Essays on the Macroeconomic Management of Foreign Aid Flows in Africa

by

Pedro Miguel Gaspar Martins

Thesis submitted for the degree of

Doctor of Philosophy
in
Economics

Institute of Development Studies
University of Sussex
Brighton, United Kingdom

April, 2010

STATEMENT

I hereby declare that this thesis has not been, and will not be, submitted in whole or in part to another University for the award of any other degree.

Pedro Miguel Gaspar Martins

____

UNIVERSITY OF SUSSEX

Pedro Miguel Gaspar Martins

Doctor of Philosophy in Economics

ESSAYS ON THE MACROECONOMIC MANAGEMENT OF FOREIGN AID FLOWS IN AFRICA

SUMMARY

The main motivation of this thesis is to contribute to the literature on the macroeconomic effects of foreign aid flows. It consists of four empirical papers, investigating the two main channels through which aid flows impact the recipient economy: (i) the fiscal sector, and (ii) the real exchange rate. The first paper is concerned with the impact of aid on government expenditure, domestic revenues and borrowing. It uses a traditional fiscal response framework with annual data for Ethiopia. The second paper also focuses on the fiscal sector but uses a recently compiled quarterly fiscal dataset and the cointegrated vector autoregression methodology. The main result arising from both papers is the strong correlation between aid inflows and domestic borrowing, possibly as a strategy to smooth unpredictable and volatile aid inflows. Aid is positively correlated with government expenditures, but there is little evidence of tax displacement. There is also evidence of aid heterogeneity, as grants and loans induce different effects. The third paper assesses the impact of foreign aid on the Ethiopian real exchange rate, which is a common measure of external competitiveness. It uses a quarterly macroeconomic dataset and applies two distinct methodologies: (i) single-equation cointegration models, and (ii) an unobserved components model. The results do not provide support for the 'Dutch disease' hypothesis. The fourth paper investigates the extent to which foreign aid is 'absorbed' and 'spent'. The empirical analysis uses a panel of 25 African low-income countries and applies recently developed panel cointegration techniques. The findings suggest that aid is fully spent while absorption is higher than previously estimated.

Acknowledgements

I would like to thank my supervisor, Dr. Ricardo Gottschalk, for his encouragement and invaluable guidance.

I am grateful to the Fundação para a Ciência e a Tecnologia for providing the financial support for this research.

I would like to thank the staff at the Ethiopian Development Research Institute for the logistical support and advice provided during my fieldwork, as well as the civil servants and academics I had the opportunity to meet in Addis Ababa. I am also grateful to Professor Terry McKinley for offering me the opportunity to visit the International Policy Centre for Inclusive Growth in 2007, where I conducted work for the first paper.

I am grateful to Dr. Dirk Willenbockel and Dr. Xavier Cirera for reading and providing comments on my thesis. I would also like to thank the participants of my three research seminars held at the Institute of Development Studies for their comments and suggestions. In particular, I wish to thank Professor Michael Sumner, Professor Sherman Robinson and Dr. Neil McCulloch for their insightful remarks.

I would also like to thank the participants of the 2008 Summer School 'The Cointegrated VAR Model: Econometric Methodology and Macroeconomic Applications' and the 2009 Summer School 'Time Series Modelling and Analysis', in particular Professor Katarina Juselius, Professor Søren Johansen, and Professor Andrew Harvey.

Finally, and probably most importantly, I would like to thank my wife for her unconditional support throughout this journey. I could not have done it without her and I am sincerely grateful for her positive energy.

Table of Contents

List of Tables

List of Figures

1. Introduction

The objective of this thesis is to contribute to the literature on the macroeconomic effects of foreign aid in African countries. A better understanding of foreign aid dynamics will enable donors and recipient governments to design better aid interventions, manage risks adequately and contain potential negative effects. Aid flows are a significant and crucial external resource for aid-dependent countries, financing socio-economic public expenditures while providing foreign exchange to support the balance of payments. However, the size and pattern of aid inflows also have the potential to destabilise recipient economies, namely through higher domestic inflation, nominal exchange rate appreciations, and perverse fiscal incentives. These macroeconomic impacts are likely to influence the effectiveness of foreign aid interventions. The challenge for policy-makers is therefore to effectively manage large inflows of foreign finance while minimising the potential negative impacts. This thesis seeks to empirically assess the occurrence and magnitude of such effects as well as examine policy responses in African countries.

The aid effectiveness literature is extensive, although the debate has often focused on the relationship between aid and economic growth. We argue, however, that a greater focus should be placed on understanding the specific mechanisms through which aid impacts on domestic economic performance. Hence, the pertinent research question seems to be *how* aid affects growth, rather than *whether* aid affects growth. With that objective in mind, this thesis investigates the two main channels through which aid influences economic performance: (i) the fiscal sector; and (ii) the real exchange rate. In order to investigate these mechanisms in more detail, we use Ethiopia as a case-study. The choice of the country was partly related to its substantial dependency on foreign resources and the availability of detailed fiscal and macroeconomic data.

Overall, the main contributions of this thesis are: (i) new insights into the fiscal dynamics of aid; (ii) an empirical investigation of Dutch disease symptoms in one of the largest aid-recipient countries in Africa; and (iii) a re-assessment of the macroeconomic responses to aid inflows for a panel of African countries. For this purpose, we construct consistent quarterly fiscal and macroeconomic datasets for Ethiopia, compile a new panel dataset for African countries, and use sophisticated empirical methodologies – namely, the cointegrated vector autoregression model, the unobserved components model, and several panel cointegration methodologies.

This thesis consists of four empirical papers that explore the potential macroeconomic effects and policy responses to an increase in aid inflows. The first (chapter 2) is concerned with the impact of aid flows on government expenditure, domestic revenues and borrowing, since these are related to interesting hypotheses about the recipient government's policy stance. We start by defining a public utility maximisation problem, from which a set of simultaneous equations is analytically derived. The estimation is performed by three-stage least squares (3SLS) with annual data for Ethiopia. The analysis is complemented by a single-equation cointegration methodology.

The second paper (chapter 3) also investigates the fiscal impact of foreign aid flows in Ethiopia. However, it uses a recently compiled quarterly fiscal dataset, which covers the period 1993Q3-2008Q2. We then apply the cointegrated vector autoregression (CVAR) methodology to analyse the fiscal dynamics in post-Derg Ethiopia. A number of interesting hypotheses are formulated and then tested. Finally, we estimate impulse response functions based on the moving average (MA) representation, and present several robustness checks to assess parameter stability.

The third paper (chapter 4) assesses whether aid flows have had a significant impact on the real exchange rate, which is a common measure of external competitiveness. We compile a macroeconomic quarterly dataset for Ethiopia, covering the period 1981Q1-2008Q1. We then use several single-equation cointegration approaches to investigate the short- and long-run determinants of the real exchange rate. To complement the analysis and provide a robustness check, a structural time series model (unobserved components) is also used.

The fourth paper (chapter 5) uses a panel of 25 African low-income countries to investigate the extent to which foreign aid resources have been 'absorbed' and 'spent'. These concepts are useful to evaluate the degree of policy coordination between recipient governments and central banks. Hence, the empirical analysis assesses whether foreign aid is correlated with increases in the non-aid current account deficit ('absorption') and the non-aid fiscal deficit ('spending'). This is done through the use of recently developed panel cointegration techniques. Moreover, the impact of aid on the accumulation of international reserves, government investment and domestic borrowing is also assessed.

The last chapter concludes this thesis by providing an overview of the empirical results and policy implications.

2. Foreign Aid and the Government Sector: The Case of Ethiopia

Abstract: The objective of this paper is to assess the impact of foreign aid on public expenditure, revenue and domestic borrowing in Ethiopia.[1] *We review the literature on the fiscal effects of aid and conclude that there are few common cross-country patterns – highlighting the important role that country-specific factors play in determining fiscal outcomes. We then apply a fiscal response model to Ethiopian annual data covering the period 1964-2005. The model disaggregates aid inflows into grants and foreign lending in order to analyse their specific (heterogeneous) impacts. The results suggest that aid is positively correlated with public investment, with loans having a stronger impact than grants. Moreover, these flows have a strong negative effect on domestic borrowing, suggesting that aid and domestic financing are close substitutes. This finding may explain the relatively weak impact of aid flows on expenditure. Finally, the evidence also appears to support the hypothesis that higher aid flows displace domestic revenues. However, this particular result does not seem to be robust across the sample. These results suggest that donors and recipients of aid should focus on ensuring greater 'aid additionality' to finance new capital spending and building up national capacities to mobilise domestic revenue.*

2.1 Introduction

The impact of foreign aid flows on government fiscal accounts is a fundamental question in the aid effectiveness debate. Since most aid inflows are provided directly to recipient governments, it is crucial to assess how these funds are allocated (e.g. investment versus consumption), the type of incentives they produce (tax effort) and the impact that they have on the fiscal deficit (debt sustainability). This is the initial and probably the most important transmission channel through which aid impacts the recipient economy. There are several hypotheses that can be tested within a fiscal response framework, which are closely linked to the following concepts:

(i) Additionality. Aid inflows are said to be additional if they entail an equivalent increase in government expenditure. Since one of the main reasons to provide aid is to finance critical investment programmes, we expect these flows to significantly contribute to increased public expenditures. However, this may not be always the case. There might be a

[1] An earlier version of this paper was published by the UNDP International Poverty Centre (now International Policy Centre for Inclusive Growth) in its Working Paper series – IPC WP41 (November 2007).

time lag between aid flows being received and the actual expenditure. This can be particularly noticeable for concessional loans (often disbursed in large instalments) and budget support grants. Moreover, aid flows may also be used for other purposes, such as retiring onerous domestic debt or reducing the tax burden. This will be discussed below.

