it is the task of citizens and their elected representatives to adjudicate between various policy options which all claim to increase social welfare, the economist can contribute to the quality of such adjudication by helping to explicitise the positive aspects of the debate. An economist should, for example, be able to provide some quantitative answers to the question of whether, in South Africa, the efficient price determination discussed in preceding chapters renders infrastructure services unaffordable to, say, more than 30% of households that have been 'connected'. Such households are, correctly, recorded as households with access to electricity in survey results and the like, but there may be reasons to question their quality of life if pricing renders consumption unaffordable. Terreblanche (2002: 28) makes this simple but compelling point. He starts by noting significant instances of progress in access to services in the post-1994 period, but then clearly draws the distinction between *access* and *affordability* of decent-quality services in pointing out that:

“Many...electricity, water and telephone connections are cut off every month because users cannot afford to pay for them. Many of the houses built are of poor quality. In 1999 the department of water affairs admitted that many of its water provision projects had fallen into dysfunction or disrepair...The improvement of health services in rural areas has led to a dramatic drop in the quantity and quality of health services in urban areas.”

Welfare economics is concerned with “the study of the social desirability of alternative arrangements of economic activities and allocations of resources” (Bannock et.al. (eds.) 2003: 404). It thus requires consideration of both efficiency *and* the basis on which to prefer certain efficient states over others. Though the efficiency approach to financing and pricing which was developed in the previous chapter establishes benchmark prices for household infrastructure services, the

rigid mechanical application of such pricing would in all likelihood result in socially and politically contentious outcomes. It is consequently necessary to articulate the implications of efficient pricing for affordability across different household income groups. Put another way: the notion of 'access' is far from univocal and, when used simplistically, can in fact too easily obscure significant distinctions between a household merely being connected to a service grid, and that household deriving the benefits of the service at an affordable cost. .

'Affordability' is, in a sense, at the centre of this chapter, yet is an inherently contestable, if not vague, notion. A number of options exist: cost may be regarded as appropriate if, for example, the average SA household pays an equivalent portion of disposable income to, say, electricity as a comparable household 'internationally', but such comparisons require so many heroic assumptions that they are largely useless. Otherwise, it can refer instead to whether the service is priced such that some minimum living level amount (whose determination needs to be aligned with the Constitution) is available to all households; this is the free basic services policy approach currently followed by Government.

From Chapter 4 it follows that the Pareto-criteria provides no means to assess the relative acceptability of various distributions of income: an infinite number of Pareto-efficient outcomes exist, corresponding to an infinite number of initial distributions of income. Production and consumption exchanges are justified only up to the point where all exchanges which benefit someone without making anyone else worse off have been completed. It may be, of course, that the initial distribution of income is such that the outcome, once all Pareto-exchanges have been made, is also equitable. This is not, however, a necessity and is, furthermore, unlikely in an economy characterised by historically grounded structural inequalities of wealth and income such as that of South Africa.

What is required, then, is some means to adjudicate between different Pareto-efficient outcomes on the basis of the maximisation of social welfare, understood here as the attainment of the highest possible measure of welfare given the resources available to society. An explicit framework for comparing alternative distributions of income is necessary in order to legitimate interventions to adjust benchmark cost-recovery prices.

Welfare can most simply be regarded as a positive linear function of the utility of individuals:  $W = f(U_1, U_2, U_n \dots)$ , with utility in turn being a positive function of income. In the simplest possible conception every rand of income would bestow an identical amount of utility on the individual who possesses it (one could think of the income-utility exchange rate as being unity) and every individual's utility would be weighed equally in determining total social welfare. It is also possible, of course, to introduce a more nuanced model where income has a declining marginal utility and the utility of different individuals are weighed differently in determining social welfare.

Clearly, if it is assumed that income has a declining marginal utility, then redistribution is justified between a richer and a poorer individual. In the simplest utilitarian approach to social welfare,<sup>65</sup> as Rosen (2005) for example discusses it, given the assumptions that total income is fixed, income is characterised by linearly declining marginal utility and the income-utility functions of individuals are identical, redistribution is sanctioned up to a point of complete income equality. These assumptions are however unrealistic and overly restrictive. Individuals value income to differing extents, income is not clearly characterised by declining

---

<sup>65</sup>The link between social welfare and individual utility can also be conceptualised in a more complex manner. Rawls (1999), in his influential *Theory of Justice*, argues that the aim should be not the maximisation of aggregate social welfare, but to maximise the welfare of the poorest members of society. This, he argued, would be most compatible with the economic dispensation people would select if they were making decisions behind a 'veil of ignorance' such that they did not know what circumstances they would find themselves in.

marginal utility, and the impact on the work-leisure tradeoff (more leisure will be preferred) for the taxed individual is likely to reduce the total amount of goods produced and available for consumption in the economy, and in this way reduce social welfare. It is, perhaps most fundamentally, also difficult to quantify utility cardinally and thus any direct interpersonal comparison of utility is not feasible. This last problem is a social equivalent of the problems associated with trying to assign cardinal values to different goods and services for an individual consumer.

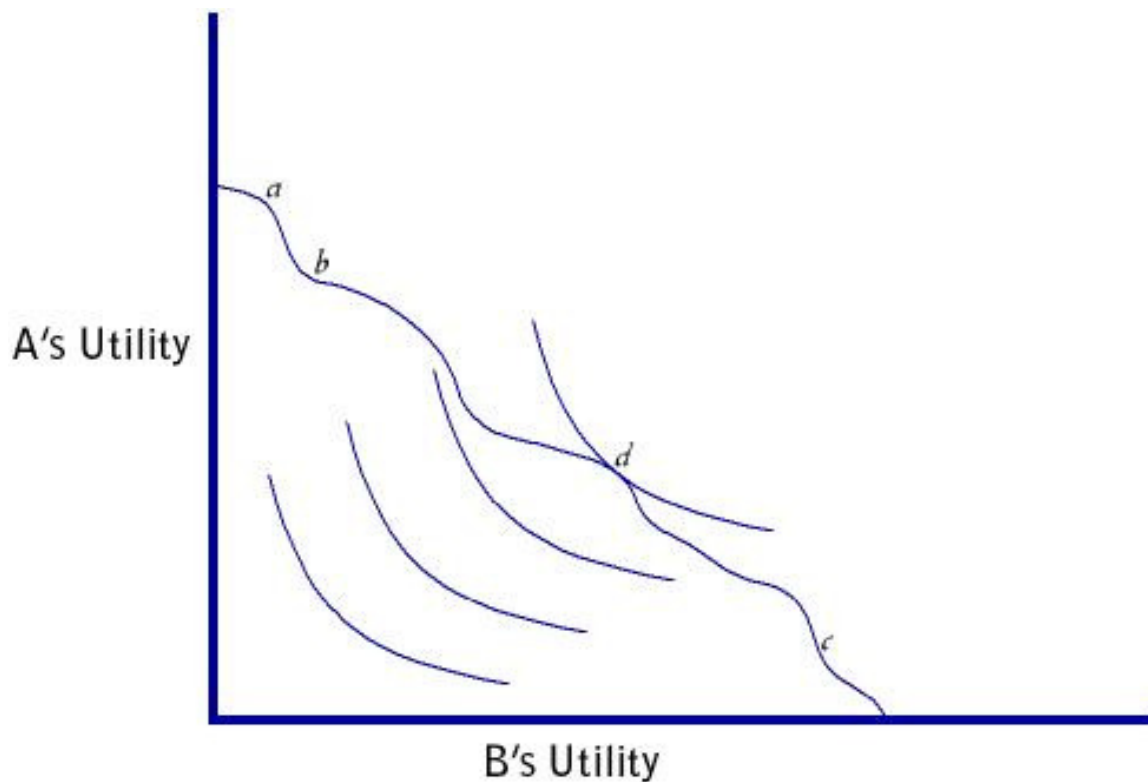
The use of social indifference curves, based on an ordinal rather than cardinal approach to utility, are often advocated as a possible solution to the social welfare problem, in much the same way that consumer indifference curves are used. Social indifference curves recognise that it is impossible to quantify the total utility associated with a particular outcome, but that it should be possible to *rank* various outcomes and specify various sets of outcomes between which society is *indifferent*. A social indifference curve can be used to map out the distributions of utility between individuals which society regards as being equally acceptable.

As income becomes more unequally distributed, an ever-increasing improvement in utility for the rich is required to compensate the loss to the poor, if society is to remain on the same indifference curve. The standard utility-possibility curve, as represented in figure 5.1, shows the possible distributions of utility given two individuals, two products and two factors of production. Any point on the frontier represents Pareto-efficiency, whilst points inside the curve represent potential Pareto-consistent improvements and points outside the curve represent currently unobtainable levels of social welfare given available resources and preferences.

The challenge for welfare economics is to determine which of points a, b and c represents the highest social welfare or 'bliss point'. This requires construction of social indifference curves which in turn requires deciding on the relative weight of A's utility compared to B and B's utility compared to A. A point of tangency

between the highest obtainable social indifference curve and a utility possibility curve represents a 'bliss point': the highest obtainable total welfare consistent with Pareto-efficiency (also referred to as the Pareto *optimum optimum*). In the figure this bliss point is given as d.

**Figure 5.1: Standard Utility-Possibility Curve**



In the 'new welfare economics', transfers of income which have an aggregate welfare-enhancing impact are permitted even if some people are made worse off. More formally, a proposed change meets the Kaldor-Hicks criteria if the winners could in theory compensate the losers and still be better off than initially. It is accordingly also called the compensation principle and underlies most forms of cost-benefit analysis. The *Penguin Dictionary of Economics* summarises the compensation principle as: "The principle that total economic welfare increases from a change in the economy if those who gain from the change could compensate those who lose from it to their mutual satisfaction. It is not necessary

for money transfers to actually take place. However, the principle has been criticised in this respect because, without actual transfers, interpersonal comparisons of utility of money are implied. Actual transfers are required if individuals are to reveal the total worth they place on their gains and losses” (Bannock et al (eds) 2003: 63, 64).

Conventional theory holds that there is a determinate threshold to the use of redistribution to enhance welfare which is given by the fact that, beyond a certain point, those who lose as a result of redistributionary measures will begin to alter their economic behaviour in a manner which reduces the total product available to society for consumption. More specifically, past a certain threshold higher tax rates (to finance greater redistributionary expenditure) would reduce the opportunity cost of leisure and alter the relative preferences for work and leisure of this group. Less work offered would entail a shrinking of the utility possibility frontier.

A country facing high levels of inequality and poverty would, in some phases of its development, probably be characterised by growing inequality, as argued seminally by Simon Kuznets and associated with the Kuznets curve (Kuznets 1955 in Meier & Rauch 2000: 382 – 386). At a higher level, however, it might well become *more* efficient given some redistribution through fiscal operations, since high inequality tends to erode social capital and generate conflict, which in turn is likely to have an adverse effect on productivity through reduced savings, investment, educational attainment and the like.

This argument implies that the marginal rate of efficiency-equity substitution is not only not constant, but also that the parameters of the curve itself vary with the degree of output and income of the economy, that is, generally, its stage of development. In a richer economy equity may come cheaper, in the sense of requiring less sacrifices of efficiency; in a poorer economy the pursuit of equity may require greater sacrifices in efficiency.



To the extent that equity-related government measures impact negatively on efficiency, this is not only because the relative opportunity costs of work and leisure change, but also because productive resources are removed both from the private sector and from those departments in government tasked primarily with the promotion of market efficiency.<sup>66</sup> The assumption that self-interested behaviour is likely to be more efficient behaviour (though also generating more externalities) and therefore that the private sector is likely to be more efficient in many realms than the public sector, goes back at least to Adam Smith's invisible hand analogy. In this study the basic assumptions of this view are assented to. Redistributive measures, then, must be defended through establishing that market transactions and government interventions aimed at efficiency do not create social outcomes that satisfy constitutional requirements. Put another way: the crucial question is whether providing basic services affordably to all, in line with the Constitutional vision of human dignity, requires a degree of redistribution which reduces the total product available, or in fact crowds out other public spending in a way which compromises the attainment of other, Constitutionally mandated, public finance objectives. Part of answering such a question is to explicitise the relationship between 'equity' and the obligations imposed on the state by the Constitution.

### **c. Equity and Constitutional Economics**

This study regards the equity dimension of assessing pricing and financing of infrastructure services as being indissolubly linked with the nature of the obligations imposed on government by the South African Constitution. However, the conceptual link between issues of equity as understood in conventional welfare economics, and the nature of the claims on the fiscus that the Constitution represents, are not simple matters.

---

<sup>66</sup> Such as the Department of Trade and Industry in South Africa.

The Constitution's Bill of Rights establishes a justiciable obligation on government to respect, protect, promote and fulfill a number of socio-economic as well as civil and political rights. With the exception of provisions relating to children, found in Section 28 of chapter 2, the nature of the obligations imposed on the state are *qualified*, rather than unqualified: government must "take reasonable legislative and other measures, within its available resources, to achieve progressive realisation of...these rights" (Constitution of South Africa, Act 108 of 1996: Sec 27 (2)). A number of legal decisions have interpreted and developed this articulation further, and have engaged particularly with the meaning of *reasonableness*.

The Constitution's Bill of Rights is not concerned directly with equitability but with ensuring access to basic services to all South Africans. This emphasis is certainly not inconsistent with a programme of active and wide-ranging redistribution, but it would arguably also not be inconsistent with static or rising levels of inequality *if* these were simultaneously accompanied by such marked economic growth and increased transfer amounts that all citizens became better off and had access to the material resources required for human dignity. In other words, the Bill of Rights is directly concerned with the elimination of absolute poverty, broadly understood to include not only income-measures but also service access and household transfers, rather than the *unqualified* reduction of inequality.

Certainly, the fact that the Bill of Rights requires the state not only to respect and protect the rights it speaks of but also to promote and fulfill them, establishes grounds for active measures on the part of the state to ensure services to those who are not able to obtain them through exchanging employment for income in the market. Such services would have to be financed predominantly through taxation of the employed portion of the population, whether those currently employed or, if debt-financing is used, those employed in the future. Thus, attempts to use the fiscus to achieve constitutional obligations inevitably have a

redistributionary incidence (that is incomes net of government spending and tax incidence will be more equal than incomes before government taxation and spending) but this does not necessarily mean that inequality will be reduced over time (that is comparisons of secondary income between years) if the incomes of the wealthy increase at such a faster rate than that of the poor even after the redistributionary budgetary incidence is factored in. In this study, then, the approach to equity is informed by the recognition that ensuring constitutionally mandated services to all requires in-year redistribution (secondary incomes will be more equal than primary ones), but that this may or may not be accompanied by reduced inequality between years.

#### **d. Equity, Access and Affordability of Utility Services in South Africa**

Regarding the delivery of household services to the historically excluded or under-serviced, the basic problem South Africa is confronted with is to provide such services on an *affordable* yet also sustainable basis. The simultaneous attainment of affordability and sustainability needs to prevail for both the maintenance and administration of *existing* services and the extension of services. That is to say, the need to finance capital expenditure affordably and sustainably in order to eradicate backlogs is a central problem in addition to the maintenance of existing service infrastructure.

Household infrastructure services such as water and electricity are potentially market-exchangeable. They can be provided at a price determined solely by the interaction of supply and demand. However, potential monopoly power would in many instances tend to distort pricing on the supply side, and externalities arise from the actual provision and consumption of the service. Rates may be higher than the socially optimal amount if a narrow cost-recovery approach is taken, since financial cost-recovery will not necessarily incorporate consideration of the positive externalities associated with adequate consumption of potable water (e.g. health externalities) and electricity (e.g. it is easier for household children to

study and be helped in their studies by adults). A broader approach would either allow for price differentiation between market segments, i.e. cross-subsidisation within a service or between services, or would be based on a shared funding. These are the *efficiency* grounds for considering regulation and public ownership as well as the use of intergovernmental financing mechanisms, rather than relying solely on the municipality and its households for financing.

If a household service in a particular municipality is efficiently priced and provides the service at a trivial charge relative to total household income, this need not concern us, since it simply implies that provision in this jurisdiction does not entail significant opportunity costs. From the perspective of social welfare, no over-consumption can result, *so long as* the price is allowed to adjust as the opportunity costs of production and consumption increase. But, though equally economically rational, if such services are efficiently priced and result in the service being provided at an exorbitant charge relative to average household income, political and constitutional considerations require some way of rendering the service affordable to all households, if only in basic amounts at a basic service level. Similarly, if municipalities are not able to recover costs either through service charges, cross subsidisation or the use of other revenue sources, then the use of transfers may become justified. It is necessary to distinguish clearly, then, between a narrow conception of access and the more complex debates which stem from also attempting to consider the affordability of services to households and municipalities and therefore the welfare implications. The basic issue may be regarded as that of whether a household can *afford* to pay for the service it now has access to. Access means a household has, say, a tap in their yard. Affordability means that the price they pay for the installation of that tap and the use of water it allows does not command an inappropriate share of their household budget. In research conducted on the provision of electricity in Soweto, the fundamental distinction between service availability and access properly understood is articulated clearly: “Being attached to the electricity grid...is only meaningful insofar as one can afford to pay for the service” (Fiil-

Flynn, 2001: 5). These questions are discussed further below with reference to water provision.

Where water is provided entirely 'free', say because a municipal metering and billing system does not exist, then questions must also be asked regarding the sustainability of service provision for the municipality, that is whether the municipality can afford this. Fankhauser and Tepic (2005) state the basic problem clearly with reference to transition Eastern European economies, though their paper appears in most respects equally relevant to South Africa:

"Improvements in the quality of infrastructure services for both rich and poor consumers – better access, more reliable supply, less wastage...will mean higher end-user prices and better billing and collection...and...an increased financial burden for low-income households, in particular if adequate social safety provisions are lacking" (Fankhauser, S & Tepic, S 2005: 1). To develop and concretize these questions further the example of water is used in the rest of this section.

According to the *2005 General Household Survey* (GHS), 68.4% of South Africans had access to piped water in their dwelling or on site in 2005 (Stats SA 2005: xxviii). This represents a steady improvement in access through 2002 (66.1%), 2003 (67.3%) and 2004 (67.8%).

As is invariably the case with such averages in developing countries, however, they do tend to mask significant regional differences. Table 5.1 below duplicates the relevant table from the GHS.

