# CHAPTER THREE: METHODOLOGY

## 3.1 Introduction

This chapter presents the theoretical framework and the model employed in this study. It is divided into six sections. Section 3.2 discusses the theoretical framework used in this study. Section 3.3 deals with a discussion of the factors that determine the structure of the public budget and presents a model to be estimated. Section 3.4 discusses the diagnostic tests that are conducted on the reported results. In Section 3.5 the estimation procedure employed in this study is discussed while Section 3.6 contains information on the data type and sources.

## 3.2 Analytical Framework

This study will use a modified and extended version of the model developed by Hewitt (1992, 1993). The original model was used to analyse the determinants of government spending on the military by identifying government spending categories as being military spending and non-military spending. In effect it adopted a public choice framework for analysing the relationship between military spending and overall government spending.

The current study extends this model by applying this framework to analyse any type of spending by partitioning government spending into the various functional categories of government spending and extending it to include debt accumulation. Accordingly, the relationship between corruption and functional spending is modeled as follows:

Let government spending, G, be a composite of the functional spending category, $g_i$, such as military spending, health spending or education spending and 'other' spending $g_j$[3], such that:

$$G = g_i + g_j \qquad 1$$

In this study the model developed by Hewitt (1992, 1993) and used by Sanjeev *et al.* (2001) is expanded on by recognising the fact that government spending is financed through taxation and through borrowing. For ease of exposition, no distinction is made between domestic and foreign debt. This suggests that the government budget constraint in period t (t=0,1) can be approximated (Beetsma & Bovenberg 1999, 2002) as:

$$G_1 = T_1 + [(1+r)d_0 + d_1] + k\pi \qquad 2a$$

where $G_1$ is the government spending in period 1 and $k\pi$ is the seigniorage revenue. Debt at time t=1 is stated as $D_1 = d_1 + (1+r)d_0$, where D is the accumulated debt, which is the sum of the debt accumulated in the current period (d1) plus the debt of the previous period together with the interest thereon. If we exclude seigniorage, $k\pi$, the government budget constraint is approximated as:

$$G_1 = T_1 + D_1 \qquad 2b$$

We also assume that the tax function, T, is stated as follows:

$$T_1 = \tau Y_1 \qquad 0 \le \tau \le 1 \qquad 3$$

In order to maximise the welfare function we assume that it follows a utility function expressed as a Cobb-Douglas utility function:

---

[3] In this case, $g_j$ is the total spending outlay less the spending on the $g_i^{th}$ category.

$$U(C, g_i, g_j) = C^{\beta} g_i^{\gamma} g_j^{\delta} \qquad 4$$

This utility function is assumed to be twice-continuously differentiable on private consumption (C) and government spending (G), with $U_f$ >0 and $U_{ff}$ <0 for $f = C, G$, where $\delta = 1 - \beta - \gamma$. Finally, for simplicity, we assume no private investment and also omit time indices for notational simplicity. The corruption free model is founded on the conventional utility maximisation problem stated as:

Max $U(C, g_i, g_j) = C^{\beta} g_i^{\gamma} g_j^{\delta}$ 5

Subject to

$Y = C + G$ and $G = g_i + g_j$ 6

The optimal values of the above problem, which in this case is regarded as a 'corruption free' optimal solution, require[4]:

$$\frac{g_i}{Y} = \frac{\gamma}{\beta}(1-\tau) - \tfrac{\gamma}{\beta}[\frac{d_1}{Y} + \frac{(1+r)d_0}{Y}] \text{ , and } \frac{g_i}{G} = \frac{\gamma}{\beta}(1-\gamma)\frac{Y}{G} - \frac{\gamma}{\beta}[\frac{d_1}{G} + (1+r)\frac{d_0}{G}] \qquad 7a$$

and

$$\frac{g_j}{Y} = \frac{\delta}{\beta}(1-\tau) - \frac{\delta}{\beta}[\frac{d_1}{Y} + \frac{(1+r)d_0}{Y}] \text{ , and } \frac{g_j}{G} = \frac{\delta}{\beta}(1-\gamma)\frac{Y}{G} - \frac{\delta}{\beta}[\frac{d_1}{G} + \frac{(1+r)d_0}{G}] \qquad 7b$$

In equation 7a, if tax rate $\tau$ is given, the share of spending category $g_i$, in income and total government spending, depends on the parameters of the utility functions $\gamma$ and $\beta$. Similarly, in equation 7b, for a given level of tax rate $\tau$, the proportion of the 'other' spending category, $g_j$, to income and total government expenditure depends on the parameters of the utility functions $\delta$ and $\beta$. This, therefore, suggests that a lower $\gamma$ relative to $\beta$, leads to a decrease in $g_i$

[4] The way these optimal solutions have been obtained is available in Appendix A3.1.

relative to private consumption. The same is true for a lower value of $\delta$ relative to $\beta$, which also leads to a decrease in $g_j$ relative to private consumption.