(ii) Aid illusion. This refers to the case when public expenditure increases by more than the amount of the net aid inflow (McGillivray and Morrissey, 2001b). Since aid projects often induce extra costs that need to be financed by domestic resources (e.g. road maintenance, staff salaries of a newly-built hospital, etc.), it is conceivable that the impact of aid is more than proportional. This can be particularly concerning if it leads to increases in public debt. Moreover, aid illusion may also take place when recipients overestimate future aid inflows or donors disburse less than originally committed when expenditures are already planned (aid unpredictability).

(iii) Fungibility. The term describes situations where earmarked aid flows are indirectly used for unintended purposes. For example, if aid finances a project that would otherwise be funded by tax revenues, then the aid inflow may (in practice) release domestic resources for unproductive non-developmental spending. This may arise in cases of asymmetric information or policy disagreements between donors and recipients. One could suggest that fungibility occurs if aid inflows increase recurrent expenditure rather than capital spending, since donors are more likely to target investment expenditures. However, it should be noted that some types of recurrent expenditure have a development focus (e.g. health staff salaries, school books, etc.) and are strongly supported by donors.

(iv) Tax displacement. Higher aid inflows may be associated with a decrease in domestic revenues. This may occur for three main reasons. Firstly, foreign aid may lower the government's incentive to increase the tax effort. This is particularly problematic if the tax base is small and the recipient country is heavily dependent on aid resources. Secondly, tax revenues could be reduced due to policy reforms linked to aid flows (McGillivray and Morrissey, 2001a:32). Trade liberalisation is likely to lower government revenues if, for example, imports do not rise enough to compensate lower duty rates. Moreover, tax reforms such as the introduction of a value-added tax (VAT) may take some time to become fully operational, while its effectiveness may depend on the quality of the tax collection system. Thirdly, the government could use the extra fiscal space provided by aid flows to lower tax rates for key sectors of the economy. The virtues of this strategy need to be assessed in light of the country's specific circumstances.

(v) Deficit financing. When aid inflows are unpredictable, governments may wish to smooth public spending. This will entail borrowing domestically when aid inflows fall short of expectations, while using subsequent (higher than expected) aid resources to reduce borrowing needs. In this case, aid flows and domestic financing can be seen as substitutes and will be negatively correlated – aid is indirectly financing public spending. Moreover, aid flows may also be used to retire onerous public loans that were, for example, incurred by previous governments ('odious debt'). This could be a good strategy in countries with a heavy debt burden, but perhaps not in general.

(vi) Spending. This concept is defined by the IMF (2005) as the widening of the fiscal deficit (excluding aid) caused by an increase in aid. It should be noted that decreases in domestic revenue and increases in public expenditure may achieve the same result, even though they correspond to rather different policy stances.

(vii) Aid Heterogeneity. Aid flows can take the form of unrequited transfers (grants) and concessional loans. Further disaggregation may be possible, for example: food aid grants, budget support grants, project aid, etc. There is an increasing belief that the impact of aid is likely to depend on the modality of aid, therefore supporting the case for the disaggregation of aid in empirical exercises.

This paper is divided into five main sections. After this brief introduction, section two provides an overview of the fiscal response literature. Section three reviews Ethiopia's recent economic performance and fiscal trends. Section four presents the theoretical underpinnings of the fiscal response model, which is then applied to four decades of Ethiopian data, 1964-2005. Two versions of the model are estimated, one with total aid flows and the other with foreign aid disaggregated into grants and foreign loans. The empirical model is then complemented by cointegration analysis, as a robustness check. Section five concludes the paper.

2.2 Literature Review

There is a significant empirical literature on the fiscal impacts of aid. This section briefly reviews the discussions and findings on the topic, with particular reference to two distinct (but related) branches of the literature: (i) categorical fungibility, and (ii) fiscal response (or aggregate fungibility). McGillivray and Morrissey (2001a) provide an excellent review of this literature, which we draw from and update. Categorical fungibility studies have a common focus on the impact of aid on the composition of government spending, often with a reasonable disaggregation of expenditure items. However, McGillivray and Morrissey (2001a:11) sub-divide this strand of the literature in two main groups: (i.a) studies that start from a utility maximisation problem (using different expenditure categories) and then derive a system of linear expenditure equations to be estimated; and (i.b) studies that do not use an explicit theoretical framework but still estimate a set of simultaneous fiscal equations.

Studies belonging to the first category usually specify the representative agent's utility function in a Stone-Geary form, which is then maximised subject to a budget constraint. Khilji and Zampelli (1991), Feyzioglu et al (1998) and Swaroop et al (2000) are some of the studies that belong to this category. For example, Feyzioglu et al (1998:33) use the following utility function with S goods purchased by the government and a single private good (c_P):

$$W = U(c_P, g_1, g_1^{NF}, \dots, g_K, g_K^{NF}, g_{K+1}, \dots, g_S)$$

where g_k^{NF} is the quantity of the kth good that the government has to purchase from the non-fungible portion of the aid allocated for good k (k = 1, …, K), while g_s is the quantity of the sth good purchased from the fungible fraction of aid (s = 1, …, K, K+1, …, S), supplemented by other revenues. We can then write:

$$g_k^{NF} = \frac{(1-\phi_k)a_k}{P_k}$$

where P_k is the price of good k and $(1-\phi_k)a_k$ is the non-fungible portion of aid that is used to purchase public good k. Hence, ϕ_k measures the extent of fungibility and is a key parameter of interest. Feyzioglu et al (1998) specify the utility function in a Stone-Geary form and then maximise it subject to the following budget constraint:

$$P_1 g_1 + P_2 g_2 + \ldots + P_S g_S = R + \sum_{k=1}^{K} \phi_k a_k$$

where P_s are prices of the public goods, a_k is the amount of aid targeted for good k, and R is revenue from all other sources. Clearly, ϕ_k is the fungible portion of aid that supplements R (it varies between zero and one). Taking p_s, ϕ_k, R and a_k as given, the government chooses S goods ($g_1, g_2, ..., g_S$) to maximise the utility function. Crucially, it is only the fungible portion of aid that affects the government's choice of S goods (g_1, …, g_S). Otherwise, the government's optimal choice would be the same for any values of ϕ (Feyzioglu et al, 1998:34).[2] After deriving the equations of the simultaneous linear expenditure model, the parameters of the system can then be estimated.

Alternatively, Pack and Pack (1990, 1993), Cashel-Cordo and Craig (1990), Gupta (1993) and Pettersson (2007) do not use any explicit theoretical framework but nonetheless estimate a set of simultaneous equations. In these studies, the government has a 'community indifference curve' and is faced by a budget constraint. They then define equations that represent demand curves (hypothetically) derived from optimising decisions. This is a more ad-hoc approach, since these equations are not derived from an explicit theoretical framework. For example, Pack and Pack (1993:260) propose the following equations (demand curves) for the Dominican Republic:

$$D_{it} = f(GDP_t, FA_{it}, OFA_{it}, DUM)$$
$$FI_t = f(GDP_t, FAT_t, DUM)$$
$$C_t = f(GDP_t, FAT_t, DUM)$$
$$R_t = f(GDP_t, FAT_t, TIME)$$

where D_i represents five categories of development expenditures, *FI* stands for financial and indirect investment, *C* is total current expenditure, and *R* is total revenues (excluding aid). In terms of the explanatory variables, *GDP* is current gross domestic product, FA_i the categorical foreign aid to expenditure in category *i*, *OFA* is aid for other expenditure categories, *FAT* is total foreign aid, *DUM* is a dummy variable to capture the effects of the structural adjustment program, and *TIME* a deterministic time trend. These equations are bound by a budget constraint, and are estimated simultaneously through seemingly unrelated regressions (SUR).

[2] Since it is assumed that all aid is earmarked by purpose for the purchase of K goods ($K \leq S$), the question is whether aid will be diverted to purchase 'unintended' goods (g_1, …, g_S).