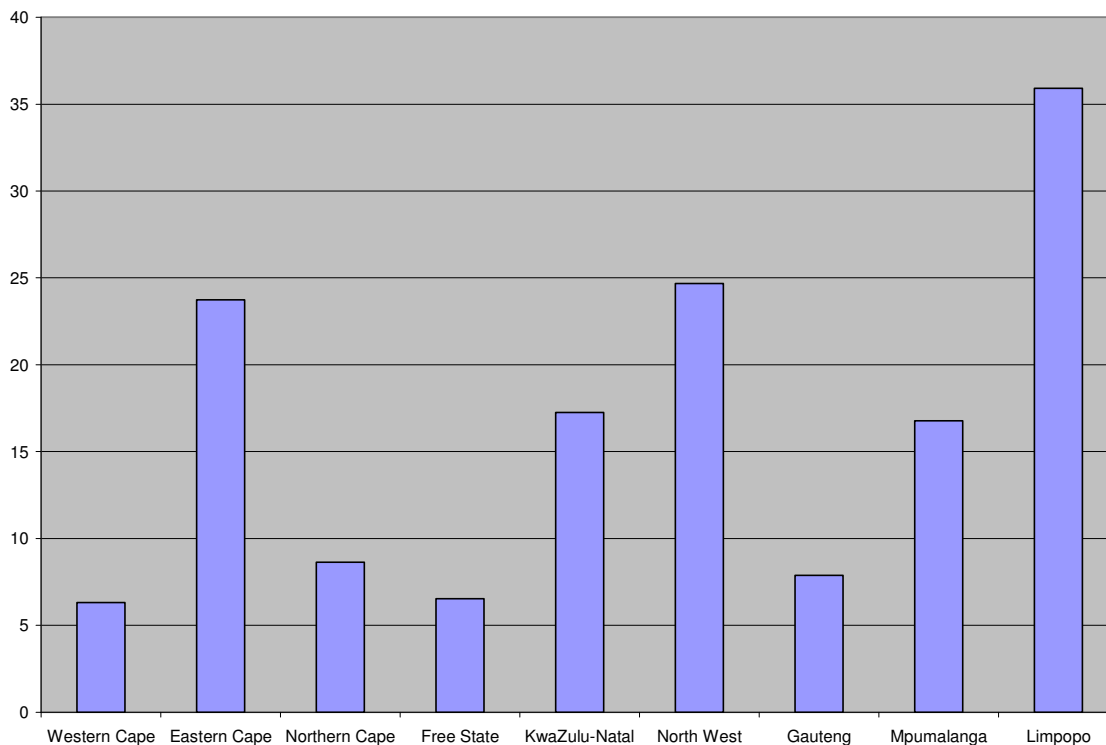
**Table 5.1: Access to Piped Water in Dwelling or On-Site, 2005**

	EC	FS	GP	KZN	LP	MP	NC	NW	WC	SA Average
<b>2002</b>	35.3	82.0	91.9	56.3	40.6	70.5	88.0	57.0	91.1	66.1
<b>2003</b>	36.1	84.1	92.0	58.5	39.2	70.6	88.6	60.4	89.2	67.3
<b>2004</b>	36.0	85.6	92.1	58.6	38.5	71.8	89.3	60.2	92.4	67.8
<b>2005</b>	41.6	88.7	89.1	62.1	36.4	67.1	86.8	60.5	92.0	68.4

Source: StatsSA (2005a: xxviii).

Of these 68.4% of households in 2005, slightly more had access in the form of piped water in their dwelling (4 926 000 or 56.6%) as opposed to access in their yard (3 775 000 or 43.4%). Of those households who did not enjoy piped access in 2005, the predominant form of access in most provinces was a public tap. In some provinces, as the figure below illustrates, a sizeable percentage of households continue to rely on this source of water provision.

**Figure 5.2: Percentages of Households Using Public Tap as Main Water Source, by Province**



Source: StatsSA (2005a: 42), own calculations.

The only other notable sources of water, and this only in some provinces, are those of flowing water/stream/river, and an off-site or communal borehole. The Eastern Cape is especially notable in this regard in that 28% of households in 2005 relied on a spring, river, stream or other form of flowing water as their main source of water.

The problem with the way data is presented in the GHS for access via public taps is that no distinction is made between public taps more than 200 meters away and public taps within 200 meters of a household. In other words, the data does not allow evaluation of access in terms of the adequacy threshold for the 'basic' water services designation in relevant water policy documents, which specifies a minimum service standard of a public tap within 200 metres. However, a subsequent table (GHS 2005: 44) is headed 'Households without water in the dwelling or on site, by time taken to reach the water source and population group of the household head' and does allow some conclusions regarding average distance from a public tap for households who primarily use these means. Only 35.5% of households nationally who lacked access to water in their dwelling or on site in fact had access to a public tap at a distance (200m or 2 minutes) which *meets* the specified basic service level. Thus, it appears that households who do not have access to water services within their house or yard are more often than not also not able to access water at a reasonable distance from where they live.

The GHS provides useful information on the extent to which South African households currently *pay* for water services, as well as indications of water pricing. On average, only 47.8% of South African households that had access to water paid for it in 2005. In three provinces, more than 60% of households who had access did in fact pay for water, namely Gauteng (65.3%), Northern Cape (63.7%) and the Western Cape (73.9%). With the exception of Limpopo (20%), in all the other provinces the percentage of paying households ranged between 30% and 50%.

It is difficult, however, to draw conclusions from these statistics without further information on the reasons for payment or non-payment. One would, *a priori*, expect such reasons to include the lack of a functioning metering and billing system, the existence of a free water services policy, and unhappiness over service delivery quality on the part of households leading to non-payment. Table 5.2 is taken from the general household survey (GHS 2005: xxxii) and gives the reasons between which households could select for non-payment. Although the values provided are percentages, they do not add up to 100%. The 2004 values add up to 103% and the 2005 values add up to 139.2%. It is not entirely clear why this is the case. Thus the results should be interpreted as broadly indicative rather than precise.

**Table 5.2: Reasons for Non-Payment for Water Services**

	<b>2004</b>	<b>2005</b>
1) No metering / billing system	40.3	56.3
2) Other reason	21.0	18.9
3) Government should supply free	11.6	12.2
4) Others get free water	10.4	12.2
5) High cost of water	6.7	19.6
6) Free basic amount only	5.9	7.2
7) Service poor	3.4	6.4
8) Irregular metering/billing	2.6	5.4
9) Meter broken	1.1	1.0
Total	103.0	139.2

For analytical purposes some of these reasons can be grouped together. Thus, reasons 1), 8) and 9) represent non-payment related to the fact that households can't currently be billed as a result of an inadequate metering and/or billing system. 62.7% of households who have access to water but don't pay for it do so for such reasons. No conclusions can be drawn from this about the percentage of households who would in fact refuse to pay *if* a metering or billing system were installed. It is furthermore interesting to note the large change from 2004 to 2005 in the percentage of households who indicated a lack of a metering or billing



system as the reason for non-payment. Unfortunately, without further investigation which falls outside the scope of this study, one can only speculate as to the reasons. It may be that many of the households who were connected in this period were connected without a metering or billing system being simultaneously installed. It may also be the result of destruction of metering systems, where a pre-paid system is used.

Reasons 3), 4) and 5) relate clearly to the matter of affordability of the service and the related perception that water should be provided free, i.e. that others should pay. In fact, there is little reason for distinguishing between respondents who cited that government should supply water free and those who cited that others get water free. In both cases the position is grounded in the view that water services are to be provided without a direct charge. 24.4% of households with access who did not pay cited this reason. If we add the 19.6 % of households who, though not necessarily demanding 'free' water, believed that the cost of water was too high, then 44% of non-paying households did not pay as a result of concerns about the cost at which it was provided to them. In other words, 23% (44% of 52.2%) of all households in South Africa have access to water services but don't pay for it because they regard it as too expensive. This is undoubtedly a high percentage.<sup>67</sup> The high percentage of non-performing municipal debt for services, referred to in a previous chapter, underscores the notion that non-payment is quite prevalent in many municipalities.

It is, however, quite difficult to rigorously answer what appear to be two simple questions: How much do South Africans pay for water? And: is water affordable for an acceptable percentage of households? Fankhauser and Tepic (2005: 1) express the situation well: "Both the development of social assistance

---

<sup>67</sup> There is of course another perspective on the high rates of non-payment for municipal services, which relates it to the culture of non-payment entrenched in the apartheid era. This hypothesis is often evoked and by its very nature is difficult to confirm or repudiate. It is not engaged with further in this study.

programmes and the policy debate on tariff reform is held back, however, by a dearth of good information about household expenditure on infrastructure services. It is surprising how little is known about their affordability, given the political sensitivity of the issue and its prominence in the policy debate.” In an explicitly equity-orientated engagement, focusing on free basic water policies and the rights of children, Mosdell and Leatt (2005 :4) concur with reference to the South African context: “There is remarkably little evidence about the cost implications for households since introducing the dual policy of payment for basic services, with a free basic component. We do not know what the impact of this has been on household expenditure across South Africa.”

In the absence of good quantitative data, debates in South Africa have at times been polarised too simplistically into camps objecting to almost any form of cost-recovery tariff-setting and those in favour of it in all instances. The problem is not only informational, however, but also conceptual. Two approaches, neither without problems, can be employed in trying to assess affordability, namely the ‘behavioural change’ approach and the ‘affordability ratio’ approach.

The behavioural change approach posits that a service is unaffordable if it leads to rational behaviour change (i.e. the maximization of household welfare for a given price and income environment) which is likely to have significant socially undesirable effects. Unaffordability has negative externalities, in other words. Expressed less technically, it means that a service is unaffordable if it drives a household into desperate, dangerous or socially inequitable alternatives. If the price of water is so high to newly connected poor households that they start collecting water from a nearby canal which passes through an industrial area before it reaches them, then one can state with some certainty that water is unaffordable.

Clearly, in the case of a resource essential to survival and well-being such as water, the income elasticity of demand, at least for low consumption levels

associated with survival, must be fairly small. Alternative, free or cheaper sources of supply are rarely available. However, survey results would not easily pick up changes in behaviour, such as a connected household resorting to public tap use because its supply has been cut off due to non-payment. The alternative appears to be to agree on a benchmark figure or 'affordability ratio' as termed above, for the percentage of household expenditure which should go to water and other infrastructure services expenditure. The discussion below develops this issue further and evaluates what can be concluded from some of the available official statistics.

The *2005 General Household Survey* includes a table which serves as a useful starting point for trying to assess affordability not from the perspective of the percentage of non-payment or behavioural change, but through trying to determine how much households typically pay for water in different circumstances. The table is headed 'Provincial distribution of monthly payment for water, 2004' and shows the percentage of households in each province who paid an amount within a certain range. The table is duplicated below as Table 5.3.

**Table 5.3: Provincial Percentage Distribution of Monthly Payments for Water, 2004**

Province	WC	EC	NC	FS	KZN	NW	GP	MP	LP
R10 or less	6.2	5.9	3.2	2.2	3.6	12.2	1.3	5.1	14.8
R11 – R20	13.3	10	6.5	8.3	7.5	11.8	3.5	8.4	9.5
R21 – R50	21.6	22.7	18.1	22.6	13.7	15.6	10.1	20.7	15.2
R51 – R100	17	21.3	18.1	16.1	20.7	13.6	11	16	8.4
R101 – R200	12.3	17.7	19.4	12.8	17.8	12.6	12.2	12.4	12.2
R 201 +	8	10.4	12.9	15.2	16.4	19.6	22.1	10.9	19
Fixed, unknown cost included in rent	17.6	10.2	21.3	21.3	19.5	11.8	37.6	21.1	20.2
Don't know	3.9	1.8	0.6	1.7	0.8	2.6	2.2	5.5	0.4
Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

Source: GHS (2005: xxxi)

Unfortunately, from an analytical perspective this information is not very useful, as there is no way to determine whether differences in monthly amounts paid are the result of differences in the unit *price* of water between provinces or in the volume of water consumed by households. The fact that in some provinces (for example Gauteng) very large percentages of households do not know how much they actually pay further renders the table's information less than ideal for purposes of comparison.

The *Non-Fin 2005*, also referred to in a previous chapter, contains a fair amount of supply-side data for water services by province. It indicates, by district municipality, the kilo-liters (kl's) supplied to the district by, respectively, a water board, own sources and other water service providers. It also indicates how many consumer units at different standards (inside yard, less than 200 m from yard and more than 200 m from yard) are being served. Finally, cost information is provided as totals and as costs per unit bought and sold.

Table 5.4 provides information for three Western Cape districts simply to illustrate how the information is presented.

**Table 5.4: Selected Water Supply Data for Three Western Cape Municipalities**

Municipality	Water Board	Other Provider	Own Sources	Supplied in Yard	Supplied Less than 200m from yard	Supplied more than 200m from yard	Ave cost per unit bought	Units lost in distribution	Ave cost per unit sold
Boland	8 213 840	15 699 366	25 077 940	104 256	5 414	500	0.88	8 824 503	2.64
Overberg	82 155	0	67 130 183	41 982	4 769	0	1.22	55 313 704	2.33
City of Cape Town	0	0	283 409 907	730 000	170 000	0	0	58 419 550	4.81

Again, however, there are numerous problems with trying to draw any conclusions from these tables on the costs of water to households. Firstly, the average cost per unit bought appears to be merely an accounting cost. In the case of the City of Cape Town, for example, the unit cost is given as zero. This is because Cape Town's water is supplied from its own water sources. It does not, of course, mean that Cape Town water provision has no opportunity costs, merely that it does not have to buy its water from others. Thus, further, in assessing the unit cost in the case of other districts, there is no way of telling what proportion of total water bought has been bought at the average unit price given. This means that the average unit cost information provided again has only limited applicability to *some* of the water it purchased.

**Table 5.5: Weighted Average Annual Cost of Expenditure on Water Services (2000 prices):<sup>68</sup>**

	<b>Africans</b>	<b>Coloureds</b>	<b>Indians/ Asians</b>	<b>Whites</b>
Water (a)	284.86	510.08	1 087.51	1 253.03
Average Annual Housing-Related Expenditure (b)	7 575.36	13 546.47	31 700.89	57166.36
Average Annual Household Expenditure on all Items (c)	22 558.00	45 146.00	72 006.00	136 141.00
Water as % of Housing Expenditure (d)	3.8%	3.8%	3.4%	2.2%
Water as % of Total Expenditure (e)	1.3%	1.3%	1.5%	0.9%

Finally, the 2000 *Income and Expenditure Survey* (IES) also provides information on average expenditure on water by households, broken down by population group. These results are provided in table 5.5 and allow some determination of the percentage of household income going to water by population group. Given the continued correlation of poverty with population group, the approach taken here is to broadly associate households with poverty according to which population group they fall into.

These results do suggest that payment for water services tends to be regressive, in the sense that poorer households spend proportionately more on water than richer households. The IES data does not allow distinction between what is usually called ‘affordability of access’ and ‘affordability of consumption’, since it provides a single figure per household. This masks what are likely to be significant differences between population groups in this regard, given the apartheid-era bias towards prioritising the provision of infrastructure in white residential areas. Black population groups, which are closely associated with lower household expenditure in the above table, are of course also more likely to

<sup>68</sup> The 2000 IES precedes the introduction of free basic services. It would be interesting to try to assess the impact of this on a table such as the one above.

live in areas characterised by infrastructural backlogs. In other words, monthly or annual household payments of this group for water would in all likelihood include some form of full or partial recovery of capital costs, whereas formerly white residential areas would tend to require only infrastructure maintenance costs to be included in the bill.

A further caveat to any simple interpretation of the above results is that the percentage of households having access to the water service differs significantly between population groups, with the black African group being the location of the overwhelming majority of households who enjoy *no* acceptable form of access to water. The average expenditure of R 284.86 for the black African group masks the fact that this group probably has a far lower median value, since a sizeable proportion of group members pay nothing for water as they do not have access, whereas in the white household group average and median values are in all probability closer. This implies that those African households who are connected pay substantially more than the average for all African households.

A final comment that needs to be made here rests on the distinction between actual household service *payments* and the amounts they were billed for. The IES data is a reasonable approximation of actual expenditure: this suggests that *billed* amounts were appreciably higher in many cases, that is it may be the case that services are, on average, affordable when expenditure data is used for analysis because rates regarded as unaffordable by households are simply not paid! Fankhauser and Tepic (2005: 13) find a similar possibility in their review of Eastern European utility services and affordability: "(O)ne of the reasons why affordability ratios are often acceptable under current conditions is the high level of non-payment." The authors seem to suggest, in other words, that non-payment acts as a self-correcting tendency when pricing is socially unacceptable. Needless to say, however, such tendencies erode the integrity and perceived legitimacy of the service system, as well as creating further inequities between households who do and households who don't pay. This argument also implies,

however, that socially problematic outcomes result from circumstances where pricing is 'unaffordable' and non-payment can't act as a correcting factor, that is to say especially where pre-paid meters and the like are installed.

Though the above percentages and affordability ratios should be treated with some caution, they do enable some comparison with other countries and existing attempts to come up with an acceptable benchmark range for what percentage of household expenditure should go to water services. Such a determination is necessarily a value judgement. Fankhauser and Tepic, for example, candidly note that a rigorous determination does not structure current debates: "(M)any governments and international financial institutions have developed *ad hoc* rules on what constitutes an acceptable level of utility expenditures...Although there is no universal benchmark...an acceptable threshold may be around 25 per cent of household expenditures for electricity, heating and water" (Fankhauser, S & Tepic, S 2005: 5). Derived from this, they articulate an indicative benchmark for water and waste water as 5 percent of household expenditures. This figure is, encouragingly, identical to that used by Chile as a proxy to assess affordability of water services for poor households, according to Gómez-Lobo (2001). The South African figures as taken from the IES, whilst regressive, compare quite well with these benchmarks.

## **e. Conclusion**

The discussion of this section has illustrated the inherent conceptual difficulties of assessing affordability in considering the equity implications of pricing. It appears that, on average, water is available at a reasonable price to households, but it does little more than that. Further, significant percentages of the population cannot access basic services under current circumstances. The next chapter evaluates the main current measures to ensure greater equitability in local service access and affordability.



## **Chapter 6**

# **Equity in Infrastructure Services: Current Policy and Financing**

### **a. Introduction**

In the previous chapter some results were provided for the affordability of local services in South Africa at present, with special reference to the provision of water. This chapter goes into more detail with reference to the current policy-environment as far as enhancing equity is concerned. The aim is both descriptive and evaluative. Where required, questions will be asked regarding the extent to which the current system does represent the best possible means of achieving a sustainable and Constitutionally legitimate balance between efficiency and equity.

Policies aiming at greater equity in access to household services and affordability across income groups can fundamentally try to target households directly or alternatively to support the ability of municipalities to implement equity-orientated policies. It is taken as axiomatic in this regard that an equity-orientated policy is funded through other means than charging beneficiary households a cost-recovery or related amount. It follows that other households, selected on the basis of an appropriate criterion such as income, will have to make up this deficit. Such households within the service jurisdiction can be charged proportionately more for consumption of the same service which is provided to beneficiary households at discounted rates. Surpluses generated on other service accounts can also be used.<sup>69</sup> A final option is to tax some households at a rate (constrained in most instances by national taxation) which generates a surplus corresponding to the 'equity deficit'.