As shown in the model used by Sanjeev, *et al.* (2001), the effect of corruption on the structure of the public budget can be studied via its effect on the parameters of equations 7a and 7b. In this regard, the association between corruption and a specific functional spending category is described as follows: Let the parameters of the utility function $\gamma$, $\beta$ and $\delta$ be affected by corruption Z such that equation 7a and 7b become:

$$\frac{g_i}{Y} = \frac{\gamma(Z)}{\beta(Z)}(1-\tau) - \frac{\gamma(Z)}{\beta(Z)}[\frac{d_1}{Y} + \frac{(1+r)d_0}{Y}] \qquad \text{8a}$$

and

$$\frac{g_i}{G} = \frac{\gamma(Z)}{\beta(Z)}(1-\gamma)\frac{Y}{G} - \frac{\gamma(Z)}{\beta(Z)}[\frac{d_1}{G} + (1+r)\frac{d_0}{G}] \qquad \text{8b}$$

and

$$\frac{g_j}{Y} = \frac{\delta(Z)}{\beta(Z)}(1-\tau) - \frac{\delta(Z)}{\beta(Z)}[\frac{d_1}{Y} + \frac{(1+r)d_0}{Y}] \qquad \text{8c}$$

and

$$\frac{g_j}{G} = \frac{\delta(Z)}{\beta(Z)}(1-\gamma)\frac{Y}{G} - \frac{\delta(Z)}{\beta(Z)}[\frac{d_1}{G} + \frac{(1+r)d_0}{G}] \qquad \text{8d}$$

Differentiating equations 8a and 8b, with respect to corruption, Z, yield:

$$\frac{\partial({}^{g_i}\!/\!_{Y})}{\partial Z} = (1-\tau)\left[\frac{\gamma_Z\beta - \beta_Z\gamma}{\beta^2}\right] - [\frac{d_1}{Y} + \frac{(1+r)d_0}{Y}]\left[\frac{\gamma_Z\beta - \beta_Z\gamma}{\beta^2}\right] \qquad \text{9a}$$

and

$$\frac{\partial({}^{g_i}\!/\!_{G})}{\partial Z} = (1-\tau)\frac{Y}{G}\left[\frac{\gamma_Z\beta - \beta_Z\gamma}{\beta^2}\right] - [\frac{d_1}{G} + \frac{(1+r)d_0}{G}]\left[\frac{\gamma_Z\beta - \beta_Z\gamma}{\beta^2}\right] \qquad \text{9b}$$

and differentiating equations 8c and 8d, yield:

$$\frac{\partial(g_i/Y)}{\partial Z} = (1-\tau)\left[\frac{\delta_Z \beta - \beta_Z \delta}{\beta^2}\right] - [\frac{d_1}{Y} + \frac{(1+r)d_0}{Y}]\left[\frac{\delta_Z \beta - \beta_Z \delta}{\beta^2}\right] \qquad 9c$$

and

$$\frac{\partial(g_i/G)}{\partial Z} = (1-\tau)\frac{Y}{G}\left[\frac{\delta_R \beta - \beta_R \delta}{\beta^2}\right] - [\frac{d_1}{G} + \frac{(1+r)d_0}{G}]\left[\frac{\delta_R \beta - \beta_R \delta}{\beta^2}\right] \qquad 9d$$

where $\gamma_Z = \frac{d\gamma}{dZ}$, $\beta_Z = \frac{d\beta}{dZ}$ and $\delta_Z = \frac{d\delta}{dZ}$. In this case,

$\frac{\partial(g_i/G)}{\partial Z} > 0$ and $\frac{\partial(g_i/Y)}{\partial Z} > 0$ if $\frac{\gamma_Z}{\gamma} > \frac{\beta_Z}{\beta}$.

and

$\frac{\partial(g_j/G)}{\partial Z} > 0$ and $\frac{\partial(g_j/Y)}{\partial Z} > 0$ if $\frac{\delta_Z}{\delta} > \frac{\beta_Z}{\beta}$.