Table 2-1: Results of Selected ‘Categorical Fungibility’ Studies

Study	Extent of Fungibility (ϕ)	Incremental Impact of Aid on: Domestic Revenue	Total Spending	Develop. Expend.	Non-Develop. Expend.	Health & Education Expend.	Invest. Expend.	Consump. Expend.
Pettersson (2007)	−0.20$^{\perp}$ 0.67^{+}			0.84$^{\perp}$ 0.25^{+}		0.35$^{\perp}$ 0.12^{+}		
Swaroop et al (2000)		0.00	0.00	0.00	0.90	0.00	0.00	
Feyzioglu et al (1998)	−0.57	n/a	0.95	0.23		0.13	0.29	0.72
Pack & Pack (1993)	0.79	−0.39	−0.27	−0.05	−0.31	0.002		0.08
Gupta (1993)	0.04	0.01	1.69	0.96	0.73			
Khilji & Zampelli (1991)	1.00	−0.01	0.26		0.74			
Cashel-Cordo & Craig (1990)		10.36† 4.25‡	12.82† −2.79‡					
Pack & Pack (1990)	0.00	0.29	1.37	1.37	0.00	0.19		0.00

Obs.: Blank means not reported (or cannot be inferred); $^{\perp}$ Full sample; $^{+}$ Countries with ϕ in the range [0,1]; † African countries; ‡ non-African countries.
Source: McGillivray and Morrissey (2001a:16) and calculated averages from Pettersson (2007)

As the table above demonstrates, the empirical evidence on the fungibility of aid is mixed. The findings of Pack and Pack (1990) for Indonesia, Gupta (1993) for India and Pettersson (2007) are encouraging in the sense that aid flows seem to significantly increase development expenditures, while there is little evidence of a fall in revenues. However, Pack and Pack (1993) for the Dominican Republic and Khilji and Zampeli (1991) for Pakistan show that aid is highly fungible and has a negative impact on domestic revenues. Moreover, Pack and Pack (1993) and Swaroop et al (2000) for India provide evidence of a negative or perhaps insignificant impact of aid on developmental expenditures. There are also some atypical results, which cast some doubt on the reliability and robustness of this empirical approach. Cashel-Cordo and Craig (1990) show a surprisingly large increase in domestic revenue (as well as total expenditure) for 46 least developed countries (LDCs), while Feyzioglu et al (1998) for 14 LDCs find a negative fungibility coefficient (‘aid illusion’). Finally, Petterson (2007) estimates fungibility coefficients for 57 developing countries and reaches the conclusion that aid is significantly fungible, with an average coefficient of 0.67. However, this is only after 28 countries with coefficients outside the [0,1] range are excluded or have their values truncated.

Although these studies have had an important influence on policy debates, especially through the publication of World Bank (1998), they have some fundamental limitations. Some of these studies use panel data methods, which assume homogeneity of the aid impacts. This is an important issue since the dynamic responses to aid are likely to be different even for countries with similar degrees of economic development and political structures. More importantly, (categorical) fungibility studies are restricted to the impact on the composition of government spending, diverting attention from the more fundamental issue, which is the broader fiscal impacts of aid over time, especially on tax effort and borrowing (McGillivray and Morrissey, 2000). Indeed, these studies assume government revenue to be a residual, not allowing for aid to explicitly influence the tax effort or domestic borrowing. They thus provide few insights into the general fiscal impact of aid.

In view of the limitations detailed above, the fiscal response (aggregate fungibility) literature has emerged as a valuable empirical alternative. Although fiscal response models use government accounts in more aggregate terms, they have the advantage of having a broader scope of analysis. Among these studies, we find Heller (1975), Mosley (1987), Franco-Rodriguez et al (1998), and Mavrotas and Ouattara (2006). These works are also centred on the utility maximisation principle, although the government's utility function differs substantially from that in the fungibility literature. The utility function is defined as the deviations of the observed variables (actuals) from government targets (budgeted figures), and is often assumed to be symmetric. The utility is maximised when all the targets are met. A set of reduced form equations is then derived from the optimisation problem and estimated simultaneously, often by non-linear three-stage least squares (3SLS). Technical details will be provided at a later stage.

Table 2-2: Results of Selected 'Fiscal Response' Studies

<table>
<tr><th rowspan="2">Main Studies</th><th rowspan="2">Countries</th><th colspan="4">Impact of Aid on</th></tr>
<tr><th>Tax Revenue</th><th>Government Investment</th><th>Government Consumption</th><th>Domestic Borrowing</th></tr>
<tr><td>Machado (2009)</td><td>Nicaragua</td><td>0.17</td><td>0.47</td><td>−1.47</td><td>−3.60</td></tr>
<tr><td>Ouattara (2006a)</td><td>Senegal</td><td>−0.68</td><td colspan="2">−0.07</td><td>0.18</td></tr>
<tr><td>Ouattara (2006b)†</td><td>Developing</td><td>0.00</td><td>0.08</td><td>0.00</td><td>−0.13</td></tr>
<tr><td>McGillivray & Ouattara (2005)</td><td>Cote d'Ivoire</td><td>−0.92</td><td colspan="2">−0.1</td><td>0.2</td></tr>
<tr><td>Franco-Rodriguez (2000)</td><td>Costa Rica</td><td>0.05</td><td>−0.02</td><td>0.07</td><td>−0.08</td></tr>
<tr><td>McGillivray & Ahmed (1999)</td><td>Philippines</td><td>−0.1</td><td>−0.02</td><td>0.02</td><td>−1.81</td></tr>
<tr><td>Franco-Rodriguez et al (1998)</td><td>Pakistan</td><td>−3.6</td><td>0.1</td><td>−2.4</td><td>0.9</td></tr>
<tr><td>Iqbal (1997)</td><td>Pakistan</td><td>0.0</td><td>0.0</td><td>1.6</td><td></td></tr>
<tr><td>Rubino (1997)</td><td>Indonesia</td><td>−1.4</td><td>−0.8</td><td>−0.7</td><td></td></tr>
<tr><td>Khan & Hoshino (1992)</td><td>5 Asian</td><td>1.2</td><td>1.2</td><td>0.3</td><td></td></tr>
<tr><td>Gang & Khan (1991)</td><td>India</td><td>0.0</td><td>0.0</td><td>0.0</td><td></td></tr>
<tr><td>Heller (1975)</td><td>11 African</td><td>−0.4</td><td>1.1</td><td>−0.1</td><td></td></tr>
</table>

Obs.: † Results from panel fixed-effects estimation of individual equations. Some authors only estimate structural equations (direct effects), neglecting feedback effects that operate within the entire system of structural equations (e.g. Otim, 1996). This table shows the total effects obtained from reduced-form parameters.
Source: McGillivray and Morrissey (2001a:28), complemented by the latest studies.

The table above summarises the main findings of this literature. It is difficult to recognise a consistent pattern of results. The impact of aid flows on government investment seems surprisingly low, sometimes even negative, as in the cases of Rubino (1997), Franco-

Rodriguez et al (1998) and McGillivray and Ahmed (1999). The effect of aid on domestic revenue is strongly negative in Ouattara (2006a), McGillivray and Ouattara (2005), Franco-Rodriguez et al (1998) and Rubino (1997), which supports the hypothesis of tax displacement. However, other studies find a negligible or positive effect, for example Machado (1999), Khan and Hoshino (1992), Iqbal (1997) and Ouattara (2006b). The results for government consumption and borrowing seem inconclusive.

Most of the studies presented above overlook the fact that different aid modalities are likely to affect the government response in different ways.[3] Mavrotas (2005) and Mavrotas and Ouattara (2006) build on the fiscal response tradition by addressing this shortcoming, using disaggregated aid data for Uganda and Côte d'Ivoire, respectively. The four categories that they use are project aid, programme aid, technical assistance and food aid.[4] The results show that most aid modalities have a positive impact on government investment, but cause a reduction in tax revenue.[5] The findings for government consumption are mixed, while domestic borrowing has a strong negative correlation with aid inflows. Hence, while there is little evidence that aid flows are used to increase government consumption, the impact on government investment seems encouraging.

Table 2-3: Results with Disaggregated Aid Variables

Study	Aid Variable	Tax Revenue	Public Investment	Public Consumption	Domestic Borrowing
Mavrotas & Ouattara (2006)	Total Aid	−0.289	−0.196	0.170	−0.761
	Project Aid	−0.413	−0.958	1.477	−1.024
	Programme Aid	−0.163	0.244	−0.067	−0.387
	Technical Assistance	0.129	0.367	−0.531	0.322
	Food Aid	0.129	0.368	−0.531	0.323
Mavrotas (2005)	Total Aid	n/a	n/a	n/a	n/a
	Project Aid	−0.001	−0.050	−0.029	−0.098
	Programme Aid	−0.019	0.197	0.251	−0.489
	Technical Assistance	−0.015	0.292	0.138	−0.493
	Food Aid	0.002	−0.301	−0.364	−0.513

Source: Respective studies.

The variability of the results presented in the previous tables (both in terms of their magnitude and sign of the impacts) seem to suggest that governments respond differently to aid inflows. It is therefore difficult to identify a broad consensus with regard to the impact of aid on fiscal accounts. This strengthens the argument that results tend to be country-specific, either because economic circumstances vary or simply because governments behave differently. We will now proceed with our analysis for Ethiopia.

[3] Although disaggregation between grants and loans (or bilateral and multilateral sources) is not uncommon.

[4] They argue that neglecting aid disaggregation would lead to aggregation bias in the results and conclusions, in cases where preferences of aid-recipient government are higher for some of these types of aid.

[5] The authors suggest that a reduction in taxation effort could benefit the private sector and individual households, so they do not see it necessarily as a bad outcome.

2.3 Economic Background

Economic Performance

Ethiopia is one of the most populated countries in Africa and one of the poorest, with 80 million inhabitants and gross national income at about $280 per capita, respectively. Moreover, Ethiopia is ranked 171 out of 182 countries on the Human Development Index for 2007.

The country has been landlocked since the independence of Eritrea in 1993, and the economy is highly dependent on the performance of the agricultural sector. Agriculture sustains over 80 percent of the population and accounts for nearly half of gross domestic product (GDP) and almost all exports (UNECA, 2002:84). Coffee remains the main source of export earnings, accounting for over 40 percent of total exports in 2005 (IMF, 2006). Ethiopia’s development efforts have often been hindered by irregular rainfall (droughts), volatility in export commodity prices, and costly wars.