---

<sup>69</sup> The Municipal Finance Management Act (MFMA) requires municipalities to run a balanced operational budget, but it does not require balanced budgets for particular service accounts.

In addition to municipal options, there is the likelihood of receiving funds from the national revenue fund, which is of course generated through a range of taxes on all residents of the country.<sup>70</sup> Table 6.1 presents this basic matrix of options for the sake of clarity.

**Table 6.1: Matrix of Local Revenue Possibilities**

<b>Financing Burden</b>	<b>Revenue Instrument</b>	<b>Public Finance Objective</b>	<b>Redistributive Incidence<sup>71</sup></b>
Beneficiary Household	User Charge	Greater Efficiency	No
Other households in Service Jurisdiction	Property Tax	Greater Equitability	Yes
Other households in Service Jurisdiction	Same or Other Municipal Services	Greater Equitability	Yes
Non-Residents	All national tax instruments; revenue transferred to municipality with or without conditions	Greater Equitability	Yes

The previous chapter indicated some concerns and trends with regards to the affordability of household services in South Africa as well as the fairly high level of service backlogs in some provinces and areas within those provinces. In this chapter the emphasis falls on exploring the conceptualisation and the funding of the primary policy intervention to enhance equity in household service access and affordability, namely the policy of free basic services. The discussion

<sup>70</sup> Some of which in turn do fall on municipal residents, both those targeted for equity measures and those not. This effect should, in principle, be allowed for in analysis, though in most cases its impact would be slight. Where, however, a consumption tax such as the VAT is levied at a very high rate and transfers are used for equity objectives it may be worth asking whether at least some of these equity objectives could not be achieved more efficiently (that is with less distortions) by further adjustments to the VAT rate.

<sup>71</sup> In the statutory sense: economic incidence would have to be determined on a case by case basis.

proceeds as follows. Section b describes and reviews the basic service debate and the free basic services policy. Information on the current state of implementation of the policy, and associated challenges, are also dealt with. Section c then turns to national transfers as a means of financing free basic services and discusses the local equitable share and the municipal infrastructure grant. These two transfers are the most significant means available in South African to address the vertical fiscal imbalances of the system, that is the financial disjuncture between revenue powers and expenditure functions assigned to the various levels of government. Section d looks at options available to a municipality from its own revenue sources for funding free basic services, namely property taxation and price intervention in the form of service cross-subsidisation.

## **b. The Policy of Free Basic Services**

Though the nature and limits of the obligations imposed by the Constitution on government with regards to basic services has been the subject of much debate, there seems to be little doubt that access to affordable basic services should be regarded as a core component of the fulfillment of socio-economic rights. This does not necessarily mean that government is required to provide or finance these services, but that government becomes responsible in those instances where it is clear that the rights of households and individuals in this regard will not be fulfilled. The Constitution makes of Government the 'default' provider, as it were. Moreover, in an economy characterised by systemic and chronic poverty and inequality, as well as a range of social ills such as high rates of crime and HIV/AIDS prevalence and incidence, there can be little doubt that government will have to take command over or redirect a substantial part of the economy's resources if the Constitutional vision is to be attained, though such 'command' is of course constrained by the potential for government failure and considerations of macroeconomic stability and the derived constraints on the national budget.

The adoption of a free basic services policy in 2001 by the South African Government (and more broadly of significant subsidisation of basic services where households pay a minimal or primarily token fee) needs to be considered in this context because it stems from the recognition that some households will be unable to purchase basic levels of those services indispensable to wellbeing even at a heavily subsidised rate. It therefore requires that these households be provided with a minimum service amount and of a minimum standard consistent with the Constitutional obligation imposed on government. This free amount has generally been identified with the 'basic' amount necessary. Of course, the designation 'free' does not mean that no costs are incurred in providing the service, merely that the beneficiaries are not required to bear those costs. Thus, financing such provision will necessarily be redistributionary to some extent; redistribution may be either internal (i.e. through the municipal budget) or external (i.e. through the national budget).<sup>72</sup> A key difference between the two options is the nature of the constraints against redistribution operating in the different spheres. It may be the case that there is less 'tolerance' for internal redistribution than there is for external redistribution.

Clearly discussion of what should constitute a basic level of a service for municipalities to aim at is likely to consist of a mix of technical and political arguments, of tensions between conceptions of the minimum household requirements for dignity on the one hand and hard resource constraints imposed by a national government which remains responsible for overall macroeconomic stability on the other hand. In South Africa inherited high levels of income and wealth inequality add further tension to the debate, especially where trade-offs need to be considered between extending water and energy services infrastructure to poor and possibly rural off-grid households and large-scale industrial expansion which is likely to have a more long-term, trickle-down effect

---

<sup>72</sup> Here, as elsewhere, arguments are presented for a two-level system; in the South African context this implies little engagement with the provincial sphere. It is assumed that this does not materially affect the analysis.

on welfare. Though the tension between ‘social goods’ and ‘economic goods’ need not be inherent, much of the critique of South African theorists who may be described as ‘leftist’ in the political economic spectrum has focused on such tensions and pointed to a perceived over-prioritisation of economic goods. This kind of question also implicitly engages with the vertical division of revenue,<sup>73</sup> since different levels of government are typically responsible to differing degrees for social and economic goods.

The essential point is that different understandings of what is meant by ‘basic’ will have different implications for funding. If a basic water service is conceived as a public tap within 200 m of a household, this has different present and future claims on the municipal budget than if basic means that all households should have a tap in their own yard. The financial requirements change not only because it costs more to install an in-yard tap on every property than communal taps in a neighbourhood, but also because a more ‘luxurious’ definition of basic will mean less households are regarded as having access, and thus will mean a higher backlog rate in the municipality.

The Department of Provincial and Local Government (DPLG)’s *The Municipal Infrastructure Grant: Basic Level of Services and Unit Costs: A guide for Municipalities* clearly expresses the fundamental financial implications of service level determination: “Municipalities depend largely on the income received from customers and this must be sufficient to cover the cost of providing the service. Higher levels of service are generally associated with higher costs, for which customers must pay more. If higher levels of service are not affordable, the ability of a municipality to recover its costs is negatively affected, threatening the viability of the municipality” (DPLG 2005: 3).

---

<sup>73</sup> That is the relative share of nationally raised revenue going to the national, provincial and local spheres of government. Such a division precedes the ‘horizontal’ division of the total allocation to a sphere amongst provinces and municipalities.

In a recent investigation of local government finance in South Africa (Reschovsky in Bahl & Smoke 2003: 197-198) it is argued that there is no clear consensus about which services are to be considered basic, but that strong arguments can be made for including the following:

- Potable Water
- Sanitation
- Solid Waste Removal
- Electricity
- Roads
- Municipal (primary) health
- Stormwater management
- Firefighting and emergency services
- Street lighting

The approach to basic services followed by these authors essentially identifies basic services as services characterised by limited spill-over effects; that is to say, following the fiscal decentralisation benchmark discussion of chapter 2, basic services are here regarded as services whose costs (or recurrent costs at any rate) and benefits are largely limited to a jurisdiction designated on spatial grounds as 'local' in the South African system of fiscal governance. The 'roads' included in the above list in all likelihood refer to roads in residential areas rather than through roads.

However, their approach to basic services is not limited to the type of service, as listed above, but also requires consideration of an appropriate level of access to the service. In fact, they articulate the key process very clearly, emphasising that the notion of basic services must comprise both a process of delineating the *types* of services included *and* specifying minimum access levels for those services which are included: "Once we have identified a list of public services, such as water and sanitation, that may be considered 'basic', the next step is to define in detail what level of service is to be defined as basic and then decide

whether the basic level of service should be different in different kinds of communities” (Reschovsky in Bahl & Smoke 2003: 198).

In conducting such a process, care should also be taken to avoid thinking of the designation ‘basic’ in absolute terms, for example thinking that ‘basic nutrition’ should be taken to denote some ahistorical minimum caloric requirement necessary for survival. It is more appropriate to regard the term as an attempt to designate a *minimum socially acceptable* or *constitutionally mandated* level of a service. Reschovsky develops this idea and argues that what is regarded as basic service levels will consequently increase positively with income over time. The meaning of basic is relative to where a society finds itself at a particular point along its developmental trajectory. Similarly, though almost certainly also more contentiously, he further argues that the content of the term ‘basic’ may legitimately vary with location, and specifically that ‘basic local services’ may mean different things in urban and rural locations (Reschovsky in Bahl & Smoke, 2003: 198, 199). Given the often far greater service area and virtually non-existent infrastructure in many rural areas, the costs of providing basic service levels via the modes associated with urban service provision are likely to be prohibitive, especially if such costs are considered in per capita terms, given often low population density in these areas.<sup>74</sup>

Basic local services, then, are understood as *socially or constitutionally determined minimum levels of public services generally associated on allocative efficiency grounds with local jurisdictions*. Policy documents are clear about the need to *prioritise* operational and capital expenditure aimed at ensuring a basic

---

<sup>74</sup> As a further example, in addition to the temporal and locational ones, of the relativism inherent in the meaning of ‘basic’ services, the Children’s Institute study on FBS argue that basic services will vary with the *health* status of beneficiaries. They cite a study which argued that HIV positive individuals require more water. Thus, if the aim is in fact to provide basic services consonant with ensuring a basic level of wellbeing various relativising factors are potentially significant. In practice, of course, information-gathering and analysing costs to effect such a nuanced approach would be huge. Conceptually, however, the point is relevant.

service level to all households. The basic amount should be provided 'free' to so-called indigent households to ensure that they also have access. Thus the designation 'basic' has, in public policy debates, increasingly become inherently associated with the free (in the sense of a transfer-funded service financed through national tax revenues) provision of basic levels. The Department of Water Affairs and Forestry (DWAFF) norms and standards, which emanate from the Water Services Act, require that water service institutions take cognizance of the following in setting tariffs (DWAFF 2001: 2, 3):

- (a) recover the cost of water purchases;
- (b) recover overhead, operational and maintenance costs;
- (c) recover the cost of capital not financed through any grant, subsidy or donation;
- (d) provide for the replacement, refurbishment and extension of water services works; and
- (e) ensure that all households have access to basic water supply and basic sanitation.

Clearly (a)-(d) concern the need to recover costs, and the need to plan for future costs. These requirements, however, stand in some sense opposed to that of (e) which requires the institution to ensure all households have access to basic water and sanitation services. Thus, the prima facie solution, on paper at any rate, would clearly be to subsidise services to targeted households, and indeed the next section explicitly sanctions this:

A water services institution<sup>75</sup> may use any source of funds, including any funds received from municipal rates and taxes or from transfers from national or provincial government or from any other source, to subsidize a water services tariff. A water services institution must consider the right of access to basic water supply and the right of access to basic sanitation when determining which water services tariffs are to be subsidized. (DWAFF 2001: 3.)

---

<sup>75</sup> 'Water service institution' includes municipalities.



As the *Guide* makes clear, any of these options may be used by a municipality to ensure the provision of basic amounts of the service in question. In seeking to realize their basic service aims, in other words, South African municipalities are not required to balance their budgets by setting own expenditure equal to own income, but also have access to both the LES and conditional transfers.<sup>76</sup>

One of the most unfortunate aspects of the current debate around FBS is the use of the term 'indigent' and the manner in which it has been conceived. As part of their FBS policy implementation municipalities are encouraged to develop an 'indigent register', which appears to be conceived as a record of households who, at some point in time, met some form of criteria for 'indigency' which qualifies them for free basic services. There are a number of problems with this notion. The first problem is that it enhances the tendency towards stigmatization of certain groups: the image that an 'indigency register' evokes is that of a group of hopeless, helpless chronically poor households who are likely to remain permanently in some kind of economic ghetto. The second problem is the potential lack of a dynamic aspect to how indigency is conceived and measured, as well as how and when a person migrates out of 'indigency'. The third problem is of course the virtually insurmountable administrative and informational hurdles which would have to be surmounted to compile and update such a 'register'. The recommendations in the DPLG's (2003: xii) *Study into the Provision of Free Basic Services* include:

- The development of a national indigent register
- An agreed definition of what is an indigent
- The framing of an indigent policy

---

<sup>76</sup> Thus the question of appropriate service level becomes one also of the relative priority assigned to local services by national government, since it is at this level that, ultimately, division of revenue decisions are made.

Thus, at this point, as these recommendations illustrate, no clear policy or even definition of indigency exists, and yet municipalities are tasked with developing their own registers and deriving a more targeted FBS policy from it! A further concern with the 'indigency' approach is that in practice it risks excluding the most vulnerable and excluded from economic opportunity and social life. One can for example imagine a municipality announcing that all households who regard themselves as indigent apply for registration as such in order to qualify for FBS. Clearly a systematic bias may well result in favour of the better off or the 'better off poor'; resulting in benefitting say the 2<sup>nd</sup> and 3<sup>rd</sup> income decile rather than the first.

The DPLG study is an attempt to address four perceived shortcomings in relation to the state of knowledge on free basic services. These shortcomings are listed in the executive summary as follows (DPLG 2003: i):<sup>77</sup>

- Statistics available did not provide a detailed picture of delivery at municipal level
- Lack of information on the challenges faced by municipalities in delivering FBS
- Unknown impact of FBS on poverty
- Lack of information on which municipalities were experiencing difficulties and would require support in rolling out FBS

An initial question on the financing of free basic services asked municipalities whether they felt they had sufficient finances to support their planned FBS initiatives. 50% felt that they did not have sufficient resources and 41% that they did (9% did not answer). Such a result may be interesting but it is far from clear what, if any, conclusions can be drawn from it. Firstly, one can assume that municipality's would tend to underestimate their own revenue potential in answering such a question. They may do so because of existing informational constraints, as well as because they would prefer to continue to receive large

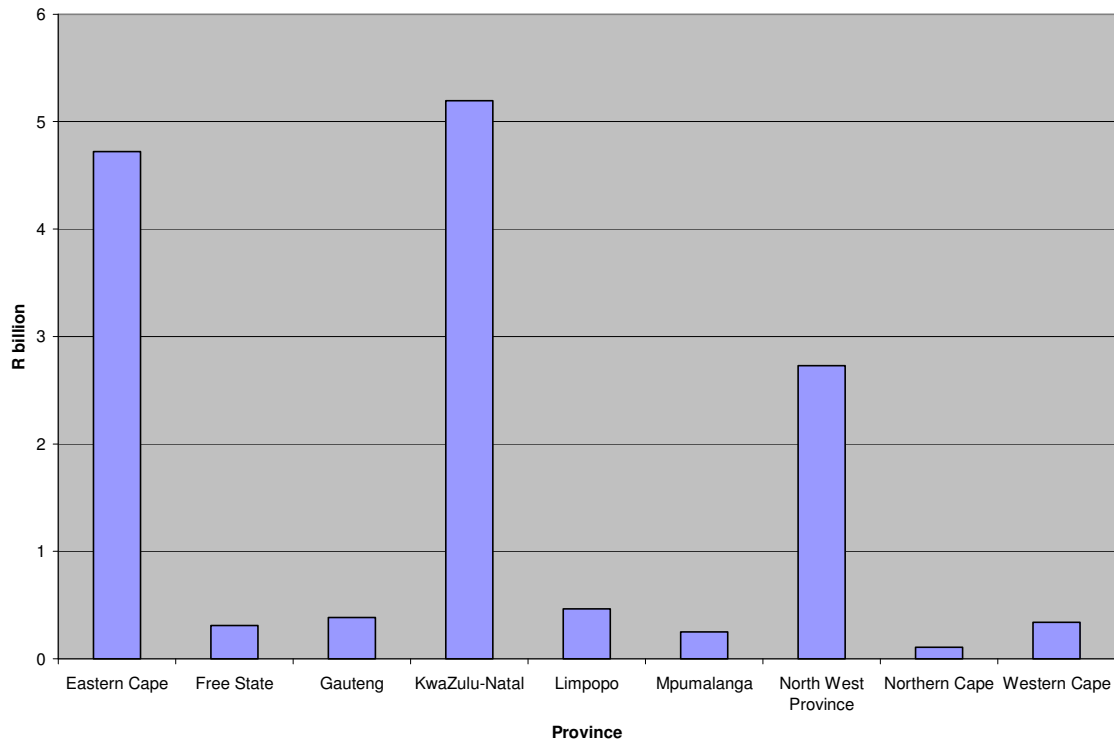
---

<sup>77</sup> Slight rephrasing has been effected in some cases.

transfers from national government rather than impose more taxes and rates on citizens whose vote they require in order to retain political power. Further, the question is phrased in such general terms that it allows no objective assessment of matters. Do 'planned FBS initiatives' represent everything a municipality would like to do in a perfect world? Or is it a pragmatic assessment of what is possible? And if the latter, then surely it should be at least broadly aligned with available resources already? There are also simply too many intermediate positions between these two extremes for these responses to be very revealing. The difficulty of interpreting the content of the questions is underscored by a subsequent question, which concerns the sustainability of FBS initiatives: 71% of municipalities felt that they would be in position to sustain their FBS initiatives. This result implies that a sizeable percentage of municipalities felt their FBS initiatives were sustainable whilst simultaneously stating that they lacked sufficient finances!

An interesting question included in the report requires municipalities to estimate the cost of backlog eradication (DPLG 2003: 19). Results, in rands, are presented by province. The question refers to 'providing the infrastructure required for FBS provision'. There is no conceptually useful distinction between infrastructure for FBS and 'other' infrastructure: connecting a household to the electricity grid or water network allows both the free and the cost-recovery provision of the service in question. Thus, the only coherent meaning which can be assigned to this question is that it simply refers to the cost of eradicating infrastructure backlogs for water, electricity, sanitation and refuse removal services. Adding up municipal responses of estimated costs in this regard provides a figure of R 14.5 billion in 2003 rands for eradication of backlogs. Figure 6.1 provides a breakdown, by province, of this amount.