In view of the above, it can be seen that corruption affects the parameters in the utility function causing a higher $g_i$ spending category as long as the utility maximiser perceives an increase this expenditure outlay as an opportunity to use public spending for private benefit[5]. In light of the above:

$$\frac{g_i}{Y} = f_1(\beta,\gamma,\tau,Z,\frac{D}{Y}) \text{ and } \frac{g_i}{G} = f_2(\beta,\gamma,\tau,Z,\frac{Y}{G},\frac{D}{G}). \qquad 10$$

Because $\beta,\gamma,\tau$ and Z are not directly observable, the impact of corruption on the $g_i$ spending category can be estimated as follows:

[5] As noted by Tanzi (1998), the simplest and most popular definition of corruption is that it is the abuse of public power for private benefit. The abuse of public power is not necessarily for one's private benefit but may be for the benefit of one's party, class, friends, family and so on. In fact, in many countries some of the proceeds of corruption go towards financing the activities of the political parties.

$$\left(\frac{g_i}{Y}\right)_{jt} = \lambda_0 + \lambda_1 Z_{jt} + \lambda_2 \left(\frac{D}{Y}\right)_{jt} + \lambda_3 K_{jt} + \varepsilon_{jt} \qquad \text{11a}$$

and

$$\left(\frac{g_i}{G}\right)_{jt} = \varpi_0 + \varpi_1 Z_{jt} + \varpi_2 \left(\frac{G}{Y}\right)_{jt} + \varpi_3 \left(\frac{D}{G}\right)_{jt} + \varpi_4 K_{jt} + \varepsilon_{jt} \qquad \text{11b}$$

t is a time index and j indexes the countries in the panel, $(\frac{g_i}{Y})_{jt}$ is the ratio of the $g_i$ spending category to the GDP, $(\frac{g_i}{G})_{jt}$ is the ratio of the $g_i$ spending category to total government spending, $(\frac{G}{Y})_{jt}$ is the ratio of government spending to the GDP, $Z_{jt}$ is a corruption indicator, $(\frac{D}{G})_{jt}$ and $(\frac{D}{Y})_{jt}$ are, respectively, the ratios of the public debt to the total public budget and the GDP, $K_{jt}$ is a vector of the state variables which are discussed in detail in Section 3.2 and $\varepsilon_{jt}$ is the error term.

By estimating equations 11a and 11b, a link can be created between the various components of the public budget as a share of the total public budget and of GDP. Equation 12 enables us to understand the role of corruption in the size of the public budget:

$$\frac{G}{GDP} = \frac{g_i/GDP}{g_i/G} \qquad 12$$

where $g_i/GDP$ is the proportion of the GDP allocated to each public economic function, $g_i/G$ is the share of the i^th economic category in the total public budget and $G/GDP$ is the total government spending's share of the GDP.

Equation 12 shows that the relationship of the total public budget to the GDP can be understood better by focusing on the shares of the components of the budget as shares of the total public budget and of the GDP. In which case, an increase in a component of the public budget as a share of the GDP accompanied by an increase of the same component as a share of the total public budget will unambiguously increase the amount of the public budget. In the literature (see Delavallade, 2006), the effect that corruption on the size of the total public budget is referred to as the quantity effect of corruption, while the effect of corruption on the distribution of the public budget is referred to as the allocation effect.

Using Equation 12, inferences are possible regarding the effect of corruption on overall government expenditure in relation to the GDP. Information concerning functional government expenditure as a share of total expenditure and of the GDP is useful in making inferences about the overall effect of corruption on the total budget as a share of the GDP. In this regard if the estimated coefficients of corruption are positive when the dependent variable is expressed as a share of the total expenditure and the GDP, it implies that the effect of corruption on functional expenditure will unambiguously lead to significant increases in total government expenditure relative to the size of the economy (GDP). However, if the estimated coefficients yield mixed signs they are bound to be either indeterminate or insignificant regarding the effect of corruption on the overall budget (Delavallade, 2006).

## 3.3 The Model

### 3.3.1 Selected factors that impact on the composition of budget spending

In terms of the determinants, there exists a pool of literature, as outlined in Chapter 2, which seeks to explain the composition of government spending. In this section we briefly discuss the factors identified in the literature that explain some or all of the functional classifications identified in Table A1.3 in the appendix.

#### 3.3.1.1 Level of corruption

As suggested by Krueger (1974), large bribes are likely to be available in conjunction with items produced by firms operating in markets where the degree of competition is low. Further more, the illegal nature of corruption and the ensuing need for secrecy imply that corrupt officials will choose goods whose exact value is difficult to monitor (Mauro 1998). It is, therefore, expected that corrupt regimes will have a tendency to tilt their budget towards sectors that procure goods and services that are specialised and have high, up-to-date technology content. Therefore, as suggested by Mauro (1998), corrupt politicians may be expected to spend more public resources on those items on which it is easier to levy large bribes while keeping those bribes secret. It is therefore expected that sectors such as defence and economic services be positively related to corruption. Other sectors such as education, health and social welfare are expected to be negatively related to corruption. However, the relationship between corruption and general public services spending may not be known from the outset.