In political terms, three main regimes in the recent history of the country can be identified: the Imperial State (until 1974), the Derg Regime (1974-1991) and the Ethiopian People's Revolutionary Democratic Front (EPRDF) coalition (1991-present). Economic performance under the Imperial State was respectable, with real GDP growing by four percent a year (on average) during the period from 1960 to 1974, while average growth per capita was about 1.5 percent (Geda, 2007). Nevertheless, several droughts afflicted the country (e.g. 1958, 1966 and 1973), causing famine and exacerbating poverty. In the early 1970s, a number of events fuelled the discontent of the population: a growing agrarian crisis (Chole, 2004), the economic consequences of the 1973 oil shock, and concerns regarding the inequitable distribution of land.[6]

In 1974, a military coup led by a committee of junior army officers (the Derg) deposed the Emperor. The new regime was characterised by a socialist (centrally planned) economic system with a strong military and discrimination against private property and entrepreneurship. Soon after its rise to power, the Derg embarked on large-scale nationalisations, including all land, private property, financial institutions and manufacturing firms, leading

---

[6] Political power was, to a great extent, related to the size and quality of the land owned (Geda, 2007).

to the 'socialisation' of production and distribution (Geda, 2007). The Derg also introduced protectionist measures to control the flows of international trade, with a view to strengthening the state's role in both exports and imports, emphasising strategic export sectors and closely monitoring the price, quantity and distribution of goods (Geda, 2001:183).

Economic performance under the Derg regime was weaker than in the past, with GDP growing at about 1.9 percent per year (1974-1990), while growth was negative in per capita terms (–0.8 percent). The policy environment, erratic performance of the agricultural sector (e.g. severe drought in 1984-1985) and a lengthy civil war were the main contributors to this sluggish economic record (Chole, 2004).

Another major change in the Ethiopian political and economic context occurred in 1991, when a coalition of rebel forces (EPRDF) succeeded in overthrowing the military regime. The EPRDF reinstated market-economy policies, some of which included the promotion of private sector development, replacement of trade quantity restrictions with tariffs, restructuring of state-owned trading enterprises and liberalisation of the exchange rate (UNECA, 2002:92).[7] During the period 1992-2000, Ethiopia's economic performance improved significantly, albeit with considerable volatility. GDP grew over four percent per year, and about two percent in per capita terms.[8]

Nonetheless, the economy remains vulnerable to climate conditions (e.g. bouts of poor rainfall, such as in 1998 and 2003) and other external shocks (such as terms of trade shocks). The export sector is characterised by a lack of diversification, with a large share of export earnings accruing from a small number of commodities. The historical reliance on the export of a few agricultural goods, with often volatile prices (such as coffee), have to some extent contributed to the instability of export earnings. However, the economy has recently shown positive signs, with real GDP growing at seven percent during the period 2000-2005, and four percent in per capita terms.

Government Accounts

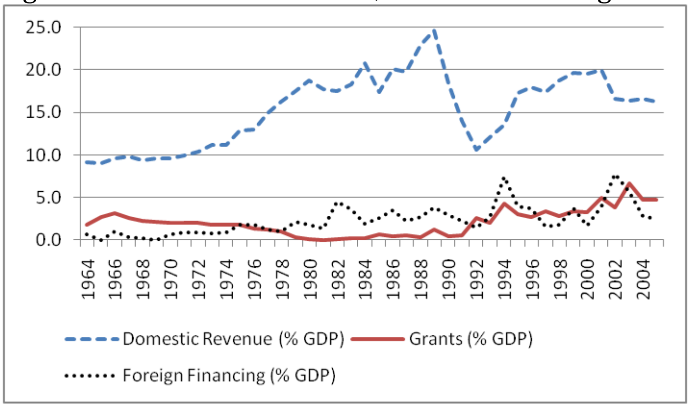
The figure below plots government sources of revenue (including grants) and foreign financing (loans) for the period 1964-2005. During the Imperial State regime, domestic

[7] However, the initial steps of the reform process had been taken by the Derg regime in its last years. "In the case of Ethiopia, the willingness to engage in meaningful reform seems to have prevailed since 1988. The Derg approached Western donors with hat in hand once it became clear that Soviet bloc aid was about to end and that the socialist experiment had failed" (Abegaz, 2001).

[8] Calculated from WDI (2007).

revenues grew slowly as a percentage of GDP, averaging about 10 percent during the period 1964-1974. With the Derg regime in power, revenues had a significantly stronger positive trend, rising from about 11-12 percent to a peak of 25 percent in 1989. In 1985, there was a fall in revenues (as a percentage of GDP), mainly due to the effects of a severe drought. A sharp decrease in revenues ensued in the period 1990-1992 as a result of the state of war in the country, which led to security and administrative problems regarding tax collection.[9] From 1992, however, government revenues managed to recover beyond their previous nominal levels and are currently stabilised at 16 percent of GDP. Foreign financing was practically negligible in the 1960s and 1970s, only surpassing the one percent level of GDP in 1975. This variable has exhibited some volatility, which seems to have increased in recent years.

Figure 2-1: Domestic Revenue, Grants and Foreign Financing (% GDP)

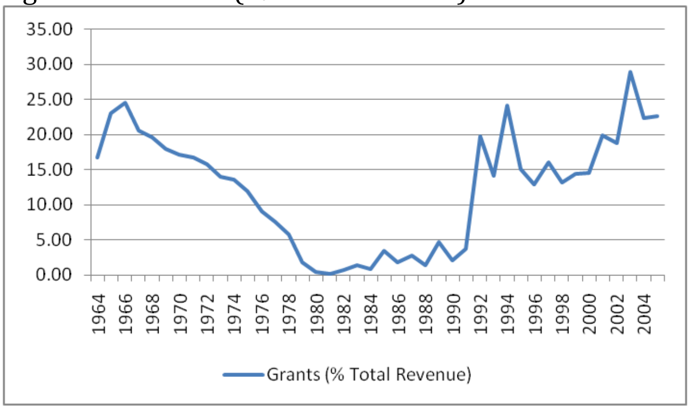


Source: IMF (IFS and Statistical Appendices)

Aid grants declined, relative to GDP, during the Derg regime, but since 1991 have increased at a fast pace, and are currently at around five percent of GDP. Taking into consideration the recent international efforts to increase foreign aid flows to developing countries in order to support the achievement of the Millennium Development Goals (MDGs), this trend seems likely to continue. In terms of their relative importance to the budget, the figure below indicates that grants represented a significant share of total revenues during the Imperial State although they had a steadily declining trend in ensuing years. Their relative importance became almost negligible during the Derg regime, owing to a decline in the volume of grants, but also due to the increase in domestic revenues. As outlined earlier, since the fall of the Derg regime in 1990, grant flows have again become an important source of revenue, currently accounting for about one quarter of total revenues.

[9] Source: National Bank of Ethiopia, Quarterly Bulletin, volume 6(4), 1990/91.

Figure 2-2: Grants (% Total Revenue)



Source: IMF (IFS and Statistical Appendices)

With regard to the composition of domestic revenues, in the 1970s taxes made up about 80 to 90 percent of the total. However, the relative importance of tax revenues declined throughout the 1980s, reaching a low of 60 percent in 1989. Recent data suggest that this trend is now being reversed, with tax revenues increasing from 68 to 79 percent of total revenues during the period 2000-2005.[10] Indirect taxes accounted for most of this recent increase (mainly import duties) while direct taxes seem to have stagnated at around 25 percent of total domestic revenues. It is also interesting to note that export taxes as a percentage of GDP have become negligible,[11] while import duties have increased substantially, mainly as a consequence of the sharp increase in imports of goods and services in the last few years.

We now turn to the composition of total government expenditure for the period 1964-2005. The figures below illustrate expenditure trends, both as a ratio to GDP and as a ratio to total expenditures. We can observe that public consumption was kept at around 10 percent of GDP until 1974 while investment expenditure accounted for about 3-4 percent during the same period. After the Derg regime came to power, and as a consequence of its (socialist) policy measures, government current expenditure rose gradually to almost 20 percent of GDP in 1989, while public spending on capital was about 13 percent in the same year. In 1983, government capital expenditure had almost doubled in nominal terms due to an increase in economic development expenditures (especially those related to 'agriculture and land settlement' and 'manufacturing'). This was also possibly due to the economic 'zemetcha' (mass mobilisation) campaigns (Abegaz, 2001).

---

[10] IMF (2006:41).

[11] Export taxes were once an important item on the revenue side although in the past few years these rates have been almost abolished (Abegaz, 2001).