**Figure 6.1: Municipal Estimates of Infrastructure Backlog Costs for FBS**



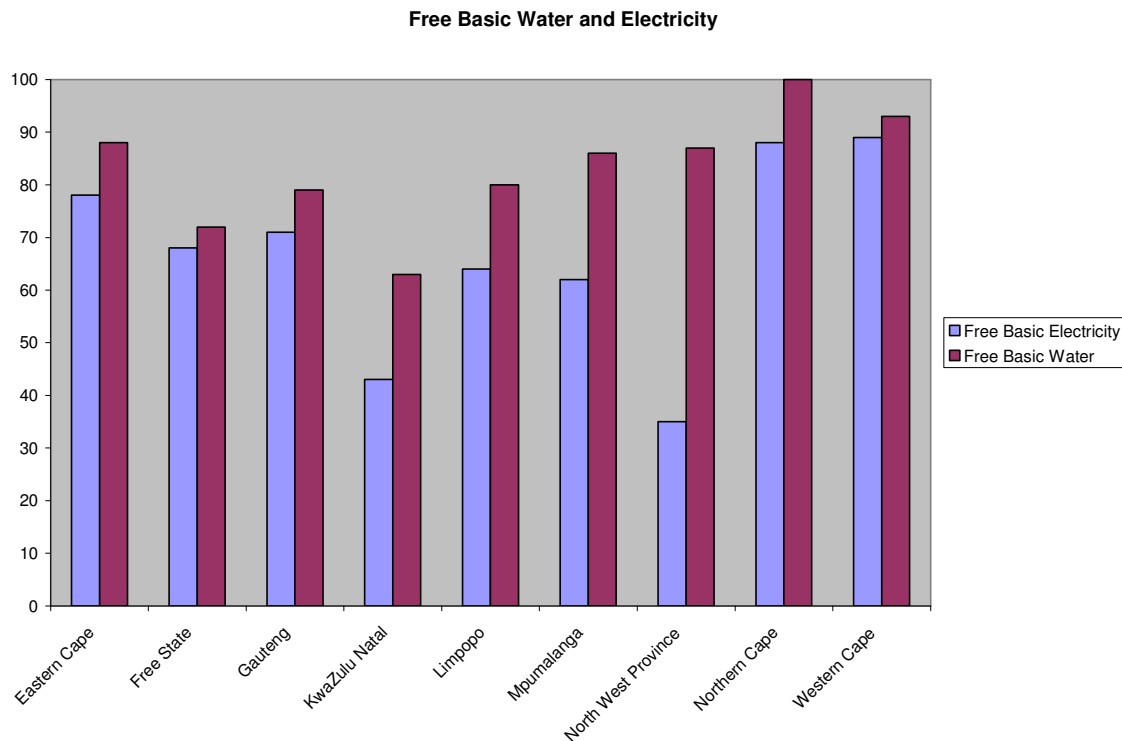
Source: DPLG (2003: 19); own calculations.

It is again quite hard to know what these indicator values represent. The percentage scores, captured visually above, appear to bear little relation to composite measures of infrastructure backlogs. Admittedly, one would not expect backlog percentages to correlate directly with estimated costs across the provinces, since other factors are involved. Specifically, costs of eradication may differ across provinces (for example according to whether backlogs are predominantly urban or rural) and may also differ according to the *level of service* aimed at in eradicating the backlog. There is no indication here whether any effort was made to ensure that municipalities were aiming at the same standard of service in estimating cost of backlog eradication for purposes of the study. A final concern with this information is that it may at times only reflect past inefficiencies: after all, a municipality's prime source of information on expected infrastructural costs would be its own past performance in this regard. If unit

costs for installing ventilated improved pit latrines (VIP's) were x, and in fact this is 15% more than the national average of all municipalities (and assuming there are no unique conditions which explain this), it is nevertheless highly unlikely that the municipality, even if it were aware of its own inefficiency, would provide a cost estimate using 85% of unit costs it has incurred to date.

A series of questions are posed with reference to the separate components of free basic services, that is free basic water, electricity, sanitation and refuse removal, including the percentage of municipalities in each province providing these services free. Results for water and electricity are provided in figure 6.2. Unfortunately, municipalities were not required to indicate on what basis they provided these free services, that is whether a basic 'lifeline' amount was provided free to all households enjoying access, or whether a targeting mechanism had been developed and was being used. Some tentative conclusions can nevertheless be drawn and are discussed below.

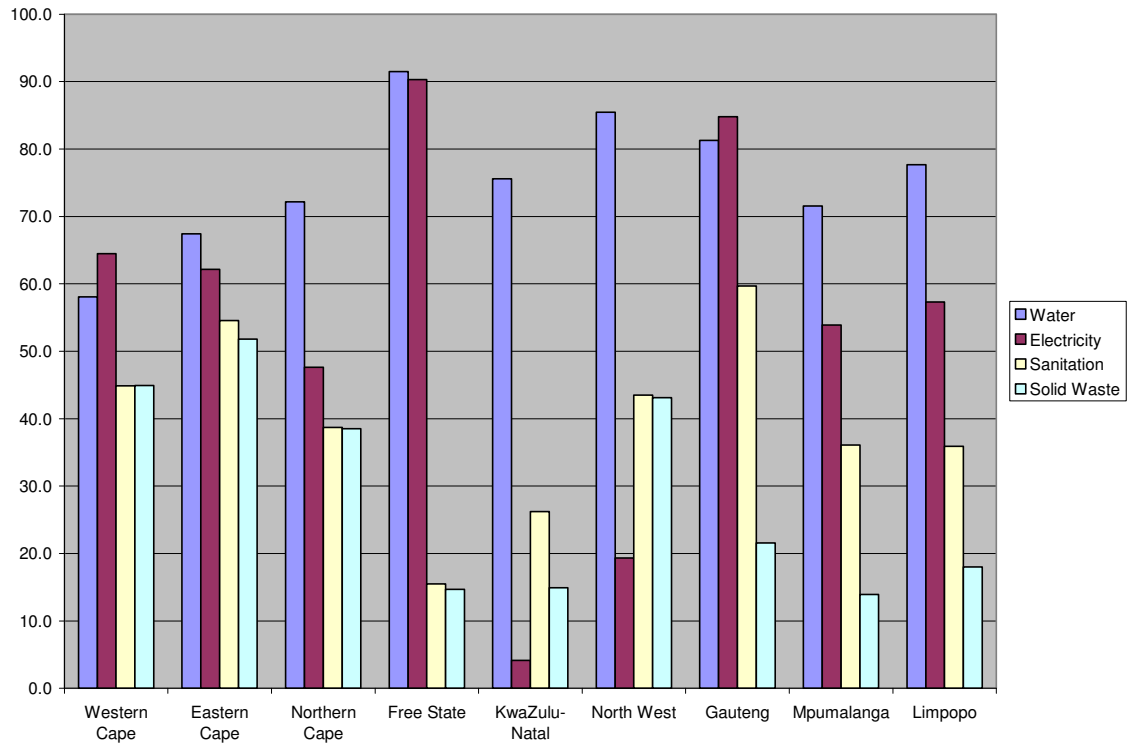
**Figure 6.2: Percentage of Municipalities in each Province that Provide Free Basic Water and Electricity**



Source: DPLG (2003: 21); own calculations.

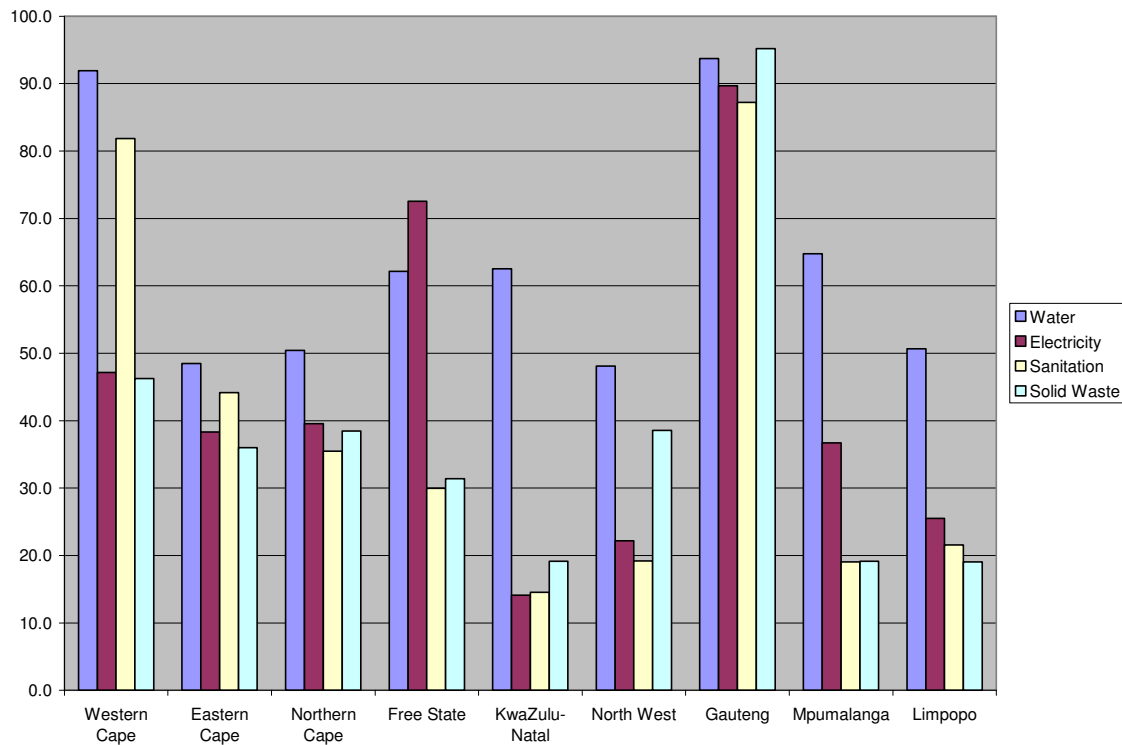
These results do not give an indication of how many households are benefiting, though this is fairly easily obtainable from the non-financial census of municipalities, where information is provided for households receiving selected services from municipalities, as well as households receiving such services *free*. Figures 6.3 and 6.4 provide an indication of trends in the percentage of households receiving municipal services who received some of them free, for 2003 and 2005, simply by dividing the number of households who are recorded as receiving these services by the number who are recorded as receiving *free* services.

**Figure 6.3: Free Basic Services: Percentage of Households Receiving in 2003**



Source: DPLG (2003: 21); own calculations.

**Figure 6.4: Free Basic Services: Percentage of Households Receiving in 2005**



Source: DPLG (2003: 21); own calculations.

It is interesting to note the extent to which some services appear to be provided with a free component, but not others, in some provinces. In the case of Gauteng, for example, it appears that solid waste management acquired a substantially higher free service component in the period between 2003 and 2005. It is also interesting to note that a substantial portion of households in KwaZulu Natal received a free amount of electricity, but much smaller percentages of households were the beneficiaries of a free amount of other basic services. In a number of provinces the percentage of households receiving a free amount declined between 2003 and 2005. In Limpopo, for example, 77.7% of households received free basic water in 2003, but only 50.7% did so in 2005. In the North West, 85.5% received free water in 2003, but only 48.1% did so in



2005. In the Western Cape, on the other hand, free basic water went from being received by 58.0% of households in 2003 to 91.9% in 2005.

These trends suggest changes in policy *towards* the broader application of FBS policies in the Western Cape, as well as, in the case of Limpopo and the North West, the introduction of improved, that is more specific, targeting mechanisms or the introduction of more broadly applied cost recovery.

In general the *Non-financial Census* and DPLG results suggest that there is not a great deal of targeted basic services provision occurring, and that in many cases municipalities prefer to meet their free basic service obligation by offering a lifeline amount of the service free to all households. This conclusion is suggested by the fact that free basic service percentages correlate quite closely with *access* percentages in most of the provinces.

Issues pertaining to financing FBS and those pertaining to targeting decisions are two sides of the same coin, in the sense that decisions made regarding targeting, in conjunction with the *amount* of support<sup>78</sup> given to identified households, will largely determine the extent of financing required for the FBS policy. In a sense the targeting issue even precedes the 'amount of support' issue, since the number of households a municipality needs to cater for will, for a given resource envelope, determine the support amount.

The basic targeting decision which needs to be made for any form of income transfer or aid-in-kind is between universal or selective support, where universal support simply means that all households falling within the jurisdiction for which the support is defined receive it. Selective targeting entails identifying households who qualify for the support on the basis of certain criteria. Firstly it needs to

---

<sup>78</sup> The presumption tends to be that the amount is fixed for all eligible households, rather than working on a scale according to income, as in a progressive income tax system. The presumption in favour of a flat amount is retained here.

choose between a universal and a targeted approach, as mentioned above. As also mentioned there, it should opt for a targeted approach only if the savings associated with reduced leakage exceed the additional costs of gathering information and administering a targeted system. Where doubts exist, the universal option should be retained, since the poor do benefit. It is certainly the case that they would benefit even more under a performing targeted system, but it is suggested here that they are likely to benefit *less* under a non-performing one. A feasible variation of the universal approach, which nevertheless includes a redistributionary aspect, would be to recoup the universal grant amount through the property tax. In a sense this would automatically entail a form of targeting, since owners of taxable property tend to have more assets and income than those who do not. If it was regarded as necessary to sharpen the targeting further then a property tax 'surcharge' could be levied on properties above a certain value. One suspects that such 'innovative' approaches may firstly meet with a fair amount of resistance amongst voters (the property tax is already often a very unpopular tax). Secondly, at present it appears as though many municipalities do not possess accurate and updated property value registers. This administrative hurdle would have to be addressed before such measures could be used effectively.

In the case of selective targeting, then, various questions will have to be addressed which a universal subsidy need not grapple with, such as the appropriate criteria and threshold values for selecting households. The universal rather than targeted approach to free or heavily subsidized provision is justified where the costs of universal provision of a free amount are less than the bureaucratic costs of a more selective system of targeting households. 'Targeting Poor Households in the Provision of Basic Municipal Services: A Guideline for Municipalities' of the DPLG is a very coherent explication of the DPLG perspective in this regard. The document emphasises that the targeted approach minimizes leakage (funds going to the non-poor) but administrative costs may be high. The universal approach, conversely, is easy to administer but may cost

more: “The universal approach remains a strong alternative to targeting, because it is simple, easy to make transparent, needs relatively little administrative and technical capacity, and is far more politically acceptable than targeting” (DPLG 2004a: Sec. 1).

In many municipalities, however, precisely those households that are likely to be *most* eligible for free basic services are also most likely to not receive them, even in a municipality with a universal subsidy or lifeline amount policy. This is because these are the households who most likely constitute the ‘infrastructure backlog’, that is who do not have the household infrastructure to receive the service in the first place. Thus the best that can be hoped for a FBS policy, in the absence of effective and efficient infrastructure extension, which enables more households to ‘come on board’, is that subsidy amounts do reach the poorest households who have access to the service. But in the poorest municipalities it is likely that significant percentages of households do not have access, and thus do not benefit from free or subsidised service provision. One option, of course, is to provide a cash transfer to all households qualifying for a free amount, rather than use a ‘method of transference’ (to use the DPLG phrase) which requires service infrastructure and administrative systems to already be in place. The DPLG document, in fact, list four methods of transference:

- Cash
- Voucher
- Deduction from Account when Billing
- Self-Selection<sup>79</sup>

Selection of the cash option would benefit all qualifying households, not only those who are connected. What matters fundamentally here is that municipal budgeting, in considering free and subsidized basic services, considers not only connected households but also unconnected households. If, as can be expected,

---

<sup>79</sup> No fee is charged but the service is so rudimentary that, it is assumed, only qualifying households would use it; a public tap is an example.

there is wide disapproval of a cash transfer, presumably in the name of a paternalistic concern for 'dependency' or irresponsible spending on the part of the poor,<sup>80</sup> then the accrued transfer amounts<sup>81</sup> can be regarded as a fund for infrastructural spending. It is admittedly naïve to assert that such 'accrued funds' would find their way into the municipal budget in the form of additional infrastructural funds: it is nevertheless a useful evaluative perspective which emphasises that the municipality has a service obligation not only to poor *connected* households, but also a 'service equivalent' obligation to poor unconnected households.

Where a targeted system is chosen the municipality will need to determine the selection mechanism. Two options appear to be feasible in this regard, namely a geographical approach and a behavioral one.<sup>82</sup> A geographical approach has the benefit of administrative simplicity. In this approach, particular areas within a municipality are identified as being, in all likelihood, characterised by household circumstances which would argue for subsidisation or FBS. Census or Income and Expenditure survey data could be used, though these may not be organised into areas in a manner useful to the municipality's needs. It is likely, however, that the municipality itself has information which can be used in this regard. Once areas have been identified, *all* households falling within the area are provided with FBS assistance, whether through a cash transfer, a voucher, or a credit on their service account. Clearly there is leakage in such a system, in the sense that some households living in an identified area would not qualify for assistance if an individual household-based indicator was used. However, the gains of the

---

<sup>80</sup> Debates on this kind of transfer would presumably follow similar trajectories to those associated with a basic income grant in South Africa; paternalistic concerns and the perceived risk of a culture of dependency,, quite apart from sustainability considerations, have featured heavily in these debates.

<sup>81</sup> Accrued in the sense that these are savings by the municipality as a result of not transferring funds which, on a strict equity criterion, it should be transferring to the poorest households who do not have infrastructure access.

<sup>82</sup> Using household income or expenditure information is not feasible, firstly because it would tend to be dated, unreliable and perhaps not even applicable, and secondly because it would be unclear how to move, technically, from this selection to providing free or subsidised services.

administrative simplicity of this approach are likely to outweigh the leakage losses, whilst still providing a system of selection that is more refined than a universal free amount approach.

An alternative is so-called self-selection and related behavioural indications that a household should receive FBS. Self-selection forms of access are generally rudimentary and communal. They should, especially where communal and rudimentary, be provided free, as the assumption is that only poor households, and especially households without any alternative forms of access to service infrastructure, would use them. The aim in the case of self-selection communal services should be that a free basic amount is provided to households who do not have other forms of access. Locationally, it is plausible to assume that such services would be concentrated in poor areas, that is the same areas identified for assistance of already connected households.

Thus, a targeted support strategy of a municipality, trying to balance administrative efficiency with allocative efficiency gains associated with accurate targeting, could try a range of scenarios as concerns subsidy amounts, areas eligible for support, and the mix of service support and rudimentary service extension. As discussed above, if equity is the objective, the free or subsidy amount provided as assistance to connected households should be balanced by measures which ameliorate the circumstances of the poorest, unconnected households. In poor municipalities these households will in all likelihood form a large part of a jurisdiction's households. This implies that subsidising consumption under such circumstances would be quite regressive, since the municipality's middle and upper income groups would reap the benefits.

### **c. The LES and MIG**

This section discusses the local equitable share (LES) and the municipal infrastructure grant (MIG), the two most significant sources of external funding to municipalities. These transfers are envisaged as supplementing the

municipality's own revenue. Furthermore, these transfers do not constitute a particularly large share of total revenue of all South Africa's municipalities. However, some municipalities are heavily reliant on the transfers as they do not currently raise appreciable revenue of their own. One of the key challenges, whenever a national government starts subsidising a sub-national government, lies in determining whether the sub-national government is in fact trying hard enough to raise own revenues, rather than 'sitting back' and waiting for transfers. This is the problem of assessing tax or revenue effort, that is the proportion of potential revenue due to the municipality which it in fact manages to collect. A given municipality may have little self-generated revenue because of the chronic poverty of its residents, but this lack of revenue may also be due to an administration that is unwilling to expend resources on establishing a functional municipal revenue collection system. The LES is also, however, a means of assisting a municipality in achieving equity objectives, since in principle it can be used to fund the financing gap which arises when some households are provided with services at less than the municipal cost of provision.