#### 3.3.1.2 Political characteristics

In the literature some arguments were found stating that the internal allocation of the public budget is largely driven by the political characteristics of a country, as measured by the level of political and civil rights enjoyed by its citizens and the level of transparency of the government. As argued by Mahdavi (2004), it is generally agreed that certain civil liberties increase the degree of public participation in and scrutiny of the resource allocation process within the public sector. Economies that are characterised by dictatorships require massive support from the military to prevent attempts on the government. The military budget in non-democratic regimes may therefore reflect the government's demand for protective services (Kimenyi & Mbaku 1995), however, as the economy becomes more liberal and accountable, there is a tendency to re-allocate the budget towards those spending categories that the public would prefer. For example, as suggested by Nader (1994), as political liberties increase there appears a shift in the budget towards health and social security. Thus, it is expected that repressive regimes will tilt the budget toward defence and public services spending, while the liberal ones tilt the budget in favour of social sectors and economic services.

#### 3.3.1.3 Political instability

Political instability or threats thereof are very important in the allocation of the budget. As suggested in the literature, a country that is under constant threat of instability tends to allocate the budget in favour of those functional categories that seek to restore stability. This is the case whether the political system in the country is democratic or dictatorial. It is, therefore, expected that countries that are under threat of or are experiencing political instability will spend more of their budgets in favour of the general public services sector and of the defence sector. Less budget allocation will be made to the social and economic sectors of the economy. For example, countries that have a history of coups, social unrests, and ethnic tensions tend to spend more on the military and on administration. On the other hand, politically stable countries tend to spend more on those sectors where social returns are high, such as education, health and economic services.

#### 3.3.1.4 Public debt

In the models developed by Tabellini and Alesina (1990), debt accumulation is instrumental in the allocation of the public budget. This is supported in the literature by Mahdavi (2004), who found that external debt has an important role to play in the allocation of the public budget, increasing the shares of some sectors of the budget while starving other sectors. For example, higher levels of public debt will tend to enhance the shares of the economic services, health and education functional categories because the funds generated through external and internal loans are usually channelled to these sectors.

#### 3.3.1.5 Level of income

In what has come to be known as Wagner's law, Aldoph Wagner (1883-1953) hypothesised that government spending would increase in the course of development to a modern society. This relationship, as argued in the literature (Mahdavi 2004), reflects a greater role for the government as the economy becomes more complex and the demand for public goods and social programmes rises. On this basis therefore, we may infer certain changes in the composition of public spending as the role of the public sector changes during the long-term process of development. For example, in the early stages of development the government may be involved in virtually all the sectors of the economy, however, as the private sector develops, the government tends to withdraw from some sectors and focus instead on the provision of pure public goods. It is, therefore, expected that as income increases the public budget will be biased towards those functions that the private sector cannot efficiently provide. Such sectors include defence, public services, economic services and social welfare. It is thus expected that level of income will be positively related to these sectors while negatively related to education and health.

#### 3.3.1.6 Demographic characteristics: population, structure and urbanisation

As suggested in the existing literature and theories, the size of the population of a country, its geographical distribution, the degree of urbanisation and the structure of the population have an important role to play in the internal allocation of that country's budget. For example, as argued by Bergstrom and Goodman

(1973) the percentage of the population above 65 years of age is important in determining the structure of government spending. Following the life cycle hypothesis, persons who are over 65 years of age tend to spend a larger portion of their current income on current consumption than younger people spend, this suggests that if the demand for public goods as a proportion of total goods does not decline with age, then one would expect an older person to demand a larger quantity of public goods than a younger person with the same income and tax share.

#### 3.3.1.7 Size of government

The relative size of the government is important in determining the structure of the budget. As observed by Mahdavi (2004), the relative size of government serves to capture the effects of more cyclical factors, such as changes in the tax base and government non-tax revenues. It is also argued that the size of government is associated with factors that may impact on the composition of total spending. These factors include the level of corruption, exposure to external risks such as trade shocks and exposure to internal risks such as political instability and social conflicts.

#### 3.3.1.8 IMF-supported programmes

The central role of the IMF's fiscal policy advice to its members has largely remained that of improving the public spending mix by urging governments to transform their budgets in favour of productive spending and reduce the share of unproductive spending. Since structural adjustment is linked macroeconomic consistency framework, it is expected that countries that are implementing IMF reforms will tend to tilt their budgets in favour of the social sectors and economic services. The functional categories of general public services and defence are expected to be negatively affected by such IMF reform programmes.