Figure 2-3: Expenditure Trends (% Total)

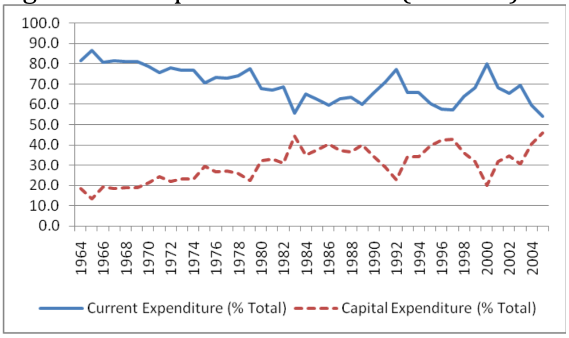
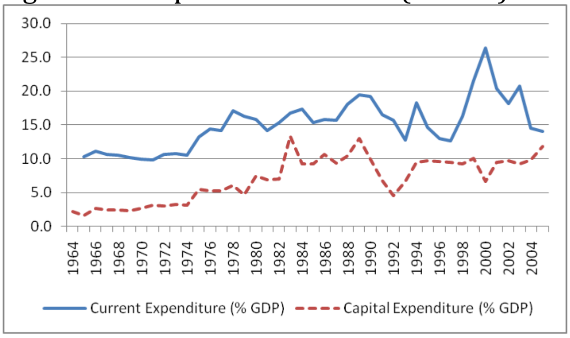


Figure 2-4: Expenditure Trends (% GDP)



Source: IMF (IFS and Statistical Appendices)

Since the beginning of the 1990s there has been some volatility in these variables. The sharp reduction in both items of expenditure in the early 1990s was caused, to some extent, by the drastic fall in domestic revenues and the end of Soviet bloc aid. In 1994, current expenditures increased, mainly due to wages and operating expenses (IMF, 2006) and the doubling of interest payments on internal debt. In 2003, a similar increase was due to a high level of external assistance (which more than doubled). The Eritrean War (1998-2000) was responsible for a sharp increase in military spending, hence the strong increase of current expenditures as a percentage of GDP during this period. Conversely, the fall in recurrent costs thereafter can be attributed to the scaling down of defence expenditures after the end of the war. Defence expenditures had been cut to less than half of their former level in nominal terms by 2002. This represents a reversal of a long increasing trend, in which military spending went from around four percent of GDP in the 1960s to an approximate average of eight percent of GDP in the 1980s, and a high of 15 percent in 2000 (Geda, 2007:13).

In 2000, the drop in capital expenditures was due to a reduction in both economic and social investments, most likely due to the severe drought that afflicted the country in 1999-2000. During the period 2000-2005, current expenditures declined relative to GDP while capital expenditures more than trebled (IMF, 2006:43). There has not been a main driving force behind this increase, but a consistent increase in various items, such as for economic development (mainly agriculture and natural resources) and social development (mainly education and urban development and housing). Current spending on agriculture, natural resources and education more than doubled. Overall, it is quite clear that capital expenditures gradually increased as a share of total expenditures during the period 1964-2005.

The analysis of the government's overall balance (after grants) reveals a high degree of volatility during the period under scrutiny (below). Until 1974, the deficit as a percentage of GDP was relatively stable, at about one percent, but during the protracted civil war, the Derg regime ran higher fiscal deficits.[12] In 1983, the government recorded a very large budget deficit, i.e. almost 12 percent of GDP, mainly due to a sharp increase in expenditures. The high deficits in 1990-1991 corresponded to the last years of the Derg regime, when the civil war worsened. The problems with revenue collection, which caused a sharp reduction in domestic revenues, were not compensated by a proportional fall in expenditures.

[12] Which were financed, to some extent, by monetisation – i.e. printing money (Bevan, 2001:4).

Figure 2-5: Fiscal Deficit (% GDP)

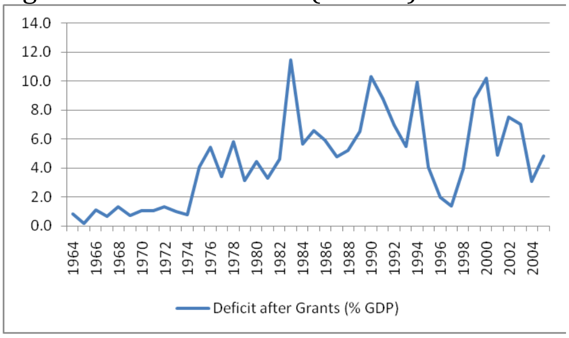
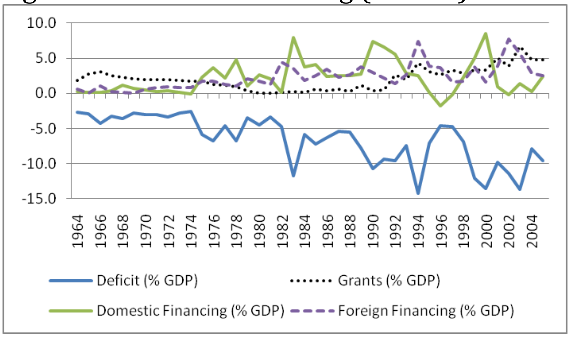


Figure 2-6: Deficit Financing (% GDP)



Source: IMF (IFS and Statistical Appendices)

In 1994, as well as in 1999-2000, the high fiscal deficits were due to a strong increase in government expenditures (below) – a sharp rise in military spending due to the war with Eritrea. Domestic borrowing shows a very strong correlation with the budget deficit before grants, with the exception of 1994, when foreign financing and grants seemed to close this financing gap. If we analyse the fiscal deficit before grants, the government deficit appears to be widening (on average).

Figure 2-7: Revenue and Spending (% GDP)

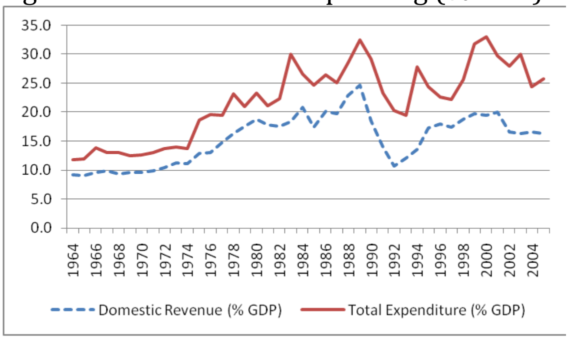
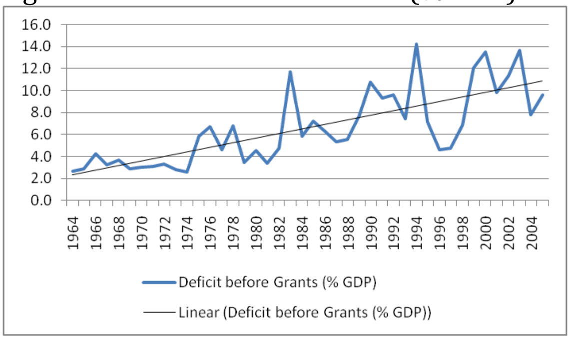


Figure 2-8: Deficit before Grants (% GDP)



Source: IMF (IFS and Statistical Appendices)

2.4 Methodology

The model used in this section follows closely those presented in Mavrotas and Ouattara (2006) and McGillivray and Ahmed (1999). For the purpose of simplicity, we will only illustrate the case in which the aid variable is aggregated (A), although estimates for a disaggregated model will also be provided and examined.[13] The model assumes that government decision-makers wish to maximise a utility function (U), which takes the following quadratic form:

[13] For more details on the analytical model with disaggregated aid, see McGillivray and Ahmed (1999). These authors disaggregate aid flows by source (bilateral and multilateral). The appendix presents the main technical debates on fiscal response models.

$$U = -\left(\frac{\alpha_1}{2}\right)\left(I_g - I_g^*\right)^2 - \left(\frac{\alpha_2}{2}\right)(G - G^*)^2 - \left(\frac{\alpha_3}{2}\right)(T - T^*)^2 - \left(\frac{\alpha_4}{2}\right)(A - A^*)^2 - \left(\frac{\alpha_5}{2}\right)(B - B^*)2 \quad [1]$$

The policy instruments available to public sector decision-makers include: government investment (Ig), government consumption (G), domestic revenue (T), foreign aid inflows (A) and domestic borrowing (B). The variables with an asterisk represent annual targets, which are set *a priori* by the government. As is clear from equation [1], utility reaches its unrestricted maximum at 0, when all variables equal their planned levels. The quadratic form ensures symmetry, i.e. both undershooting and overshooting a target confer the same degree of disutility.[14] The alphas (α) represent the relative weights that decision-makers place on each of their goals, and are all assumed to be positive.

Foreign aid (A) is endogenous, based on the assumption that the government has some degree of control over aid disbursements (Franco-Rodriguez et al, 1998). According to the proponents of this formulation, donors and recipient governments engage in negotiations to determine the planned level of aid, with the government having some bargaining power in terms of how these targets are set, as well as with regard to the amount of aid that is actually disbursed. Hence, decision-makers formulate targets for different expenditure items (in this case Ig* and G*) as well as the revenue components (including aid) during their fiscal planning exercises. The utility maximisation framework then suggests that during the fiscal year they will aim to achieve those same targets.

The utility function [1] is then maximised subject to the following constraints:

$$I_g + G = B + T + A \quad [2]$$

$$G \leq \rho_1 T + \rho_2 A + \rho_3 B \quad [3]$$

Equation [2] is the traditional government budget constraint, which indicates that government expenditures must equal total receipts (revenues and borrowing). This equation must always hold. However, by simply constraining the utility function to equation [2], we are allowing for total fungibility since there is no restriction on how each revenue source is allocated amongst the various expenditure items (Franco-Rodriguez, 2000).

[14] Although some authors have argued against this specification, their proposed alternatives (e.g. linear-quadratic) are not without problems. In fact, the quadratic specification has been the most widely used specification in peer-reviewed publications.

For this reason, inequality [3] is added to the traditional budget constraint, where the rhos (ρ) represent the maximum proportion of domestic revenue, aid flows and domestic borrowing allocated to consumption, with these coefficients assumed to lie between 0 and 1. Conversely, 1–ρ stands for the proportion allocated to investment. However, the interpretation of ρ_2 should not be taken as the extent of fungibility, as in categorical fungibility studies. The reason is that some aid flows intended for development purposes might be recorded as consumption expenditures, and the inequality does not need to hold in every period (Franco-Rodriguez, 2000).