The LES is an unconditional transfer from national government to local government. It meets the Constitutional requirement that, through an Act of Parliament (the Division of Revenue Act), nationally raised revenue be distributed equitably among the national, provincial and local spheres of government. Both the LES and the PES can be regarded as Constitutional responses to the reality that the system of intergovernmental fiscal relations (IGFR) which developed in the post-apartheid economic dispensation was likely to be characterised by vertical and horizontal fiscal imbalances. The system was likely to have vertical imbalances because of the mismatch between Constitutional expenditure function and revenue power assignment between the three spheres of government. The system was likely to exhibit horizontal imbalance because of the vastly differing financial capacities of municipalities and the differing degrees of ability to pay for public services within them. Thus, the South African system of intergovernmental fiscal relations (IGFR), similarly to many such systems, uses

transfers from the national revenue fund to sub-national governments firstly in order to ensure that they are able to fulfill their expenditure functions and secondly in order to promote greater financial *convergence* between provinces and between municipalities. The ultimate objective in this regard is that South Africans in all municipalities have access to Constitutionally adequate services.

The LES share amount going to a particular municipality is currently the result of two fiscal decisions: firstly, the vertical division of revenue between the three spheres of government, and secondly and subsequently the horizontal division of revenue between the various municipalities. Constitutionally, the process of vertical and horizontal division is required to result from consultation with provincial governments, organised local government and the Financial and Fiscal Commission (FFC). The Constitution specifically requires that the recommendations of the FFC be considered.<sup>83</sup> The Division of Revenue Act, in articulating the vertical and horizontal division of nationally raised revenue between three spheres of government over the three-year period represented by the medium-term expenditure framework (MTEF), is required by the Constitution to take into account a range of factors (Constitution of South Africa, Act 108 of 1996: Chapter 13, Sec 214(2)(a)-(j)).

The Constitution also provides explicitly for the possibility of additional transfers, in Sec 214(1)(c) which indicates that the Division of Revenue Act needs to provide for “any other allocations to provinces, local government or municipalities from the national government’s share of that revenue, and any conditions on which those allocations may be made”, as well as Section 227 on sources of provincial and local government funding which indicates that local government and each province “may receive other allocations from national government

---

<sup>83</sup> The Intergovernmental Fiscal Relations Act, 1997 provides a framework for the various stakeholders and their interaction in generating the division of revenue; most pertinently, it sets out the timeframes and content of the recommendations of the FFC in the vertical and horizontal division of revenue.

revenue, either conditionally or unconditionally”. A clear Constitutional link is therefore established between the LES and municipal capacity to provide basic services and perform their functions, although the Constitution does not provide further detail.

The vertical division of revenue is determined by strategic shifts in policy interacting with the benchmark allocations represented by the preceding year’s budget. That is, given the need for predictability in public finances and the fact that much departmental expenditure is tied up in personnel commitments, legally binding social transfers<sup>84</sup> and the like, the best clue to a coming year’s division of revenue is probably the division of revenue of the previous year. The South African division of revenue has seen some increased shifts in the share going to local government at the expense of national government, but this has been primarily due to the need to compensate municipalities for the loss of RSC levies rather than a shift in the prioritisation of functions. Though there have been some calls, notably from the FFC (2001), for a more formulaic approach to the vertical division of revenue, the kind of policy rigidity stemming from a vertical formula would probably be undesirable since it would make it difficult for fiscal policy to be adequately responsive to exogenous changes.

Determination of both the PES and the LES (that is the horizontal division of revenue), however, is based on a formula which seeks as far as possible to base transfer amounts on factors which fall beyond the power of the provincial government or municipality to manipulate, which are an adequate indication of municipal household *need* and municipal own revenue generation potential, and which do not remove incentives for the municipality to collect all revenues due from its own revenue bases. The formula’s underlying aim is to close the financial gap between municipalities, that is the difference between the cost of delivering

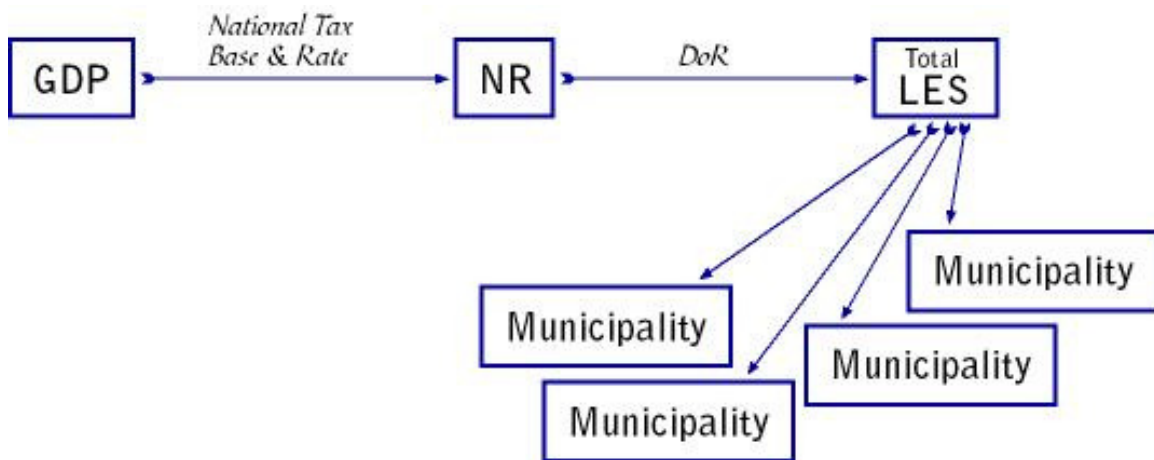
---

<sup>84</sup> For example the child support grant and the state pension: primary care givers to children and pension age individuals who meet a means test have a legal claim on a direct income transfer to the amount government has committed itself to at that point in time.



their functions and the revenue they are able to raise through the property tax and charging for services. The extent to which national government is able to close the financial gap for all municipalities and for any given municipality is of course constrained firstly by total national revenue raised and secondly by the slice apportioned to the local sphere in the division of revenue. A basic flow chart would represent the process as follows for municipalities<sup>85</sup>.

**Figure 6.5: From GDP to Municipal LES**



Annexure E of the *Budget Review* every year is entitled ‘Explanatory memorandum to the division of revenue’ and generally contains a review of the PES and LES as well as setting out the executive’s responses to the proposals of the FFC. The descriptive parts of the LES discussion below draw heavily on Annexure E of the *Budget Review 2007*.<sup>86</sup> The current formula was introduced in 2005 and the 2005 *Budget Review* provides a useful summary of some of the

<sup>85</sup> ‘NR’ refers to nationally raised revenue.

<sup>86</sup> A simple distinction which often seems to cause confusion needs to be cleared up at the outset. The LES formula is used to calculate the total share of local sphere funds going to a municipality. It has no inherent relationship to how the municipality *actually* spends that money, since this proceeds through the politically quite autonomous municipal budgeting process and is also a function of the quality of municipal financial management (that is the extent to which spending outcomes approximate the budget).

thinking behind it as well as the debates between the National Treasury and the FFC (National Treasury 2005: 225-272).

The current LES formula is as follows for a given municipality A:

$$LES_A = BS + D + I - R + C$$

The basic services component (BS) aims to support water reticulation, sanitation, refuse removal and electricity reticulation as the core services. The subsidy amount is derived from the number of poor households (income of less than R 800 per month) in the municipality. A full subsidy amount is attributed for households who have access but income below the R 800 threshold value. An amount of one third of this is provided for each poor household without access. The actual (updated) subsidy amounts, for 2007/2008, are given in a table in annexure E which is duplicated below:

**Table 6.2: LES Basic Service Component Amounts**

	<b>Serviced Households</b>	<b>Unserviced Households</b>
<b>Electricity</b>	R 40	R 15
<b>Water</b>	R 30	R 10
<b>Refuse</b>	R 30	R 10
<b>Sanitation</b>	R 30	R 10
<b>Total</b>	R 130	R 45

Source: National Treasury (2007: 254)

D is called the development component. According to the annexure it “has been set at zero since the inception of the current formula and will remain inactive until a suitable factor can be found that adequately captures the developmental needs of local government” (National Treasury 2007: 252). This is an exceedingly vague

formulation and it is suspected that such a component will always be vague. It should be scrapped or subsumed under other components of the formula.

The institutional support component (I) seeks to ensure that all municipalities possess enough resources to maintain at least the basic institutional requirements of local government. The annexure points out that many poor municipalities are “unable to raise sufficient revenue to fund the basic costs of administration and governance” (National Treasury 2007: 251). The allocation has a base component for all municipalities, a population component and an amount derived from the cost of maintaining councilors (National Treasury 2007: 252). R is termed a revenue-raising capacity correction and is included to create an incentive for municipalities to collect revenue locally, that is to enhance both their tax effort and the effectiveness of billing households for municipal services. C is termed the correction and stabilisation factor and essentially guarantees that no sweeping changes will be made over the MTEF: this enables a municipality to plan on the basis of a fairly stable and predictable revenue flow over the medium term.

Additional complications are associated with the fact that basic service responsibility is not always assigned identically between metros, local and district municipalities, as a result of historical factors as well as current capacity constraints. Thus, the formula is perhaps best regarded as a starting point for LES determination. Furthermore, as the Annexure E of the *Budget Review 2007* points out, the rand amounts determined according to this process do not necessarily match the vertical share allocated to the local sphere (National Treasury 2007: 253). Thus, provision is made for ‘rescaling’ of the BS, D and I components of the formula to fit the local sphere share. Since D is currently set at zero, this essentially means that the BS and I amounts are adjusted to accommodate the overall local sphere allocation.

The re-scaling aspect mentioned at the end of the previous paragraph essentially means that cabinet-level decision-making on the *vertical* division of revenue has large significance for the extent to which national revenues supplement municipal own finances. There is no formula or policy which explicitly commits cabinet to a particular division of revenue. Thus the approach is largely a 'top-down' one: the vertical division is made by cabinet and from this the formula determines the distribution of LES shares. This approach can be contrasted with that advocated by the FFC (FFC 2001) which would give more weight to deriving the local sphere share from estimates of its resource requirements if it is to deliver on its constitutionally assigned mandate.

Although there is thus no explicit or formal commitment from the executive to ensure that the local sphere receives adequate support, it is clearly necessary for the viability of the fiscal system in its entirety that the local sphere receives appropriate support. This support, however, should enable the municipality to achieve equity objectives rather than diminish its incentive to collect revenue from the property tax and for utility services rendered.

It is quite difficult to assess the process analytically in more detail than this. As mentioned above, the fact that utility services responsibility in some cases resides with a district municipality and in some cases with a local municipality means that a particular LES may or may not correlate with a particular selection of indicators such as poverty and population. This is one instance of the difficulty of trying to come to conclusions about the way South African local government is expected to and succeeds in dealing with distributional matters, namely that the caveats, exceptions and circumscriptions often outweigh the conclusions which can be derived.

There is clearly a redistributory component represented by the BS component, which explicitly links the share going to a municipality to its number of poor households. The LES however, does not prioritise allocations to municipalities

with infrastructure backlogs: a municipality with 5000 poor households, all of whom enjoyed service access, would receive a larger amount, all else being equal, than a municipality with 5000 poor households of whom 2500 had access and 2500 still didn't. In this sense the LES is a transfer towards the recurrent expenses of the municipality. It is therefore essential to regard it in conjunction with the municipal infrastructure grant (MIG), which is aimed at backlog eradication.

The MIG<sup>87</sup> is a conditional grant from national government which supports the municipal capital budget and consolidates a number of infrastructural grants into a single transfer. The vision of the MIG is to provide all South Africans with "at least a basic level of service by the year 2013 through the provision of grant finance aimed at covering the capital cost of basic infrastructure for the poor" (DPLG 2004b: 1).

As in the case of the LES, the MIG formula consists of a number of components, The horizontal division determined by the formula "takes account of poverty, backlogs, and municipal powers and functions" (National Treasury 2007: 255). Specifically, the formula consists of the following components:

$$\text{MIG} = \text{B} + \text{P} + \text{E} + \text{N} + \text{M}$$

The key provided in the annexure explains each component as follows (National Treasury 2007: 256):

- B - Basic residential infrastructure
- P - Public municipal service infrastructure

---

<sup>87</sup> The descriptive part of this section on the MIG draws on two policy documents: *The Municipal Infrastructure Grant: Basic Level of Services and Unit Costs: A Guide for Municipalities* (DPLG 2005) and *Policy Framework for the introduction of the Municipal Infrastructure Grant (MIG), concise version* (DPLG 2004).

- E - Allocation for social institutions and micro-enterprises infrastructure
- N - Allocation to all nodal municipalities
- M - Negative or positive allocation related to past performance of each municipality relative to grant conditions

The MIG is explicitly regarded as a transfer intended to *fully* subsidise capital costs of basic services to poor households, through the provision of bulk, connector and internal infrastructure in key services (DPLG 2004: 3). A key challenge, consequently, is that of appropriately targeting the MIG, not only between municipalities but also between various possible uses within the municipality which conform to the nationally set conditions for the grant. Like the LES, the MIG is determined by formula.<sup>88</sup> However, unlike the LES, the inclusion of M makes specific provision for tying receipt of grant funds to municipal performance in the use of such funds. Municipalities are required, over a medium-term period, to spend funds on basic infrastructure services roughly in proportion to their weighting in the formula employed to determine the MIG amount. The component M of the formula provides for a “(n)egative or positive allocation related to past performance of each municipality relative to grant conditions” (National Treasury 2007: 256). The table below shows MIG amounts over the MTEF together with total LES amounts for the same period.

**Table 6.3: Budgeted and indicative MIG and LES amounts for 2007/08-2009/10 in current rands**

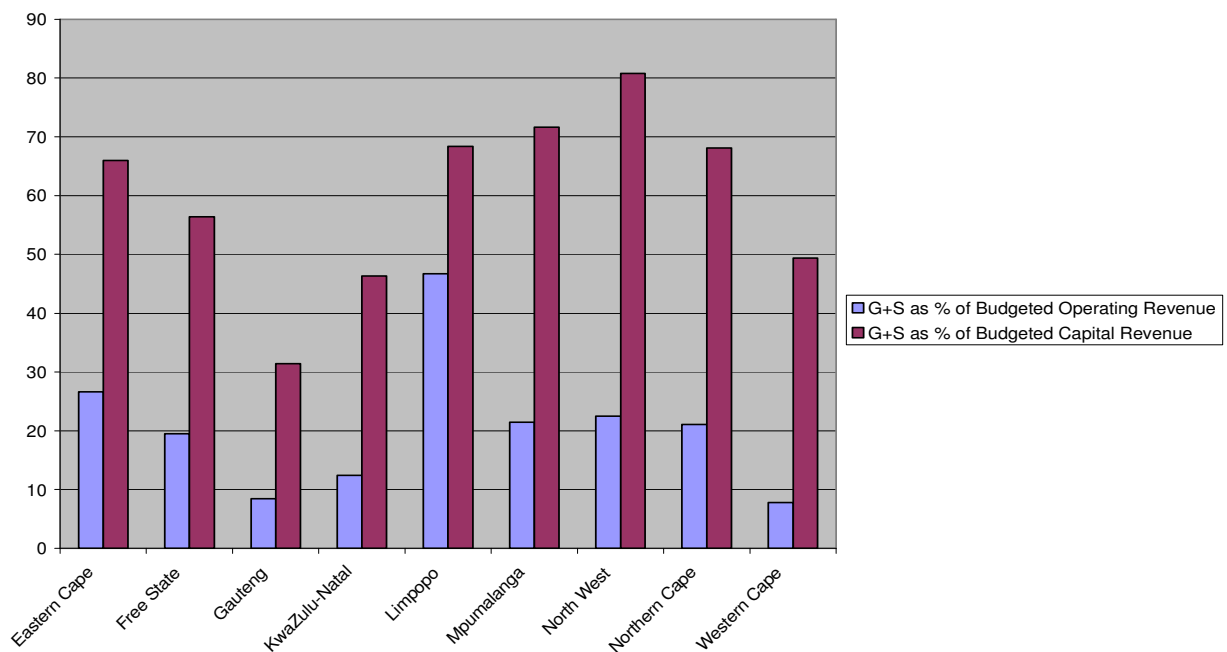
	<b>2007/2008</b>	<b>2008/2009</b>	<b>2009/2010</b>
<b>Total LES Allocation</b>	20 675 620	23 774 767	29 444 144
<b>Total MIG Allocation</b>	7 548 564	8 053 090	9 130 230
<b>Ratio (LES / MIG)</b>	2.74	2.95	3.22

Source: Division of Revenue Act, Act No. 1 of 2007: Schedule 1&2.

<sup>88</sup> The formula and discussion can be found in the *Budget Review 2007* (National Treasury 2007: 256).

One of the primary conclusions that can be drawn about the MIG at present, and about municipal capital spending more generally, is that it remains heavily reliant on transfers, that is the 'Grants and subsidies' category of budgeted revenue in the capital budget summaries provided in the National Treasury's *Local Government Budgets and Expenditure Review 2001/2002-2007/2008*. Figure 6.6 compares share of grants and subsidies as a percentage of budgeted revenue by province for both the capital and operating budgets for 2005/2006.

**Figure 6.6: Budgeted Grants and Subsidies as Percentage of Total Budgeted Municipal Operating and Capital Revenue, 2005/2006**



Source: National Treasury (2006: 85).