### 3.3.2 Model specification

In view of the framework discussed in section 3.1 and the previous discussion, the general basic equations for the relative share of the i[th] functional category for the j[th] country at time t are stated as:

$$\left(\frac{g_i}{G}\right)_{jt} = \alpha_i + \beta_i (Lypc)_{jt} + \gamma_i (Lgov)_{jt} + \delta_i (DEM)_{jt} + \eta_i (POL)_{jt}$$
$$+ \lambda_i (Acc)_{jt} + \kappa_i (Cor) + \omega_i (IMF) + \hbar_i \left(\frac{D}{G}\right) + \mu_{ijt} \qquad 13a$$

and

$$\left(\frac{g_i}{Y}\right)_{jt} = \alpha_i + \beta_i (Lypc)_{jt} + \delta_i (DEM)_{jt} + \eta_i (POL)_{jt} + \lambda_i (Acc)_{jt} +$$
$$\kappa_i (Cor) + \omega_i (IMF) + \hbar_i \left(\frac{D}{Y}\right) + \mu_{ijt} \qquad 13b$$

Lypc is the real per capita GDP that serves as a proxy for the level of development; Lgov is the ratio of the total government spending to the GDP that measures the relative size of the government; DEM is a vector of demographic characteristics such as population, population structure, density and urbanisation; POL is the political instability index which measures the extent of any political instability in the country; Acc is the Voice and Accountability Index, which measures the extent of political and civil rights and of democracy in a country; and Cor is the corruption control index which measures the state of corruption in a country. IMF is the IMF dummy which stands proxy for the degree of reform in a country.

Equation 13a, which depicts the spending by category as a percentage of the total public budget imposes the following adding up constraint:

$$\sum_{i=1}^{7}\left(\frac{g_i}{G}\right)_{jt} = \sum_{i=1}^{7}\alpha_i + \sum_{i=1}^{7}\beta_i(GDPK)_{jt} + \sum_{i=1}^{7}\gamma_i(GEXGDP)_{jt} + \sum_{i=1}^{7}\delta_i(DEM)_{jt} + \sum_{i=1}^{7}\eta_i(POL)_{jt} + \sum_{i=1}^{7}\lambda_i(LIB)_{jt} + \sum_{i=1}^{7}\kappa_i(KOPT)_{jt} + \sum_{i=1}^{7}\omega_i(IMF)_{jt} + \sum_{i=1}^{7}\tau_i\left(\frac{D}{G}\right)_{jt} + \sum_{i=1}^{7}\mu_{jt} = 1 \qquad 14$$

Equation 14 implies that the error terms in various equations are correlated because relative spending shares in the j[th] country at time t must necessarily add up to unity. This adding up restriction has implications on the estimated parameters of Equation 14 as follows:

$$\sum_{i=1}^{7}\beta_i = \sum_{i=1}^{7}\gamma_i = \sum_{i=1}^{7}\delta_i = \sum_{i=1}^{7}\eta_i = \sum_{i=1}^{7}\lambda_i = \sum_{i=1}^{7}\kappa_i = \sum_{i=1}^{7}\omega_i = \sum_{1=1}^{7}\tau_i = 0 \qquad 15a$$

The restriction also has implications with regard to the error term, since for the j[th] country in period t, underestimation of the share of one of the spending categories is associated with overestimation of the remaining shares, the sum of the error terms from all the share equations will be zero. This is formally stated as:

$$\sum_{i=1}^{7}\mu_i = 0 \qquad 15b$$

In view of equation 15a, which imposes restrictions on the estimated parameters of the share equations and equation 15b, which gives the expected value of the errors from the share equations it then follows that:

$$\sum_{i=1}^{7}\alpha_i = 1 \qquad 15c$$

This therefore, suggests that the error terms across the share equations comprise a system of seemingly unrelated equations. As suggested by equation

15b, the sum of the error terms from the system of equations will sum to zero, which implies that since our system comprises of seven equations only six will be independently estimated and the seventh one will be recovered by using the restrictions suggested above.

### 3.4 Diagnostic tests

To test the validity of the estimated models in Chapters 4-11, a battery of diagnostic tests is required, the tests are discussed in this section.

#### 3.4.1 Testing the joint validity of fixed effects

Before reporting the estimation results on panel econometrics, the joint validity of fixed effects needs to be tested. This test is conducted to decide whether or not the cross sections that are included in the study can be pooled. Traditionally, panel estimation involved pooling all the members of the cross section and then estimating them. Concern regarding whether or not the members of this panel had similar enough characteristics to warrant estimation as a pool occasioned this test. In panel econometrics, literature testing for suitability for pooling or, the validity of the fixed effects model, is based on the F-statistics. In this regard the null hypothesis to be tested is that each individual cross section is not unique and, therefore, the members of the panel can be pooled together. The alternative hypothesis is that the individual members of the panel have unique characteristics and are therefore cannot be pooled. The F-test conducted in this exercise uses the residual sum of the squares of the restricted model (pooled model) and those from the unrestricted model (fixed effects model)[6]. The null and alternative hypotheses are formally stated as:

[6] Where N is small we use the Least Squares Dummy Variable (LSDV) estimation results. It is important to note that while N stands for number of cross sections and T stands for time periods, if one is interested in treating time periods as cross sections, then the F test can be adjusted appropriately by interchanging N and T.

$$H_0 : \mu_1 = \mu_2 = ........ = \mu_{N-1} = 0 \qquad 16$$

$$H_A : \mu_1 \neq \mu_2 \neq ........ \neq \mu_{N-1} \neq 0$$

and the F-test statistic is given as:

$$F = \frac{(RRSS - URSS)/(N-1)}{URSS/(NT - N - K)} \sim F_{(N-1),(NT-N-K)} \qquad 17$$

Where RRSS is the restricted sum of squares and URSS is the unrestricted sum of squares. N is the number of cross-sections, T is the number of time periods and K is the number of parameters to be estimated.

The null hypothesis of no individual effects (suitability for pooling) is accepted when the test statistic is less than the appropriate critical value. Accepting the null hypothesis admits estimation of a pooled model and rejection of the null hypothesis leads to estimation of the fixed effects model.

### 3.4.2 Testing for random effects

Testing for random effects is also conducted to establish whether or not there are individual random effects that must be taken care of rather than estimating the model using a pooled or fixed effect approach. The test for random effects is conducted using the Lagrange Multiplier (LM) test. This test is conducted by first estimating the restricted model (pooled model), obtaining the residual sum of squares, which is then utilized in the LM test statistic. In this case the null hypothesis is that there are no random effects and the alternative is that there are random effects. This is formally stated as:

$$H_0 : \sigma_\mu^2 = 0$$

$$H_A : \sigma_\mu^2 \neq 0 \qquad 18$$

and the LM test statistic is stated as:

$$LM = \frac{NT}{2(T-1)}\left[\frac{\sum_{n=1}^{N}\left[\sum_{t=1}^{T} e_{it}\right]^2}{\sum_{n=1}^{N}\sum e_{it}^2} - 1\right]^2 \sim \chi^2(1) \qquad 19$$

The LM value obtained from the above expression is compared with a chi-square with 1 degree of freedom. A test statistic value less than the critical chi-square with one degree of freedom leads to acceptance of the null hypothesis, thus admitting estimation of a pooled model, rejecting the null hypothesis leads to estimation of the random effects model.

### 3.4.3 The choice between a fixed effects model and a random effects model

Section 3.4.1 explains how a decision is taken when one is confronted with a choice between a pooled model and a fixed effects model while Section 3.4.2 explains the decision with regard to choosing between a pooled model and a random effects model. In a situation where the pooled model is not accepted, in Section 3.4.1 a fixed effects model is preferred while a random effects model is preferred in Section 3.4.2. If the results in Sections 3.4.1 and 3.4.2 prefer fixed effects and random effects models respectively, a decision has to be taken regarding the model to use. As suggested by Baltagi (2001) there is controversy in the literature regarding the appropriate model to use. In this regard a test suggested by Hausman (1978), which is based on the difference between the fixed and random effects estimators, is utilised to identify the preferred model.

## 3.5 Estimation procedure

The estimation is done within the panel econometrics framework at two levels. Firstly, equations 13a and 13b are estimated in a manner that allows each of the spending categories to be estimated independently of the other categories. In so doing, I am able to allow for the inclusion of the determinants that are specific to each of the spending categories both as a share of the total public budget and of the GDP.

Secondly, equation 13a is estimated in a manner consistent with the systems estimation and the results are reported in a separate chapter. In view of the expected simultaneity in Equation 13a, estimations will be conducted within a system panel framework in a manner consistent with the Seemingly Unrelated Regression (SUR) estimation method proposed by Zellner (1962).

## 3.6 Description of the data

### 3.6.1 Data sources and type

This study seeks to investigate the spending behaviour of African governments for the set of countries listed in Appendix A1.2 over the period 1995-2004. The variables used in this study are as follows:

- The general public service, defence, education, health, social welfare, economic services, and 'other' spending sectors are the dependent variables. They are expressed as a ratio of the total public budget and of the GDP in the various estimations. The data is obtained from IMF country reports which are obtained from www.imf.org/external/country/index.htm.
- The variable, the 'defence spending of neighbouring countries' is obtained by finding the average defence spending of a country's neighbours. It is expressed as a ratio of the total public budget and of the GDP. The data is obtained from IMF country reports which are obtained from www.imf.org/external/country/index.htm
- The corruption control index, political stability index, and voice and accountability index are the governance indicators used in this study to proxy, respectively, corruption, political stability and freedom and

transparency of government. These indicators are obtained from the World Bank at www.worldbank.org/wbi/governance/puts/govmatters+html.