Inequality [3] suggests that there are external constraints that influence how governments allocate their resources amongst expenditure items. The expression can be best understood as actions undertaken by donors or domestic interest groups, which impose the values of the rhos (ρ) on decision-makers, with no guarantee that the targets can be met, even if revenues satisfy the standard budget constraint. If equation [3] is not binding, the utility function [1] is maximised subject to equation [2]. In this case, it is possible to reach the unconstrained maximum (zero) if revenues are sufficient. However, if equation [3] is binding, external and domestic pressures prevent the attainment of this maximum, even when revenues are sufficient, because at least one expenditure target cannot be met (Franco Rodriguez et al, 1998).[15]

Maximising the utility function [1] subject to the constraints [2] and [3] and solving the first-order conditions of the Lagrangian gives the following system of structural equations:[16]

$$\begin{aligned} I_g = {} & (1-\rho_1)\beta_1 I_g^* + (1-\rho_1)\beta_2 G^* + (1-\rho_1)[1-(1-\rho_1)\beta_1 - \rho_1\beta_2]T^* \\ & + [(1-\rho_2) - (1-\rho_1)(1-\rho_2)\beta_1 - (1-\rho_1)\rho_2\beta_2]A \\ & + [(1-\rho_3) - (1-\rho_1)(1-\rho_3)\beta_1 - (1-\rho_1)\rho_3\beta_2]B \end{aligned} \quad [4]$$

$$\begin{aligned} G = {} & \rho_1\beta_1 I_g^* + \rho_1\beta_2 G^* + \rho_1[1-(1-\rho_1)\beta_1 - \rho_1\beta_2]T^* + [\rho_2 - \rho_1(1-\rho_2)\beta_1 - \rho_1\rho_2\beta_2]A \\ & + [\rho_3 - \rho_1(1-\rho_3)\beta_1 - \rho_1\rho_3\beta_2]B \end{aligned} \quad [5]$$

$$\begin{aligned} T = {} & \beta_1 I_g^* + \beta_2 G^* + [1-(1-\rho_1)\beta_1 - \rho_1\beta_2]T^* - [(1-\rho_2)\beta_1 + \rho_2\beta_2]A \\ & - [(1-\rho_3)\beta_1 + \rho_3\beta_2]B \end{aligned} \quad [6]$$

---

[15] Other specifications of the constraints have been used in the literature (e.g. split budget constraint), but since the rhos (ρ) should be the outcome of the maximisation problem, and not be imposed *a priori*, this seems to be a better formulation (Franco-Rodriguez, 2000:432).

[16] Equation [3] is transformed into an equality (corner solution), and B* is set to zero.

$$A = \beta_3 I_g^* + \beta_4 G^* - [(1-\rho_1)\beta_3 + \rho_1\beta_4]T + [1-(1-\rho_2)\beta_3 - \rho_2\beta_4]A^* - [(1-\rho_3)\beta_3 + \rho_3\beta_4]B \quad [7]$$

$$B = \beta_5 I_g^* + \beta_6 G^* - [(1-\rho_1)\beta_5 + \rho_1\beta_6]T - [(1-\rho_2)\beta_5 + \rho_2\beta_6]A \quad [8]$$

where,

$\beta_1 = \alpha_1(1-\rho_1)/\Phi_1$ $\quad\beta_2 = \alpha_2\rho_1/\Phi_1$ $\quad\beta_3 = \alpha_1(1-\rho_2)/\Phi_2$

$\beta_4 = \alpha_2\rho_2/\Phi_2$ $\quad\beta_5 = \alpha_1(1-\rho_3)/\Phi_3$ $\quad\beta_6 = \alpha_2\rho_3/\Phi_3$

and,

$\Phi_1 = \alpha_1(1-\rho_1)^2 + \alpha_2\rho_1^2 + \alpha_3$ $\quad\Phi_2 = \alpha_1(1-\rho_2)^2 + \alpha_2\rho_2^2 + \alpha_4$

$\Phi_3 = \alpha_1(1-\rho_3)^2 + \alpha_2\rho_3^2 + \alpha_5$

Since the alphas (α) are expected to be positive and the rhos (ρ) to lie between 0 and 1, it can be concluded that the betas (β) are also expected to be positive. These structural equations capture the direct impact of aid flows on the remaining fiscal variables although it is argued that the estimation of these equations ignores indirect feedback effects, which operate through the simultaneous-equation system. To capture the total impacts (direct and indirect), the reduced-form equations need to be derived. When the previous structural equations are solved simultaneously (ensuring that only exogenous variables remain on the right-hand-side), the system of reduced-form equations can be represented as follows:

$$I_g = \delta_1 I_g^* + \delta_2 G^* + \delta_3 T^* + \delta_4 A^* \quad [9]$$

$$G = \delta_5 I_g^* + \delta_6 G^* + \delta_7 T^* + \delta_8 A^* \quad [10]$$

$$T = \delta_9 I_g^* + \delta_{10} G^* + \delta_{11} T^* + \delta_{12} A^* \quad [11]$$

$$A = \delta_{13} I_g^* + \delta_{14} G^* + \delta_{15} T^* + \delta_{16} A^* \quad [12]$$

$$B = \delta_{17} I_g^* + \delta_{18} G^* + \delta_{19} T^* + \delta_{20} A^* \quad [13]$$

The deltas (δ) are specific combinations of rhos (ρ) and betas (β), and ultimately alphas (α). What is clear from the equations above is that the reduced-form specification will be able to give information about the impact of aid targets (often commitments), and not aid disbursements. By looking at the structural equations [4]-[8], we note that the impact of A* on A will be decisive in assessing the impact of aid flows on the remaining fiscal variables.[17]

However, there is a practical problem. Since the system represented by the reduced-form equations [9]-[13] is over-identified, it is not possible to estimate it simultaneously (McGillivray, 2000:162). Hence, it is necessary to estimate the system of structural equations [4]-[8],[18] and then substitute the estimated coefficients back into the equations before solving the system simultaneously.

2.5 Data

The data were collected from the International Financial Statistics (IFS) database of the International Monetary Fund (IMF), which was complemented by several issues of the IMF Statistical Appendices for Ethiopia. The data cover the period 1964-2005. The variables collected include: government revenues excluding grants (T), grants (A1), government total expenditure (GE), government capital expenditure (Ig),[19] financing from abroad (A2), domestic financing (B), imports (M), exports (X) and GDP (Y). All the variables are expressed in million Birr and transformed to real values by the GDP deflator.

The consistency of the data was checked by applying the standard budget constraint, where total government expenditures must equal the sum of domestic revenues (T), grants (A1), foreign finance (A2) and domestic borrowing (B). There were only minor differences (decimal points), but for the sake of consistency the domestic borrowing variable was derived as a residual by using the traditional fiscal constraint. For the remainder of this paper, the terms 'capital expenditure' and 'investment', as well as 'current expenditure' and 'consumption', will be used interchangeably.

---

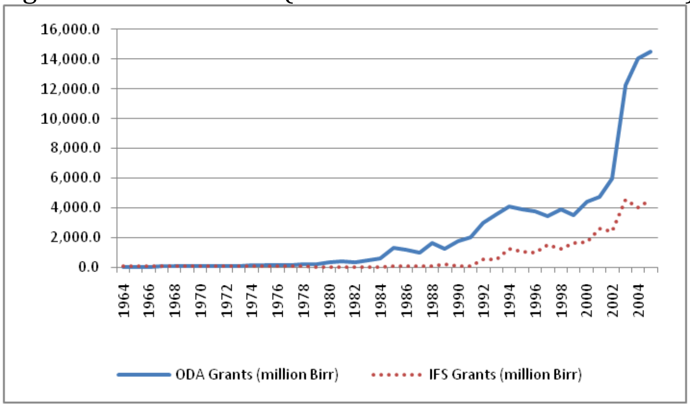
[17] There is a vast literature on aid uncertainty suggesting that aid commitments are poor predictors of disbursements. This might, to some extent, compromise the impact of A* on aid flows A, and hence the results of the model.

[18] Due to the budget constraint, one of these equations will be redundant, so in practice we have to drop one of them. Equation [5] is the main candidate since all of its coefficients are present in the other equations.

[19] This was initially derived as the difference between total expenditure (GE) and government consumption from the national accounts. However, since the latter, by definition, does not include interest payments and government transfers, the series was corrected to the extent possible by the IMF Statistical Appendices for the period 1992-2005. A thorough comparison with other sources (e.g. NBE Quarterly Reports) suggests that this new series is accurate.

While most of the studies in this literature obtain their foreign aid variables from the OECD-DAC statistics, this paper argues that this is not adequate for Ethiopia. The figure below compares IMF data with the OECD-DAC statistics for aid grants.