It should be noted that the actual (i.e. budget outcome) percentage of grants and subsidies as a share of capital budget revenue is likely to be even *larger* than the budgeted share, since grants and subsidies in general represent a legal obligation on national government to disburse what has been allocated, whereas estimates of revenue from internal income are less certain and may be overly

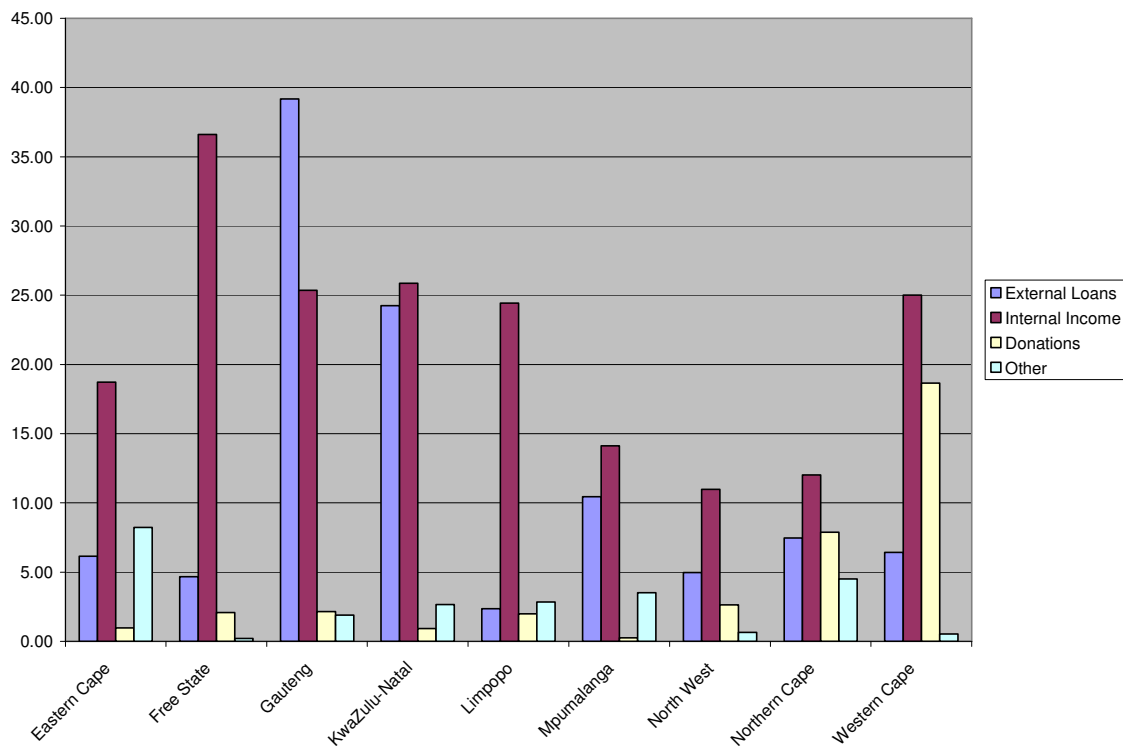
optimistic. Thus the comparison between operating and capital revenue budgets provided below, which is already revealing, is likely to be exaggerated were operating and capital budget *outcomes* to be compared.

The question of whether municipalities are allocating enough of their own funds to capital spending, that is to say whether their budget prioritisation is appropriate, is very difficult to answer rigorously, not least because of difficulties in determining an appropriate measure for allocating municipal resources between operating and capital requirements. Figure 6.6 suggests that, in general, municipalities are not allocating enough of their own funds to capital spending. The figure provides some indication of how small the MIG allocation is relative to the LES: yet the MIG constitutes a larger source of capital financing for most municipalities than own sources. Although nothing *prevents* a municipality from financing capital assets from its operating revenue, this does not appear to be the case. Thus, the problem of inadequate municipal own capital expenditure, notwithstanding questions of prioritisation and poor revenue effort, must also be linked to difficulties in borrowing experienced by municipalities.

External loans made up an almost negligible part of budgeted capital revenue in 2005/2006 for all the provinces except Gauteng and KwaZulu-Natal. Since grants and subsidies are allocated on the basis of a set of criteria quite different from that of a credit rating agency, municipalities that are unable or unwilling to borrow must finance out of their internal income or, as frequently happens it seems, do not finance much capital spending at all from their own income. Figure 6.7 below shows external income, internal income, donations and 'other' as percentages of total budgeted capital revenue in 2005/2006.



**Figure 6.7: Non-Grant and Non-Subsidy Capital Budget Revenue, 2005/2006**



Source: National Treasury (2006: 85).

A last point to note, which was strongly emphasised by the World Bank (1994) in their review of infrastructural issues concerns the problem of inadequate allocations to *maintenance* of physical capital in the operating budget. It is difficult to say with any certainty what a ‘good’ maintenance budget would look like in the South African context. Anecdotal evidence also suggests that many municipalities themselves don’t know because they lack sufficient engineering and planning specialists.

The LES and the MIG do promote equity because they are based on a formulaic approach which prioritises transfers to poorer municipalities and municipalities characterised by greater infrastructural backlogs. In both cases the amounts going to particular municipalities are constrained by the vertical division of

revenue, which in turn is determined both by nationally collected revenue<sup>89</sup> and the cabinet-level prioritisation of sectors and associated spheres of government.

#### **d. Equity-Orientated Municipal Policy Options: Price Intervention and Property Taxation**

The use of transfers such as the LES and the MIG essentially shift the burden of equity-orientated policies with respect to basic services onto *all* tax payers, both current ones and also future ones if debt-financing is used to expand the amount of resources available to the national government. The conventional view, discussed in Chapter 2 in the context of fiscal decentralisation and the three traditional objectives of public finance, is that local redistribution tends to be counterproductive given the greater mobility of residents between jurisdictions compared to that between countries. This greater mobility, as emphasised by Tiebout, means that local redistribution would in fact lead to worse social outcomes: poor residents would move to redistributory municipalities and rich residents (ignoring altruistic motives) would move away from them. Inequality between individuals would not be reduced as much as if national government were responsible, and inequality between jurisdictions would in all likelihood increase significantly.

However, previous sections of this chapter have suggested that the transfers flowing to a municipality for extending and maintaining basic services to poor households (that is free or heavily subsidised provision) are inadequate to address backlogs and enable an appropriate level of consumption for connected households. The fundamental reason for this is two-fold and simple: firstly, South Africa remains a middle-income developing country with a relatively small GDP compared to developed economies (though huge compared to average sub-Saharan Africa economies): as such the aggregate resources available to

---

<sup>89</sup> Itself a function of economic growth, tax policy, tax effort and the 'tax culture', that is the extent of compliance and evasion.

Government are limited; secondly, the division of revenue has tended to not prioritise national transfers to the local sphere compared to transfers to the provincial sphere,<sup>90</sup> for two reasons. Firstly, it has been assumed that the revenue instruments available to local governments should in many cases generate sufficient resources to allow fulfilment of their functions. Secondly, the functions associated with the provincial sphere, such as education, health and social development, are pressing social ones which also require significant resources.

If the share of nationally raised revenue going to the local sphere is limited by broader economic trends, cabinet-level decision-making on the division of revenue, and the urgent demands of the provincial sphere, then it may become necessary for local governments to engage in some form of redistributory financing themselves in order to attain more equitability in both access to and consumption of local services within the jurisdiction. The costs of redistribution for the economy as a whole may even be reduced if it is assumed, following the fiscal decentralisation benchmark model, that decentralised redistribution would allow customised or differentiated free basic service levels and that the aggregate degree of redistribution would then be lower than if national government attempted to provide free services at a uniform level to everyone. A final reason why local redistribution could be warranted is that the provision of basic services to poor, formerly excluded households is likely to generate significant positive externalities whose effects are limited to the municipal jurisdiction.

The use of a progressive unit rate structure is a form of pricing which may further equity objectives through cross-subsidisation from higher to lower consumption households.<sup>91</sup> Under such a price structure, different amounts of a service are

---

<sup>90</sup> In the simple sense that total conditional and unconditional transfers to the provincial sphere are much larger than those to the local sphere.

<sup>91</sup> Deedat et al (2001) is a critical evaluation of the application of this approach in one South African municipality.

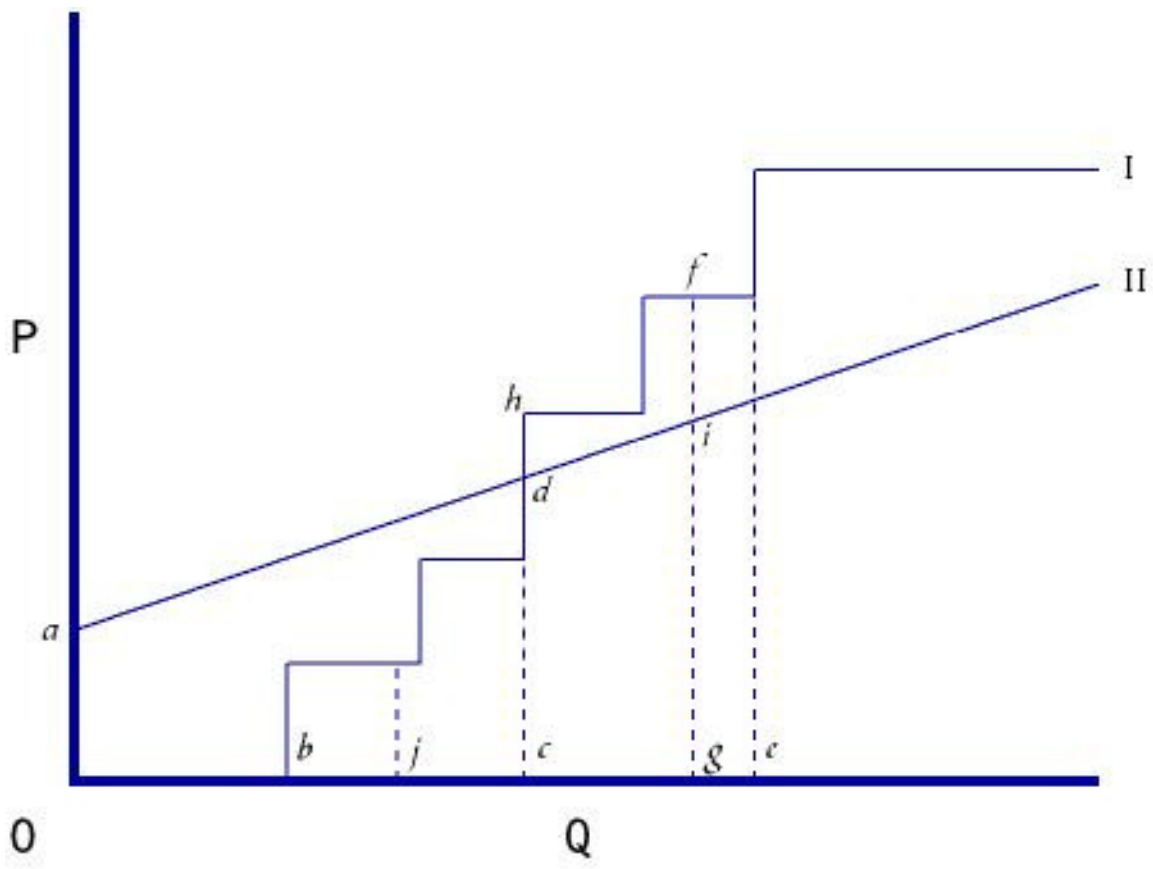
priced at different rates, with the unit rate increasing with consumption. In principle pricing could be specified as a smoothly rising rate where each unit would be charged a tiny amount more. In practice, for administrative reasons unit rates tend to apply to 'blocks' of consumption. Thus, the system is also called a 'block tariff' system.<sup>92</sup> If higher consumption is identified (as appears reasonable in most though not all cases) with greater wealth and income, the fee schedule is progressive in the same manner as a progressive income tax schedule is. Escalating rates can fund internal cross-subsidisation (that is subsidisation within a service) if the rates charged per unit of consumption at the upper range are sufficiently in excess of the cost-recovery rate to generate a revenue surplus from the service which can fund the 'lower than cost' rate at the bottom end.

Figure 6.8 represents a block-tariff design where rates go up with units consumed. In this version, a free amount  $O_b$  is supplied to *all* households (ie a universal subsidy approach is followed), which would presumably correspond to a lifeline amount as articulated in national and local policy.  $I$  represents the supply schedule for the purposes of pricing,  $II$  represents the true supply schedule to the provider and  $O_a$  represents fixed costs, after which costs rise continuously; no economy of scale benefits are assumed, in other words. This representation can be used, from a municipal planning point of view, for scenarios including households with and without access, or only for households with access. Where non-connected households are to be included the fixed cost component would include capital costs and the curve  $II$  would rise at a steeper rate to represent greater costs per unit as more households are serviced and assuming no economies of scale.

---

<sup>92</sup> World Bank (2006) *Water, Electricity and the Poor: Who benefits from utility subsidies* provides useful discussions, especially Chapter 2, 'A Typology of Consumer Utility Subsidies'.

**Figure 6.8: A Representation of Cross-Subsidisation Structure Within a Single Service**



Focusing for the moment on households with access, that is curve II, it should be apparent that up to a consumption amount  $O_c$ , the marginal cost of provision is in excess of the unit cost charged to households, as represented by the difference between the two lines. If total consumption equalled  $O_c$ , the total loss to the municipality would accordingly be represented by the area  $aObd$ . Point  $Ad$  is the pivot point: here the marginal cost charged to the household for the first time equals the marginal cost of provision by the municipality. For consumption amounts at and beyond  $c$ , the unit rate exceeds the municipal cost of provision. At some consumption amount the unit rate will presumably attain a maximum and will remain constant: this is represented in Figure 6.8 by amounts greater than

Of. If the objective of the service is to be redistributionary but simultaneously to 'break even', that is to recover operating costs through operating revenue, then the 'loss' represented by consumption below an amount of  $O_c$  needs to be recovered from the 'profit' per unit consumed beyond  $c$ . Imagine, for example, that the dotted line  $fg$  represents total consumption. Then the area represented by  $dhfi - adbO$  would represent the loss or profit incurred by the municipality *on this service account*. It follows, therefore, that if the municipality wishes to recover costs on this particular service, it needs to both design the unit rates and, crucially, estimate demand in such a way that the two areas are equalised.<sup>93</sup>

Estimating demand will, as always, also require working estimates of the unit price elasticity of demand. If the municipality gets its sums wrong, for example in underestimating demand elasticity, the repercussions may be quite severe. The service may, for example, run at a loss which needs to be made up elsewhere or through painful belt-tightening measures further down the line. Where demand is inelastic but users can't afford the post-block tariff rate where it exceeds the cost recovery rate, users may simply consume and not pay. The municipality's balance sheet may look reasonably healthy until a closer look is taken at the percentage of non-performing debts on its books.

There is no reason why unit rates need to rise smoothly as consumption increases: a municipality may, for example, decide to provide the amount  $bj$  at a low rate to benefit lower income households especially, and it may decide to charge very steep rates beyond a certain consumption amount. In this example only households as users of services such as water and electricity are considered: the analysis would remain formally similar but would become more complicated if the behaviour of firms was also introduced. The demand of bigger firms could provide valuable additional service-generated revenue to the municipality as bigger firms would in all probability consume amounts at the upper range of the schedule; on the other hand, there may be some grounds for

---

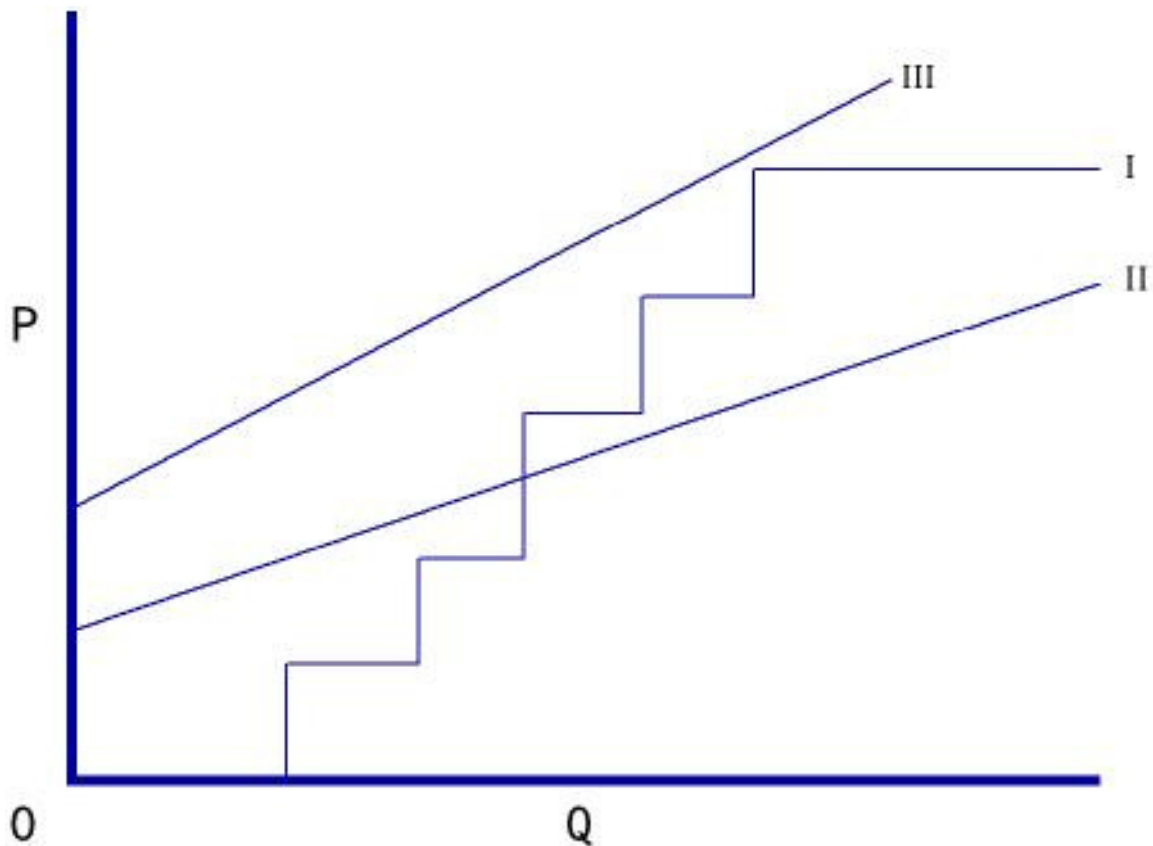
<sup>93</sup> The lifeline amount to be provided free is treated as exogenous in this example.

preferential treatment of firms through lower unit rates, especially in attracting prospective firms to the municipality. Such decision-making would again be a complicated mix of politics and economics. The National Electricity Regulator of South Africa (NERSA) suggests in this regard that, in fact, the price determination of electricity be conducted independently of consideration of the *use* it is put to (NERSA 2004: 1). This amounts to an assertion that high-consuming firms would pay a higher average rate than households for utility services consumed.

In principle this approach can also quite easily accommodate subsidisation *between* services. The challenge then becomes to equalise total 'losses' with total 'profits', rather than do so for a particular service. Formally the municipality would have more room to manoeuvre, since it could set unit rates more aggressively for services with lower price elasticities of demand. Also, and more contentiously, the consequences of 'getting the prices wrong' for a particular service would be less severe since there would still be a chance that unanticipated in-year operating shortfalls could be made up from surpluses on other service accounts.