- The variables, 'population size', 'population density' and 'urbanisation rates' have been obtained from the World Bank: African Development Indicators, various issues.
- The variables, 'size of government', 'level of income' (GDP) and 'external debt' have been obtained from the World Bank: African Development Indicators, various issues.
- The IMF dummy is constructed on the basis of information available from the IMF. A country is assigned a value of 1 if IMF programmes have been implemented in a given year and a value of 0 if not. Data for constructing this dummy is obtained from www.imf.org/external/country/index.htm
- The variable 'military personnel per 1000 people' is the proxy for the staffing levels of the defence department. The data is obtained from the Stockholm International Peace Research Institute (SIPRI).

### 3.6.2 Choice of governance indicators

A number of organisations produce governance data using different methodologies and for diverse reasons; these data sets have been used by various studies touching on governance matters. In this study we use the World Bank data set for the following reasons: firstly, the World Bank data set is obtained from the data collected by 31 firms that construct governance indicators which makes it a hybrid index encompassing all the attributes of these individual indicators; secondly, individual firms use different methodologies to construct their indices for different uses. The World Bank data set performs better in this regard because it is drawn from many different sources which makes it more reliable overall; thirdly, the number of countries and territories used by individual firms is less than the number available from the World Bank database. For example, the Political & Economic Risk Consultancy uses only 10 countries, the Political Risk Services uses 140 countries and Afrobarometer uses 18 countries.

Furthermore, there is no guarantee that the countries covered in these individual data sets include the countries of interest in this study; fourthly, while some firms report data for many countries they only started doing so recently, these data sets are therefore not suitable for this study which requires a long time series.. For example, IJET Travel International covers 167 countries, but only for 2004. Using data from these individual sources would have reduced the number of countries in this study because, besides being limited in terms of coverage they are also limited in terms of the periods covered; fifthly, compared to, for example the Corruption Perception Index (CPI) of Transparency International (TI), the World Bank data set is superior because it does not use lagged data when current data is not available, which is what TI does when constructing the CPI (Kaufmann, Kraay and Mastruzzi. 2006).

The governance indicators developed by the World Bank are constructed using data from 31 different sources as shown in Appendix A1.4. From these sources the World Bank constructs six categories of governance indicators for 213 countries and territories: voice and accountability; political stability and absence of violence; control of corruption; rule of law; effectiveness of government; and regulation quality. In this study the first three are of interest and are discussed in detail below.

#### 3.6.2.1 Corruption control index

The corruption control index is a proxy for the level of corruption. It captures the extent to which public power is exercised for private gain, including both petty and grand forms of corruption, as well as a ‘capture’ of the state of elites and their private interests. It is constructed in such a way that a country which demonstrates the least effort in the fight against corruption (and therefore a higher level of corruption) is assigned a value of -2.5 while one showing a greater effort in fighting corruption (and therefore a lower level of corruption) is assigned a value of +2.5.

#### 3.6.2.2 Political stability index and absence of violence

The political stability index is a proxy for the level of political stability in a country. It measures perceptions of the likelihood that the government will be destabilised or overthrown by unconstitutional or violent means, including political violence and terrorism. The political stability index is constructed in such way that a country which is most politically unstable is assigned a value of -2.5 while one that is politically stable is assigned a value of +2.5.

#### 3.6.2.3 Voice and accountability index

The voice and accountability index measures the extent to which a country's citizens are able to participate in selecting their government, as well as the level of freedom of expression, freedom of association and freedom of the media. The voice and accountability index is constructed in such way that a country which ranks poorly in voice and accountability is assigned a value of -2.5 while one that is ranked highly is assigned a value of +2.5.

## 3.7 Univariate analysis

The role of descriptive statistics is well documented in econometrics literature. Descriptive statistics show the individual characteristics of the variables that are used in the estimations. These include knowledge of the first, second and third order moments. More importantly, aspects such as skewness, kurtosis and consequently normality are exposed. Table 1 contains summary descriptive statistics of the dependent variables used in this study[7].

[7] Descriptive statistics of other variables are presented on Appendix 2.