Figure 2-9: Aid Grants (OECD-DAC versus IMF-IFS data)



Source: IMF (IFS and Statistical Appendices) and IDS-DAC[20]

The DAC reported values for ODA grants exceed those reported by the IMF by a considerable margin. This discrepancy may arise from the fact that some aid flows are not reported in the government budget – usually known as 'off-budget' items (see MacKinnon, 2003:9). These tend to be grant-funded projects that are implemented without the knowledge of the government planning office, either due to the lack of reporting of aid-funded activities by donor agencies or lack of communication between sectoral ministries, regional offices and the central government. For this reason, the use of IMF data will be more appropriate for this study since it is compiled from government sources and reports the amount of grants that government decision-makers are actually aware of.[21]

These are the values that will be taken into consideration when formulating and revising the budget, hence influencing the fiscal decisions undertaken by the government. The DAC data would be better suited for other macroeconomic questions, such as the impact on the exchange rate, where total aid flows are more pertinent. Another problem arising from the use of DAC statistics is the interpretation of the domestic borrowing variable (B). Since this is often constructed as a residual from equation [2], so as to ensure that the fiscal data are consistent, borrowing from domestic sources could be substantially underestimated, potentially biasing the results of our model.

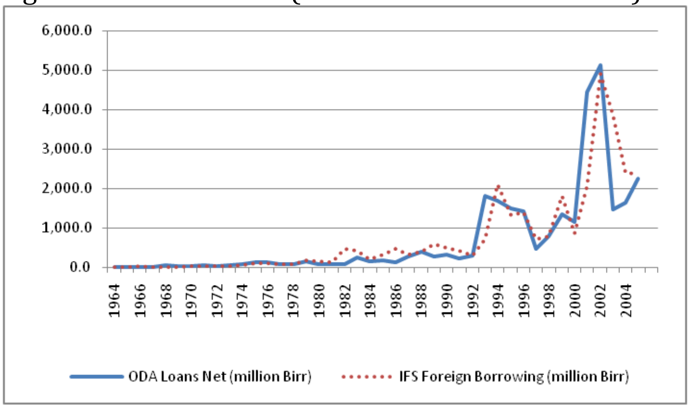
---

[20] The IFS (IMF) exchange rate was used to convert the DAC USD values into Birr.

[21] It is important to note, however, that being reported in the budget does not necessarily indicate that the funds are channelled through the treasury or that the government is directly in charge of the implementation of the project or programme.

With regard to loans, the IMF reports foreign financing, which includes all government borrowing from abroad. In theory, this item would incorporate both concessional lending (e.g. IDA 'soft' loans) and borrowing at competitive (or commercial) rates. Although it is not feasible to directly assess the proportion of concessional lending, it is possible to compare the IMF data with the DAC values for ODA loans. Since all foreign loans will be, in principle, reported in the budget (as opposed to grants), these two series are likely to be identical if most borrowing is, as expected for a country such as Ethiopia, from bilateral and multilateral concessional sources.[22]

Figure 2-10: Aid Loans (OECD-DAC versus IMF-IFS)



Source: IMF (IFS and Statistical Appendices) and IDS-DAC

The figure above shows a strong correlation between the two variables: ODA (concessional) loans and foreign borrowing. This suggests that most foreign loans to the government were provided on concessional terms, which is not particularly surprising due to the difficulty that some African countries face when borrowing from international markets. The higher IMF estimates during the 1980s might be explained by the existence of loans from the Soviet bloc, which are not included in the total ODA figures. From 2000 onwards, there might be discrepancies regarding the recording date. Hence, the grants variable (A1) can be lumped with foreign financing (A2), and it can be argued with some confidence that this is a proxy for total aggregate aid flows (A).

Since it is difficult to obtain data for the target variables, the standard procedure followed in the literature is to construct proxy variables. These approximations are, in fact, the fitted values of long-run cointegrating relationships between each of the fiscal variables and a set of exogenous regressors. The rationale is that, since government targets are based on expectations about the future values of revenues and expenditures, the fitted

[22] Aid loans usually need to be approved by MoFED. See also MacKinnon (2003:11).

values of a cointegrating relationship will represent the values that could be obtained in equilibrium (Franco-Rodriguez, 2000).

For this purpose, these targets will be approximated by estimating an unrestricted error-correction model (ECM) and testing for cointegration. This methodology combines the long-run information and the error-correction mechanism in the same equation. One of the main advantages is that the variables of the cointegrating relationship can be either stationary or integrated of order one. In this particular case, this is important since preliminary tests could not reject the hypothesis that some of the relevant variables are stationary.[23] The generic (unrestricted) ECM specification is based on an autoregressive distributed lag model (ARDL). The error correction representation of an ARDL(p, q, q, …, q) is outlined below:

$$\Delta y_t = \phi y_{t-1} + \beta' X_{t-1} + \sum_{j=1}^{p-1} \lambda_j \, \Delta y_{t-j} + \sum_{j=0}^{q-1} \delta_j' \, \Delta X_{t-j} + \varepsilon_t$$

where X is a vector of explanatory variables, β contains information about the long-run impacts, ϕ is the error correction term (due to normalisation), and δ_j incorporates short-run information. The cointegration test presented here is the bounds test approach proposed by Pesaran et al (2001:307). This is obtained from the application of coefficient restrictions on the long-run coefficients (Wald test). If these coefficients are (jointly) statistically different from zero, then it is suggested that there is cointegration, hence a long-run relationship amongst the variables. The critical values used to assess cointegration are those from Pesaran et al (2001). The target for government revenues (T*) is obtained by regressing T on exports (X) and imports (M). Export taxes were until recently an important source of revenue while import duties have accounted for a substantial share of tax revenue. The target for government investment (Ig*) is obtained from the fitted values of the long-run relationship among GDP, loan commitments and Ig.[24]

Since it was not possible to find a cointegrating relationship for government consumption expenditure, we use a regression with its lagged values and a time trend – see Franco-Rodriguez (2000). With regard to the aid variables, we use the OECD-DAC data on ODA commitments as a proxy for aid targets, a procedure employed in several fiscal response

[23] The remaining variables are integrated of order one. See Appendix.
[24] The explanatory variables for these target approximations have to be exogenous to the model. This means that in this case domestic revenue and aid flows cannot be used on the right-hand side of the equation.

studies. The implicit assumption is that governments have a certain degree of bargaining power in setting the amount of these commitments, and so these values can be taken as aid targets. Although this procedure seems to be sensible for the loan target, it might not be adequate for the grants target since this will be overestimated. Hence the Creditor Reporting System (CRS) database of the DAC was used to proxy for this variable.[25]

2.6 Empirical Results

2.6.1 Three-Stage Least Squares

Once the target variables are constructed, it is possible to proceed to the estimation of the system of structural equations. It is standard practice to use the non-linear three-stage least squares (N3SLS) estimation method since the system of equations [4]-[8] is not linear in the coefficients, and contains cross-equation restrictions. The table below presents the results.

Table 2-4: Estimation Results

Model	Parameter	Estimate	Std. Error	t-Statistic	P-value
Aggregated Aid	ρ_1	0.547***	0.031	17.755	0.000
	ρ_2	0.833***	0.067	12.372	0.000
	ρ_3	0.892***	0.092	9.716	0.000
	β_1	0.201	0.191	1.054	0.294
	β_2	1.605***	0.161	9.964	0.000
	β_3	0.155	0.109	1.431	0.154
	β_4	1.165***	0.105	11.094	0.000
	β_5	0.141	0.112	1.265	0.208
	β_6	1.101***	0.098	11.246	0.000
Disaggregated Aid	ρ_1	0.626***	0.045	13.898	0.000
	ρ_2	0.704***	0.203	3.467	0.001
	ρ_3	0.554**	0.236	2.346	0.020
	ρ_4	0.902***	0.122	7.386	0.000
	β_1	0.172	0.211	0.815	0.416
	β_2	1.478***	0.175	8.433	0.000
	β_3	0.165*	0.085	1.926	0.056
	β_4	0.676***	0.158	4.270	0.000
	β_5	0.236**	0.101	2.341	0.020
	β_6	−0.636	0.448	−1.419	0.158
	β_7	0.138	0.128	1.077	0.283
	β_8	1.083***	0.128	8.438	0.000

Obs.: ρ4, β7 and β8 in the disaggregated model correspond to ρ3, β5 and β6 in the aggregate model (i.e., parameters associated with domestic borrowing). The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

[25] This was done by subtracting sector allocable and emergency aid commitments from the grant total.

The first step is to analyse the results of the aggregated model. As mentioned earlier in the paper, the rhos (ρ) represent the proportion of revenues, aid flows and domestic borrowing allocated to consumption. The estimated coefficients suggest that 55 percent of domestic revenues are allocated to consumption, with the remaining 45 percent going to public investment. This result seems very plausible, as most developing countries will want to ensure that an important share of government recurrent costs is covered by domestic revenue sources.

The second coefficient suggests that 83 percent of foreign aid is allocated to pay for recurrent costs, which seems considerably high. Finally, almost 90 percent of domestic borrowing is allocated to consumption expenditure, which is perhaps a concern since loans should ideally be used for investment purposes. However, this may suggest that the government often turns to the domestic borrowing market to close the financing gap (the deficit).

The betas (β) do not have an intrinsic interpretation, but can be used to identify the utility weights (alphas). They are all positive, as expected. The p-values in the last column indicate that β_1, β_3 and β_5 are not statistically significant even at the 10 percent level of confidence, while the remaining coefficients are strongly significant. Some authors would choose to restrict insignificant coefficients to zero, as well as limit the rhos (ρ) to lie within the interval [0, 1] and the betas (β) to be only positive numbers. However, this procedure might have the implication of biasing the results.[26]

Turning now to the disaggregated model, we can see the impact of disentangling the effects of grants and foreign borrowing. All the rhos (ρ) are statistically significant, and lie within the expected interval [0, 1]. Corroborating the previous results, a large share of domestic revenues seems to be devoted to recurrent expenditure (roughly two-thirds) while 70 percent of grants are allocated to consumption. Although this would seem to support the argument that aid (in this case grants) is fungible, it is important to bear in mind that a portion of current expenditures will be developmental (e.g. for salaries of school teachers). Moreover, total grants to the country are not likely to be so 'pro-consumption' due to the existence of off-budget investment grants.