A curve III has been included in figure 6.9 to visually represent the challenges associated with including infrastructure costs. Here, the capital expenditure on providing infrastructure for household services is regarded as a component of fixed cost for a given planning period, say one year. Because additional connections mean additional demand, the supply curve will also tend to have a steeper slope. The danger is that, as with curve III as drawn here, the municipality ends up with a situation where cost recovery rates (now higher) are no longer equalised at any consumption amount with costs, and therefore significant financial losses are incurred by the municipality.

**Figure 6.9: A Representation of Cross-Subsidisation Structure Within a Single Service, Including Infrastructure Costs**



There are a number of arguments against the manipulation of pricing to achieve equity objectives. Fundamentally, however, they can be divided into concerns around information and concerns around incentives. As the above discussion has probably made clear, the use of pricing to achieve cross-subsidisation requires good estimates of the values of a number of key variables. Some of this information is simply difficult to obtain, in the sense that it requires allocation of municipal resources. Others, such as the response of households in the upper consumption ranges to significant increases in the unit rate for consumption of utility services, is dependent on so many possibly factors that estimation



becomes difficult, if not impossible. Although the framework can be set out and discussed as above, the reality is likely to be far more complex. On the other hand, a brand new, untested cross-subsidisation approach is not the only way of getting some form of redistributionary pricing into municipal provision. Most municipalities who provide household services have practiced some forms of differential pricing and therefore possess some empirical information on household behaviour and other relevant circumstances in their jurisdictions. In many cases the best approach will therefore be to make equity-orientated adjustments to these kinds of existing practices.

A more significant objection to the use of cross-subsidisation is that, as with all approaches which distort the working of price, which the most fundamental unit of information in economics, is that it conveys the wrong information to users on the relative scarcity of the good or service they are consuming. Where price is below cost households will tend to over-consume, where it is above cost they will tend to under-consume, from the perspective of social welfare. There is no reason to assume that the *total* production and consumption which results is optimal or is the same as that which would follow from a neutral two-tier pricing approach, such as outlined in chapter 4. Furthermore, the diagrammatic illustrations provided above are misleading in that they convey the impression that there is a true supply schedule and then, separate from this and clearly and explicitly derived from this, an 'equity-orientated' schedule. In reality, once equity-style distortions are introduced, it becomes more and more difficult to ascertain what 'true' supply and demand schedules for a particular service might look like. This appears to be the central object of NERSA to the use of cross-subsidisation to achieve equity objectives. This concern brings the discussion straight back to the fundamental efficiency-equity trade-off issue, since in essence the concern here is that price intervention for the sake of equity reduces efficiency to an unacceptably high extent.

The second option for intra-municipal financing of basic services on a non-cost-recovery basis is to use local tax revenue, that is, in the South African context, the property tax. The value of land and/or improvements is in many cases a good indicator of ability to pay, and the immobility of the tax base ensures that it constitutes a viable local tax. The key question related to the property tax, from the perspective of this study, is whether it is currently utilised appropriately and whether it constitutes a potentially significant enough source of revenue.

In SA the legal aspects of property taxation are established by the *Municipal Property Rates Act* (Act no 6 of 2004). A great deal of debate and re-drafting preceded the enactment of this legislation, with a key point, not surprisingly, being the distribution of power between the national and local governments when it comes to determining property tax bases and rates.<sup>94</sup> Local governments in South Africa are given a large degree of autonomy to set tax bases and rates. What is perhaps most pertinent is the extent to which the property tax could provide significant *additional* revenue to local governments seeking to expand resources in order to provide free basic services and eliminate infrastructural backlogs more rapidly. As Bahl (2001) points out, citing the Department of Finance, the property tax constitutes a significant share of local revenue collected in South Africa by international standards. He points out that, at the time of his study, although the property tax did not constitute a significant share of total government revenue, it did constitute a large (11%) share of total provincial and local government expenditures: that is, it is a significant source of sub-national, and specifically local, financing (Bahl 2001: 5). In fact, South African property tax revenue in 1995 amounted to 1.95% of GDP, whilst its level of per capita GDP would lead to a *predicted* share of only 0.62% of GDP (Bahl 2001: 7). Thus, South Africa exerts a high 'property tax effort', and it cannot be stated that the property tax is being underutilised as a source of revenue, at least not when an international comparative approach is used to generate benchmark values.

---

<sup>94</sup> Bahl (2001) is a useful summary of some of the theoretical issues in this regard in South Africa.

On the other hand, doubts exist as to the extent to which property tax revenues collected are optimised since in many cases valuation of property is not up to date and historical values are used. In a suggestive statement, Bahl (2001: 5) argues that “(w)here the property tax fails is in the area of fairness in administration at reasonable cost. Identification of the tax base, and assessment, are very expensive propositions, and properly done, collection and appeals impose significant costs. In South Africa these costs have been avoided in some sense by failure to upgrade the property tax rolls, and failure to force a high rate of compliance.” One conclusion stemming from this would be that property rates in South Africa are quite high but that property tax effort appears to be quite low, in that various administrative lapses result in collected revenue to be significantly less than what is due. It may then also be the case that significantly improved tax effort (in the case of the property tax this would include updated valuations which in fact expand the value of the tax base) would lead to greater tax resistance amongst tax payers. The challenge, from the perspective of an individual municipality, is to maximise revenue by selecting the optimal tax base and rate combination, subject to considerations of the revenue elasticity of the effective tax rate.

Unfortunately, little empirical data exists on property tax collection rates as well as the *realism* of municipalities in estimating how much revenue they will collect in this way in the coming year, that is the realism of property tax revenue estimates as a component of total budgeted revenue. The tables in the *Local Review* require municipalities to provide an estimate of the ‘expected collection rate’ and compares this to a previous year’s estimated collection rate, but no comparisons are made of these figures with actual collection rates.

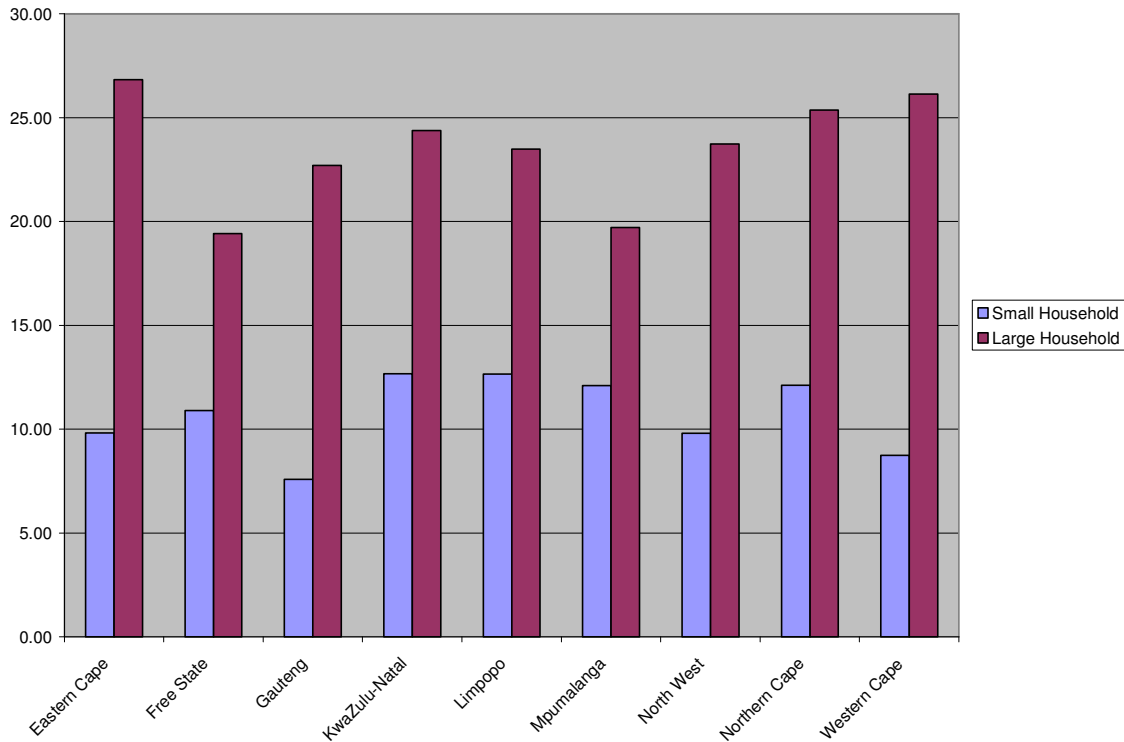
A further issue is the extent to which municipalities are (and should) make a greater effort to bring households in historically non-white areas into the property tax environment. The *Local Review* is quite interesting in this regard, as it

presents tables for monthly municipal accounts for both ‘small’ and ‘large’ households. Small and large households are distinguished as follows:

- Small: 300m<sup>2</sup> erf, 48m<sup>2</sup> improvements, 498 units electricity per month and 25 kl water.
- Large: 100m<sup>2</sup> erf, 150m<sup>2</sup> improvements, 1000 units electricity, 30 kl water.

All else being equal, smaller households would tend to be poorer and larger households wealthier, though any conclusions derived from this conclusion can be no more than tentative. On the basis of this, average monthly accounts for rates and services are provided by province. Figure 6.9 below provides this information.

**Figure 6.10: Property Tax as Percentage of Small and Large Household Monthly Accounts**



Source: National Treasury (2006: 112 & 135).

Of course larger households pay a larger absolute amount of property tax since the size of the erf as well as scope of improvements here constitute one of the defining attributes of 'large households'. However, what is less easily explained is the difference in the percentage contribution made by the property tax. It is hard to think of a theoretically valid reason why the ratio of property tax to the total monthly account should differ so clearly for small and large households. These values suggest one of the following problems:

- Smaller households are not paying enough property tax
- Smaller households are paying too much for trading services
- Larger households are paying too much property tax
- Larger households are not paying enough for trading services

The final chapter deals with challenges and this question is explored further there.

## Chapter 7

### Conclusion: Issues and Recommendations

#### a. Introduction

This study set out to identify and investigate key issues in the provision of basic local services in South Africa from the perspectives of both efficiency and equity. The preceding chapters described and evaluated the current state of municipal financial and non-financial performance and the policy context of national government support to local governments. The study also spent some time defining a benchmark model of fiscal decentralisation and grappling with both the efficiency and equity perspectives on the provision of basic local services.

The purpose of this final chapter is to identify and further elaborate on some of the matters which arise from the discussions of the preceding chapters and which in many cases were already highlighted in those chapters. Where possible, recommendations are also provided. Some of these issues concern the quality of the information currently available to researchers, and some concern issues of policy. Section b addresses the former and section c addresses the latter.

#### b. Quality of Information on the Financial and Non-Financial Performance of Municipalities

In preceding chapters a number of discussions referred to problems in drawing usable conclusions from the way information has been presented and the underlying conceptualisation involved. Cases were also identified where the conclusions in reports cited did not appear to follow very clearly or straightforwardly from the information used.

One problem which appears to exist is the degree of leeway municipalities are given in reporting on key aspects of free basic services policy and the extent to which the required infrastructure exists. In most cases used in this study the

problem is primarily with the vagueness and/or the generality of questions posed. Thus the *Non-Financial Census of Municipalities* required municipalities to simply indicate whether they had a free basic services policy 'in place'. Responses to such a vague question do not mean much from a research and evaluation point of view.

Similarly, municipalities were asked whether, for particular services, they had 'infrastructure in place'. Again, this is far too vague a question and not much can be done with statistics generated through this. Most municipalities who are assigned a particular function would have some infrastructure 'in place' for delivering the required service: few, if any, would have infrastructure in place to accommodate *all* households who might want the particular service. Thus, in the absence of further specification by the compilers of the questionnaires, municipalities are free to interpret the sense of the question as they wish. The question does not deal with infrastructure backlogs, which of course is what really needs to be known.

It is perhaps unfair to criticise a statistical agency when the manner in which it presents its information does not meet the requirements of a particular research study. However, it is suggested that the problems mentioned here do not concern the particular informational needs of this study, but impede an accurate sense of the state of municipalities more generally. It appears that the designers of the questions not only did not have the kinds of questions in mind that researchers might reasonably want answered, but were also inadequately informed on the nature and quality of data which local governments required in order to plan and deliver services which ensure public value for money.

This problem was especially pertinent in the attempt in chapter 5 to assess the affordability of water to South African households. There is admittedly an inherently subjective aspect to what one regards as 'affordable,' and the statistical authorities cannot be blamed for this. However, both the *General*

*Household Survey 2005* and the *Non-Financial Census of Municipalities 2005* were less useful than expected. In the former case, estimates were provided of monthly payments for water, but with no consideration of the distinction between the role of price and volume in determining cost differentials. In the case of the latter, the main problem was that accounting cost rather than economic cost was used in estimates of water purchased by municipalities. As discussed, this made it difficult to draw any conclusions of economic significance.

It is recommended that the manner in which questions are posed to municipalities compels them to provide more specific indications of progress made in, for example, free basic services policies, and especially more specific estimates of infrastructural backlogs. In addition, the manner in which information is conceptualised and presented needs to ensure that clear and meaningful economic conclusions can be drawn from them.

A related problem is that questions were not sufficiently informed by an awareness of the strategic intent municipalities might bring to bear in answering vaguely formulated questions. Thus, for example, asking whether municipalities feel they have enough finances to support planned FSB initiatives may be quite useless since many would respond in the negative in the hope of receiving more funds or at least not jeopardising planned increased funding to their jurisdictions!

A further challenge to any study which takes a 'bird's eye' view of trends in local finances and delivery is the heterogeneity of local government circumstances. A more comprehensive study might do well to select a few 'case study' municipalities (perhaps one metro, one 'urban' local municipality and one rural local municipality) and delve more deeply into the unique circumstances and problems they are experiencing. Of course, and unavoidably, such a case study approach would then have to be treated with caution regarding the extent to which conclusions could be generalised. It is, however, suggested that a case



study approach would be a useful complement to the broader perspective taken here.

A final concern associated with the conceptualisation, collection and presentation of information is the whole notion of 'indigency' and the erroneous approach it establishes for targeted free basic services provision. As mentioned in the discussion in chapter 6, such an approach, in addition to being administratively problematic, creates an unnecessary stigmatisation of free basic services.

### **c. Policy Issues and Recommendations**

#### **i. Budgeting for the Maintenance of Assets**

It is one thing to suggest that municipalities are not budgeting adequately for the maintenance of their assets, including their infrastructural assets, and that this results in a costly capital rehabilitation and replacement need which jeopardises service delivery. It is, however, quite difficult to assess firstly the adequacy of capital maintenance allocations in a given budget, and secondly the adequacy of planning for medium- and long-term capital maintenance allocations. The latter, particularly, depends on a number of variables, such as trends in municipal demand as a result of migration, changes in income, and changes in preferences, as well as changes in the revenue effort of the municipality and trends in capital transfers to the municipality from national government.

Guidelines to municipalities need to ensure that municipalities have a clear sense of the challenges in planning for capital maintenance allocations, given the variables listed above. However, the problem of capital maintenance allocations is equally, if not more, attributable to the politics of budget-making and the fact that allocations which appear to always be shiftable to the 'next budget' are all too easily shifted in favour of more 'urgent' salary demands and the like. Unless

there is a political commitment to ensuring adequate maintenance, and this commitment is allied to clear guidelines regarding appropriate maintenance budgeting, the problem is likely to remain.

It needs to be recognised that relations between local and national governments are strategic. The national government has funds to transfer but is at an informational disadvantage relative to the municipality and must necessarily rely to some extent on the municipality's own reports regarding its circumstances. There are clearly benefits for the local government to exaggerating the scope and depth of poverty within its jurisdiction.

In the case of capital assets and their maintenance, a further strategic consideration may come into play, one which would certainly constitute an interesting research question in the South African context. Asset maintenance is financed through the operating budget, which is predominantly funded through own revenue. On the other hand, for many municipalities the capital budget is primarily funded through national transfers. Local governments may consequently have an incentive to *under-provide* for maintenance, knowing that asset replacement and rehabilitation costs will not fall on residents of their jurisdiction to the same extent that maintenance costs do. Such strategic behaviour becomes more likely the larger the information asymmetry between national and local government regarding local conditions and finances.

## ii. Revenue Effort

As discussed, the LES and the MIG provide for conditional and unconditional transfers to municipalities. A number of other conditional transfers also exist whose aim is to promote municipal attainment of nationally shared objectives. Both the LES and the MIG are valuable means of supplementing municipal own income. However, as in any system of intergovernmental fiscal relations which includes transfers, the question is to what extent such transfers in fact reduce

incentives for municipalities to increase their own revenue efforts. Whilst the larger municipalities appear to make fairly successful use of both local taxes and user charges, smaller municipalities rely to a much larger extent on transfers. They tend to be poorer, which entails a smaller revenue base, as well as being characterised by larger infrastructural backlogs, which means limited revenue can be generated through service provision. These factors tend to argue against requiring greater revenue effort than that exhibited at present.

The fact that household poverty is a key component of determining LES amounts can act to reduce revenue effort and can also create an incentive against poverty alleviation initiatives. In the long-term, what is desirable is that local governments make full use of the tax bases at their disposal. In the short- and medium-term this is bound to be a painful adjustment process as local revenue efforts in the municipality are sharpened and the municipality's residents come to bear more of the burden of their own development.

The problem is, however, that no clear sense seems to exist in South Africa currently of how to assess the local tax gap (i.e. revenue collected as a share of revenue due). Therefore little information also exists on the extent to which perverse incentives against increased local revenue effort currently prevail. Further research seems warranted in this regard.

### iii. Mobility and Local Redistribution

The constraints against local redistribution have been discussed in this study and relate primarily to the fact that mobility between local jurisdictions is greater than that between countries. Therefore it is often assumed, following Tiebout (1956) that household migration patterns between jurisdictions may be sensitive to policy differences between jurisdictions. If this assumption holds, then aggressive local redistribution policies would result in a concentration of poor households in some municipalities and of rich ones in others. Redistributive municipalities

would in fact become increasingly powerless to effect such policies as their tax base migrates and average household income decreases as a result of in-migration of poorer households. It may be argued that property prices will adjust so as to induce reverse migration, but such adjustment may take quite long and it is unlikely that property prices are a major determinant of migration trends.