**Table 1: Descriptive statistics of the ratio of dependent variables as ratios of the total public budget**

| | General public services | Defence | Education | Health | Social welfare services | Economic services | Other |
|---|---|---|---|---|---|---|---|
| Mean | 22.37 | 12.88 | 16.30 | 6.30 | 6.54 | 15.02 | 20.72 |
| Median | 21.46 | 9.01 | 17.71 | 5.96 | 3.84 | 12.82 | 18.75 |
| Maximum | 58.91 | 65.84 | 41.52 | 15.78 | 24.84 | 51.73 | 85.39 |
| Minimum | 3.28 | 1.53 | 1.19 | 1.00 | 0.17 | 0.68 | 0.85 |
| Std. Dev. | 10.57 | 11.33 | 6.75 | 2.79 | 6.54 | 10.88 | 15.39 |
| Skewness | 0.46 | 2.17 | -0.24 | 0.71 | 1.20 | 1.33 | 1.06 |
| Kurtosis | 2.52 | 7.88 | 2.96 | 3.40 | 3.34 | 4.55 | 4.42 |
| Jarque-Bera | 12.43 | 498.24 | 2.79 | 25.45 | 68.08 | 110.31 | 75.33 |
| Probability | 0.00 | 0.00 | 0.25 | 0.00 | 0.00 | 0.00 | 0.00 |
| Sum | 6262.37 | 3606.02 | 4563.90 | 1762.96 | 1830.00 | 4206.16 | 5801.08 |
| Sum Sq. Dev. | 31154.00 | 35810.96 | 12703.50 | 2165.12 | 11921.53 | 33035.17 | 66113.28 |
| Observations | 280.00 | 280.00 | 280.00 | 280.00 | 280.00 | 280.00 | 280.00 |
| Cross sections | 28.00 | 28.00 | 28.00 | 28.00 | 28.00 | 28.00 | 28.00 |

Table 1 shows the descriptive statistics for the dependent variables as a share of the total public budget. From the table it is evident that the mean budget allocation to the general public services functional category had the highest allocation. The highest share in this category was reported at 58.9% while the minimum was 3.3%. However, the second moments show that this variable is both skewed and has a kurtosis of 2.52, with the Jarque-Bera statistic showing that the variable is not normally distributed. The 'other' category was second highest and the second moments show that the variable is not normally distributed. The budget allocation for defence stood at 12.9%, for education at 16.3%, for health at 6.3%, for social welfare at 6.5% and for the economic services sector, at 15.02%. The other statistics show that all the variables, except the budget share for education are not normally distributed. This, therefore, suggests that these variables need transformation to approximate normality (Hamilton 1992).

**Table 2: Description statistics of dependent variables as ratios of GDP**

| | General public services | Defence | Education | Health | Social welfare | Economic services | Other |
|---|---|---|---|---|---|---|---|
| Mean | 7.72 | 4.55 | 5.34 | 2.10 | 2.71 | 4.89 | 6.47 |
| Median | 5.21 | 2.11 | 4.78 | 1.70 | 0.94 | 3.64 | 3.97 |
| Maximum | 37.51 | 37.96 | 32.59 | 13.71 | 25.86 | 29.10 | 35.68 |
| Minimum | 0.58 | 0.12 | 0.31 | 0.09 | 0.03 | 0.16 | 0.09 |
| Std. Dev. | 7.08 | 6.10 | 4.36 | 1.70 | 4.19 | 4.56 | 7.18 |
| Skewness | 1.84 | 2.98 | 2.38 | 2.40 | 2.95 | 1.97 | 1.82 |
| Kurtosis | 6.37 | 13.18 | 11.97 | 13.47 | 13.60 | 8.03 | 5.82 |
| Jarque-Bera | 291.21 | 1623.76 | 1204.68 | 1549.06 | 1717.55 | 477.28 | 246.67 |
| Probability | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Sum | 2161.37 | 1273.63 | 1495.11 | 589.24 | 759.06 | 1368.04 | 1812.55 |
| Sum Sq. Dev. | 13995.26 | 10379.75 | 5302.37 | 807.22 | 4899.87 | 5809.58 | 14368.73 |
| Observations | 280.00 | 280.00 | 280.00 | 280.00 | 280.00 | 280.00 | 280.00 |
| Cross sections | 28.00 | 28.00 | 28.00 | 28.00 | 28.00 | 28.00 | 28.00 |

Table 2 shows the dependent variables expressed as shares of the GDP. From the table it is evident that the budget share for public services accounts for 7.7%, while the health category accounts for 2.1% of the GDP. The other descriptive statistics suggest that all the variables suffer from either positive or negative skewness. The Kurtosis measurements, which show whether the variables are peaked or not and the levels thereof, show that all the variables are peaked. The Jarque-Bera statistics show that these variables are not normally distributed.

## 3.8 Summary

In this chapter the methodology used in this study has been discussed. A public choice framework is adopted which is consistent with Sanjeev *et al.* (2001) and adds value by including public debt. The functional spending categories that are used have been identified the factors that are identified in the literature that explain these functional spending categories are discussed. The chapter also discusses the econometric estimation procedures and the diagnostic tests performed on all the estimated models. Lastly, the chapter discusses the data type and sources and gives a brief overview of the descriptive statistics of the dependent variables.