To identify a mobility constraint on 'aggressive' local redistribution, however, does not mean that a *degree* of redistribution through the municipal budget would have the same effect, for at least two reasons. Firstly, it is suggested that whilst households may consider municipal policies in making decisions to move or stay put, these are unlikely to play a large part in such decisions, at least below a threshold value. In other words, a mild degree of redistribution in one municipality, when compared to no such redistribution in another municipality, is unlikely to significantly induce major migration, simply because many households are unlikely to expend the effort needed to acquire this kind of information. In other words, high transaction costs may be a deterrent. This consideration does however apply only to redistribution below a 'threshold incidence': there must in principle be a point at which aggressive redistribution is so clearly recognisable in its effects that households who have been burdened with its financing decide that the benefits of relocation now outweigh the costs. It is unlikely that, in the current South African system, any municipality would engage in such active redistribution that this local redistribution 'tolerance' is breached. The magnitude of such a tolerance, as well as the broader question of the sensitivity of households to differential incidence of local spending and taxation between jurisdictions, would however be an interesting question for further analysis and empirical study.

A second reason why local redistribution may well be more feasible than is suggested by those who argue that this could induce nullifying migratory patterns concerns the fact that household utility curves may to some extent incorporate positive externalities associated with a degree of local redistribution. If richer households are willing to trade some income for social benefits (or removal of

social costs) associated with reduced poverty and infrastructural backlogs in their jurisdiction, then the research question which policy-makers need answers to concerns the *extent* to which such considerations feature in the utility curves of richer households.

In one sense the South African system already offers a clear challenge to the 'mobility hypothesis'. Both the provincial equitable share (PES) and the local equitable share (LES) are explicitly redistributionary. Although they do not redistribute income between households, they do redistribute between jurisdictions and aim at a greater convergence of the quantity and quality of public services available in various jurisdictions. It is possible that these forms of inter-jurisdictional redistribution have had little if any impact on household migration trends, and therefore that other factors than mild redistribution determine such trends.

Further research with a view to establishing how sensitive households are to municipal finances and their impact on location decisions, as well as the extent to which rich households tend to incorporate positive externality effects into their utility curves, should contribute to a clearer understanding as to whether local redistribution is *possible* to a greater extent than strict adherence to the 'mobility constraint' suggests.

Such investigation does not, however, provide a reason *why* local redistribution should be pursued in a country such as South Africa, where national government spending and taxation is already redistributionary.<sup>95</sup> In fact, if the basic premise is accepted that the efficiency gains of fiscal decentralisation stem from the greater ability of such a system to customise policies and provision to local preferences, then it becomes an intriguing question (briefly alluded to in chapter 6 in the context of free service provision) whether greater local redistribution might effect an increase in overall welfare. This possibility arises since municipalities would

---

<sup>95</sup> Cf. Van der Berg (2005).

provide a customised redistributionary budget as opposed to nationally standardised redistribution. The case for local redistribution may arguably be stronger if municipalities have a fairly clear sense of the redistribution tolerance of richer households, and could customise tax and spending accordingly, whilst national government could only offer a standard mix through national tax policy.

#### iv. Targeting, Access and Affordability

The issue of targeting and the relative merits of a universal and selective approach were discussed previously. It was emphasised in the discussion that targeted support in the form of consumption subsidies, whilst benefiting poorer households, would not benefit those poorest households who were not yet connected to the required infrastructure. In more affluent jurisdictions this would not represent a large number of households. In poorer jurisdictions, however, the consumption subsidy could well have a fairly regressive incidence given the large numbers of households who are not in a position to benefit. These considerations serve to emphasise the need to ensure infrastructural provision as a necessary step not only to service consumption paid for through user charges, but also to secure for households the equity-benefits associated with access to a free basic consumption amount.

But even where this is the case, it remains difficult to assess the extent to which the problem of *affordability*, as opposed to mere access, is a serious one in the current environment. Access is fairly easy to determine. Affordability, on the other hand, is an inherently normative notion. As the discussion in chapter 6 emphasised, payments for water services in South Africa are slightly regressive, but compare favourably with some international benchmarks in terms of their share of household expenditure. However, an issue which stems from this analysis concerns the extent to which these findings should be adjusted for differences in service backlogs between population groups. Infrastructural backlogs are concentrated amongst black households. The fact that black

households, on average, spend a larger percentage of income on water services than white households suggests, in fact, that paying black households pay substantially more, since this average includes a significant percentage of households who don't pay because they aren't connected. It would be an issue for further research to determine, in fact, whether the water services payments environment still compares favourably with international benchmarks if these factors are taken into consideration.

#### v. Basic Local Services and the Trade-Off between Efficiency and Equity

The vital and also most difficult question, which underlies virtually all the debates presented in preceding chapters, concerns the relationship between efficiency and equity considerations in the current local services and intergovernmental fiscal relations environment. It is an economic truism that greater equity comes at a cost. The unconditional pursuit of equitability in local service access between households and between jurisdictions would be likely to exact an efficiency cost which would in fact, over the medium- and long-term, reduce the resources available for redistribution as households alter their economic behaviour. It is also recognised, though, that efficiency too comes with a cost, in that the strict application of efficiency criteria, pre-eminently through efficiency-based pricing as discussed in chapter 4, would have a negative impact on equity. This would, in all likelihood, reduce the efficiency of the economy through factors such as reduced social capital, increased crime, and the like.

Thus, the fundamental debate in this regard concerns the trade-off between efficiency and equity. As with similar debates at the level of national policy, it is virtually impossible, in the local context, to determine what an optimal mix would consist of. Such a question is inherently a complicated and sometimes messy mix of normative and positive issues, of politics and economics. Thus, this study has tended to emphasise one aspect of the broader question, namely the extent to which the current policy and resource environment is orientated towards

ensuring that the minimum obligations imposed on government by the Constitution are being met. Even such a question, however, consists of highly normative questions such as the appropriate content of a package of lifeline 'basic services.' A useful question, with which this study concludes, is to ask what further research might promote a more rigorous treatment of the efficiency and equity dimensions of local infrastructural services.

One missing ingredient in this regard appears to be the absence of local incidence studies. The challenges of such studies are vast and may be close to insurmountable at a local level given the data available in South Africa at present. An associated problem is the differing circumstances and policies of municipalities. However, a clearer sense of the incidence of local finances in a few case study municipalities should add value. Such work would be useful if only for the recommendations regarding ways to ensure a better such study in future! The results of such work would be especially interesting if read together with current estimates of national fiscal incidence, such as Van der Berg (2005). The Van der Berg study emphasises social spending occurring in the provincial sphere. Similarly, a local study might limit itself to drawing some incidence conclusions for utility services such as water and electricity.

It is suggested that such attempts at quantitative results pertaining to the local sphere are necessary because they would add a degree of concreteness to discussions on the efficiency-equity trade-off. In the absence of such results many of the most interesting debates about local finances remain more conjectural than they need be. Thus, for example, the question of local redistribution in South Africa has been discussed at a few points in this study, but without any reference to any results on its *current* extent. In the absence of such information, even in a rudimentary form, research can only note general possibilities and probable outcomes, without incorporating quantitative indicators of the South African experience. The rigour of debate on local government



finances, including the ability to select the best policy alternatives through comparison of likely costs and benefits, suffers as a result.

# Bibliography

## a. Books and Papers

Ajam, T. 2001. 'Intergovernmental Fiscal Relations in South Africa'. In Levy, N & Tapscott, C (eds). *Intergovernmental Relations in South Africa: The Challenges of Co-operative Government*. 125-142. Cape Town: IDASA: Cape Town & UWC.

Arrow, KJ & Debreu, G. 1954. 'Existence of an Equilibrium for a Competitive Economy'. *Econometrica* 22 (3), 265-290.

Bahl, R. 2001. 'Fiscal Decentralization, Revenue Assignment, and the Case for the Property Tax in South Africa'. Working Paper 01-7. International Studies Program. Andrew Young School of Policy Studies. Georgia State University. Available at [www.dplg.gov.za](http://www.dplg.gov.za).

Bahl, R & Linn, J. 1992. *Urban Public Finance in Developing Countries*. Oxford: Oxford University Press...

Bahl, R & Smoke, P (eds). 2003. *Restructuring Local Government Finance in Developing Countries: Lessons from South Africa*. Cheltenham: Edward Elgar Publishing.

Bannock, G & Baxter, RE & Davis, E (eds). 2003. *The Penguin Dictionary of Economics*. London: Penguin.

Barankay, I & Lockwood, B. 2006. 'Decentralization and the productive efficiency of government: Evidence from Swiss cantons'. IZA Discussion Paper NO. 2466. Available at [www.ssrn.com](http://www.ssrn.com).

- Bird, R. 2004. 'User Charges in Local Government Finance.' Intergovernmental Fiscal Relations and Local Financial Management Program. Washington D.C.: World Bank Institute. Available at [www.worldbank.org/wbi](http://www.worldbank.org/wbi).
- Brennan, G. & Buchanan, M. 1980. *The Power to Tax: Analytical Foundations of a Fiscal Constitution*. Cambridge: Cambridge University Press.
- Canning, D & Bennathan, E. 2000. 'The Social Rate of Return on Infrastructure Investments.' World Bank Policy Research Working Paper. Available at [www.worldbank.org](http://www.worldbank.org).
- Cashdan, B. 2001. 'Local Government and Poverty in South Africa.' Municipal Services Project. Background Research Series. Available at [www.queensu.ca/msp](http://www.queensu.ca/msp).
- Deedat, H & Pape, J & Qotole, M. 2001. 'Blocked Tariffs or Blocked Access: The Greater Hermanus Water Conservation Programme.' In: McDonald, D & Bond, P (eds). *Municipal Services Project*. Occasional Papers Series No. 4. Available at [www.queensu.ca/msp](http://www.queensu.ca/msp).
- DoJCD. 1998. *The White Paper on Local Government*. Available at [www.info.gov.za](http://www.info.gov.za).
- DPLG. 2003. *Study into the Provision of Free Basic Services*. Directorate Free Basic Services. Available at [www.dplg.gov.za](http://www.dplg.gov.za).
- DPLG. 2004a. *Targeting Poor Households in the Provision of Basic Municipal Services: A Guideline for Municipalities*. Available at [www.dplg.gov.za](http://www.dplg.gov.za).

- DPLG. 2004b. *Policy Framework for the Introduction of the Municipal Infrastructure Grant (MIG), concise version*. Available at [www.dplg.gov.za](http://www.dplg.gov.za).
- DPLG. 2005. *The Municipal Infrastructure Grant: Basic Level of Services and Unit Costs: A Guide for Municipalities*. Available at [www.dplg.gov.za](http://www.dplg.gov.za).
- Du Mhango, G. 2000. 'Municipal Infrastructure Services: A Planning and Pricing Model for Capital Investment'. Chapter 6 in *Infrastructure Mandate for Change 1994-1999*. Khosa, M (ed). Cape Town: HSRC Press:. 85-107. Available at: [www.hsrc.ac.za](http://www.hsrc.ac.za).
- DWAF. 2001. *Norms and Standards in Respect of Tariffs for Water Services in terms of Section 10 (1) of the Water Services Act (Act No. 108 of 1997)*. Available at [www.dwaf.gov.za](http://www.dwaf.gov.za).
- DWAF. 2003. *Strategic Framework for Water Services*. Available at [www.dwaf.gov.za](http://www.dwaf.gov.za).
- Ebel, R & Yilmaz, S. 2001. 'Concept of Fiscal Decentralization and Worldwide Overview.' International Symposium: Quebec Commission on Fiscal Imbalance. 13, 14 September 2001. Available at [www.worldbank.org](http://www.worldbank.org).
- Fankhauser, S & Tepic, S. 2005. 'Can Poor Consumers Pay for Energy and Water? An Affordability Analysis for Transition Countries.' EBRD Working Paper 92. Available at: [www.ebrd.com](http://www.ebrd.com).
- Ferguson, P& Ferguson, G. 1998. *Industrial Economics: Issues and Perspectives, 2<sup>nd</sup> edition*. New York: Palgrave.

FFC (Financial and Fiscal Commission). 2001. *A Costed Norms Approach for the Division of Revenue: Consultation Document*. FFC: Pretoria. Available at [www.ffc.co.za](http://www.ffc.co.za).

FFC. 2006. *Review of Transfers in the Intergovernmental Fiscal Relations System in South Africa: Research Reports in Support of the FFC Submission for the Division of Revenue 2007/2008*. Josie, J & Khumalo, B & Ajam, T (eds). Available at [www.ffc.co.za](http://www.ffc.co.za).

Fjeldstad, O. 2001. 'Intergovernmental Fiscal Relations in the Developing World: A Review of Issues'. In . Levy, N & Tapscott, C (eds): *Intergovernmental Relations in South Africa: The Challenges of Co-operative Government*: 143-162. IDASA: Cape Town & UWC: Cape Town.

Fil-Flynn, M. 2001. 'The Electricity Crisis in Soweto'. In: McDonald, D & Bond, P (eds). *Municipal Services Project*. Occasional Papers Series No. 4. Available at [www.queensu.ca/msp](http://www.queensu.ca/msp).

Fourie, J. 2006. 'A Review of Economic Infrastructure in South Africa: Whither, What and Where?' Unpublished M.Comm Thesis. University of Stellenbosch.

Fox, W & Edmiston, K. 2000. 'User Charge Financing of Urban Public Services in Africa'. Working Paper 4. International Studies Program. Andrew Young School of Policy Studies. Georgia State University.

Gildenhuys, JSH.1997. *Introduction to local government finance: A South African Perspective*. Pretoria: JL van Schaik Publishers.

Gómez-Lobo, A. 2001. 'Making water affordable: Output-based

- consumption subsidies in Chile'. In: Brook, P & Smith, S (eds). *Contracting for Public Services: Output-Based Aid and its Applications*. Washington D.C.: World Bank. Available at [www.worldbank.org](http://www.worldbank.org).
- Hulten, C. 1996. 'Infrastructure Capital and Economic Growth: How well you use it may be more important than how much you have.' NBER Working Paper 5847. Available at [www.nber.org](http://www.nber.org).
- Hyman, D. 2005. *Public Finance: A Contemporary Application of Theory to Policy*. Ohio: South-Western..
- Kahn, A. 1970. *The Economics of Regulation*. Boston: MIT..
- Kuznets, S. 1955. In Meier, G & Rauch, J. 2000. *Leading Issues in Economic Development*. 382-386. Oxford: Oxford University Press.
- McClure, C & Martinez-Vasquez, J. 2003. 'The Assignment of Revenues and Expenditures in Intergovernmental Fiscal Relations'. Topic 5: Intergovernmental Fiscal Relations and Local Financial Management Program. Washington D.C.: World Bank Institute. Available at [www.worldbank.org/wbi](http://www.worldbank.org/wbi).
- Mosdell, T & Leatt, A. 2005. 'On Tap: A Review of the Free Basic Water Policy'. Children's Institute. Cape Town: UCT. Available at [www.ci.org](http://www.ci.org).
- Momoniati, I. 2001. 'Fiscal Decentralisation in South Africa: A Practitioner's Perspective'. Available at [www.dplg.gov.za](http://www.dplg.gov.za).
- Musgrave, R. 1959. *Theory of Public Finance: A Study in Public Economy*. New York: McGraw-Hill.

- National Treasury. 2005. *Budget Review 2005*. Pretoria: Communication Directorate. Available at [www.treasury.gov.za](http://www.treasury.gov.za).
- National Treasury. 2006. *Local Government Budgets and Expenditure Review 2001/02 – 2007/2008*. . Pretoria: Communication Directorate. Available at [www.treasury.gov.za](http://www.treasury.gov.za).
- National Treasury. 2007. *Budget Review 2007*. . Pretoria: Communication Directorate. Available at [www.treasury.gov.za](http://www.treasury.gov.za).
- NERSA. 2004. *National Electricity Regulator Guideline on Electricity Pricing in South African Electricity Distribution Industry*. Available at [www.ner.org](http://www.ner.org).
- NERSA. 2007'. Description of the ROR methodology'. Available at [www.ner.org](http://www.ner.org).
- Oates, W.1973.*Fiscal Federalism*. New York: Harcourt Brace Jovanovich.
- Oates, W. 1999. 'An Essay on Fiscal Federalism'. *Journal of Economic Literature* XXXVII (September 1999). 1120-1149.
- OECD Secretariat. 2005. *A Policy Framework for Investment: Infrastructure and Financial Services*. Background Document in support of OECD conference: 'Investment for Development: Making it Happen'. Available at [www.oecd.org](http://www.oecd.org).
- Prud'Homme, R.1995. 'The Dangers of Decentralization.' *The World Bank Research Observer* 10 (2). 201-220.
- Rawls, J.1999. *A Theory of Justice, Revised Edition*. Harvard: Belknap Press.
- Rosen, H. 2005. *Public Finance, 7<sup>th</sup> edition*. Singapore:Irwin/McGraw-Hill.

RSA (Republic of South Africa). 1994. Pretoria: *White Paper on the RDP*.

Shah, A.1999. 'Balance, Accountability and Responsiveness: Lessons about Decentralization.' Available at [www.worldbank.org](http://www.worldbank.org).

Shah, A.2004. 'Fiscal Decentralization in Developing and Transition Economies: Progress, Problems and the Promise'. World Bank Policy Working Paper 3282. Available at [www.worldbank.org](http://www.worldbank.org).

Statistics SA. 2000. *Income and Expenditure Survey 2000*. Available at [www.statssa.gov.za](http://www.statssa.gov.za).

Statistics SA. 2005a. *General Household Survey 2005* .Available at [www.statssa.gov.za](http://www.statssa.gov.za).

Statistics SA. 2005b. *Non-Financial Census of Municipalities 2005*. Available at [www.statssa.gov.za](http://www.statssa.gov.za).

Stein, JL.1982. *Monetarist, Keynesian and New Classical Economics*. Oxford: Basil Blackwell.

Stiglitz, J. 1982. 'The Theory of Local Public Goods Twenty-Five Years After Tiebout: A Perspective.' NBER Working Paper 954. Cambridge, Massachusetts: NBER.

Terreblanche, S. 2002. *A History of Inequality in South Africa, 1652-2002*. Sandton: University of Natal Press: Scottsville and KMM Review Publishing.



Tiebout, C. 1956. 'A Pure Theory of Local Expenditures.' *The Journal of Political Economy* 64 (5). 416-424.

Van der Berg, S. 2005. *Fiscal Expenditure Incidence in South Africa ,1995 and 2000: A Report for the National Treasury*. Cape Town: University of Stellenbosch.

World Bank. 1994. *World Development Report 1994 :Infrastructure for Development*. World Bank: Washington D.C.

World Bank. 2006. *Water, Electricity and the Poor :Who Benefits from Utility Subsidies*. Washington D.C.: World Bank.

## **b. Legislation**

Constitution of the Republic of South Africa, Act No. 108 of 1996

Division of Revenue 2007, Act No 1 of 2007

Intergovernmental Fiscal Relations Act, Act No 97 of 1997

Municipal Finance Management Act, Act No. 56 of 2003

Municipal Property Rates Act, Act No 6 of 2004

Municipal Systems Act, Act No 32 of 2000

Public Finance Management Act, Act No 1 of 1999