The coefficient for loans, by contrast, looks more 'pro-investment', with 45 percent of loans allocated to capital expenditures. Finally, the figure for borrowing reinforces the

[26] See Ouattara (2006:1113-4).

previous finding that most domestic borrowing is used to pay for recurrent costs. There are three betas (β) that do not seem to be statistically significant (at the 10 percent confidence level), but since most of these estimated values are relatively low, they are not expected to bias the results significantly. Nevertheless, the models have also been solved for the case where β_1, β_6 and β_7 take the value of zero, and this does not change the final results and conclusions significantly.

Direct Impact of Aid Flows

Now that we have estimated the system of equations, we can substitute the estimated coefficients into the structural equations in order to obtain the direct impacts of aid flows. The table below summarises the relevant impacts.

Table 2-5: Direct Impacts of Foreign Aid

Model	Impact		Estimate
Aggregated Aid	A on Ig	$(1-\rho_2) - (1-\rho_1)(1-\rho_2)\beta_1 - (1-\rho_1)\rho_2\beta_2$	−0.454
	A on G	$\rho_2 - \rho_1(1-\rho_2)\beta_1 - \rho_1\rho_2\beta_2$	0.083
	A on T	$-(1-\rho_2)\beta_1 + \rho_2\beta_2$	−1.371
	A on B	$-(1-\rho_2)\beta_5 + \rho_2\beta_6$	−0.941
Disaggregated Aid	A1 on Ig	$(1-\rho_2) - (1-\rho_1)(1-\rho_2)\beta_1 - (1-\rho_1)\rho_2\beta_2$	−0.113
	A1 on G	$\rho_2 - \rho_1(1-\rho_2)\beta_1 - \rho_1\rho_2\beta_2$	0.021
	A1 on T	$-(1-\rho_2)\beta_1 + \rho_2\beta_2$	−1.092
	A1 on B	$-(1-\rho_2)\beta_7 + \rho_2\beta_8$	−0.803
	A2 on Ig	$(1-\rho_3) - (1-\rho_1)(1-\rho_3)\beta_1 - (1-\rho_1)\rho_3\beta_2$	0.111
	A2 on G	$\rho_3 - \rho_1(1-\rho_3)\beta_1 - \rho_1\rho_3\beta_2$	−0.006
	A2 on T	$-(1-\rho_3)\beta_1 + \rho_3\beta_2$	−0.896
	A2 on B	$-(1-\rho_3)\beta_7 + \rho_3\beta_8$	−0.661

As mentioned earlier, it is important to bear in mind that these results represent only partial effects. The impact of aggregate aid flows on capital expenditure seems to be negative, as is the effect on domestic tax revenue collection. Government current expenditures are positively correlated with aid flows, while the last coefficient suggests that foreign aid is a substitute for domestic borrowing. Nonetheless, if specific types of aid have different fiscal impacts, these results will be biased due to the aggregation of aid.

The disaggregated model seems to support this proposition, as the direct impacts of aid grants and loans on expenditure are significantly different. Grants appear to have a negative impact on public investment and a weak positive effect on public current expenditure while loans have a positive effect on public investment and an almost negligible impact on government current expenditure.[27] Moreover, both grants and loans seem to have a negative effect on revenue collection, with the negative impact of grants being stronger than that for loans. Similarly, both forms of aid have a negative relationship with domestic borrowing. These results suggest that aid flows are a substitute for both taxes and domestic borrowing.

Direct and Indirect Effects of Aid Flows

Although the results so far have given us interesting insights into the fiscal response to aid flows, these are only partial (direct) effects. To obtain the total (direct and indirect) effects that run through the simultaneous system, we need to analyse the estimates for the reduced-form equations. This is done by simultaneously solving the system of structural equations (with the estimated coefficients).

Table 2-6: Total Impacts of Foreign Aid

Impact	Ig	G	T	B
Total Aid (A*)	0.016	−0.001	−0.059	−0.524
Grants (A1*)	0.064	−0.004	−0.474	−0.435
Loans (A2*)	0.295	0.033	−0.154	−0.457

The results reported above suggest that an increase in aid commitments (which are likely to be followed by aid disbursements) has a weak positive impact on public investment, while it reduces domestic revenues. The impact on recurrent expenditure is almost negligible. Finally, the impact on borrowing is strongly negative, suggesting that aid and domestic borrowing are close substitutes. However, as stated before, these coefficients are likely to be biased due to the aggregation impact.

The disaggregated model provides further insights into the impact of aid inflows on the fiscal aggregates. Both grants and foreign loans have a positive impact on public investment, with loans having a stronger impact. This may happen because loans tend to be associated with capital projects, while grants captured in the budget figures might be biased towards consumption due to the existence of grant-funded projects not recorded in the budget (i.e. off-budget items). The impact on current expenditure is positive, but weak. The negative results for domestic revenue seem to corroborate the conclusions of the aggregated aid model, although the coefficient on grants is significantly greater than for loans.

[27] The negative effect of grants might arise from the fact that one of the grant components is positively correlated with natural disasters (e.g. drought relief), while investment expenditures are likely to decrease in such events.

This effect suggests that grants produce a stronger tax disincentive effect than loans, perhaps because they do not have to be paid back. However, it could be the case that the model is capturing some 'exogenous' or indirect effects rather than a disincentive caused by higher aid flows, such as droughts or aid conditionality.[28] Moreover, this result may also reflect the observation that donor support (especially grants) decreased during the Derg regime, precisely as tax revenues were significantly increased.[29] Finally, the impact of both forms of aid on domestic borrowing is strongly negative, suggesting that they both act as substitutes for domestic borrowing, with similar magnitudes.

The results presented above suggest that foreign aid directly finances new public expenditure. We may argue that by substituting for domestic borrowing and revenue, foreign aid also does so indirectly. In fact, similar conclusions were reached in a recent study on Ghana (Osei et al, 2005), which finds that aid does not have a strong direct effect on the volume of government spending but is treated as a substitute for domestic borrowing.

Although some donors would be concerned with the lack of additionality of aid resources, to the extent that a birr of aid does not generate an equivalent increase on the expenditure side, substituting aid for domestic borrowing or even using aid flows to retire onerous debt might be a desirable strategy in cases where the debt burden is high.

Taking into consideration that domestic borrowing is often seen as an expensive last resort to balance the budget, this strong negative correlation with aid might suggest that a significant share of aid flows is not incorporated into the budget planning process (notwithstanding being recorded). This may arise in cases where aid unpredictability and volatility undermine long-term fiscal planning, prompting large fluctuations in the levels of domestic borrowing. This might explain why foreign aid does not have a stronger (direct) impact on the other fiscal variables, especially on the spending items.

Although the use of a fiscal response model has improved our capacity to analyse specific fiscal dynamics in Ethiopia, we need to be aware of its limitations. One of the limitations concerns the underlying assumptions about government behaviour that are embedded in the model (e.g. utility maximisation and symmetric functions). If these assumptions are

[28] The volume of grants may increase during natural disasters while tax collection is likely to be reduced. Another possibility is that the model is capturing some degree of aid conditionality. The fiscal data show that aid flows to Ethiopia (especially grants) increased substantially from the early 1990s, roughly when economic reforms started to be implemented.

[29] Since this model is static, it is difficult to infer causality.

incorrect, the results and conclusions from this framework could be misleading. Another weakness relates to the use of proxies for the target variables since data on budgeted (i.e. planned) figures are not readily accessible. It is often acknowledged by researchers that these models might be sensitive to the way targets are approximated. Finally, even though the results and conclusions seem plausible, not all of the results appear to be robust to changes in the sample size. This can arise from the existence of outliers in the data or perhaps structural breaks in the time series.

2.6.2 Cointegration Analysis

In order to complement the previous analysis, and test the robustness of the results obtained, this paper also estimates cointegrating relationships for each of the main fiscal variables.[30] This approach is more straightforward than the fiscal response framework, and addresses some of the shortcomings of the previous methodology: there is no need to make strong assumptions about government behaviour or to estimate proxies for the target variables. It also allows testing for structural breaks and outliers in the data.

Nevertheless, the single-equation approach presented in this section may ignore potential interactions among fiscal variables, which can be important if fiscal decisions are taken simultaneously. This is also referred to as the 'endogeneity problem', which occurs when, for example, the level of expenditure influences revenue collection, which in turn influences spending decisions. This is one of the main reasons why the single-equation approach is not popular in the fiscal response literature. However, the purpose of this exercise is to strengthen our understanding of the fiscal dynamics in Ethiopia. If the results corroborate the conclusions of the previous fiscal response analysis, they could suggest that the relationships in the data are fairly strong, and that the model specification is not 'driving' the results.

The relationships between fiscal aggregates and other relevant macroeconomic variables are studied in an unrestricted error-correction model (ECM). This methodology has some advantages over other cointegration approaches: (i) the variables can be either stationary or integrated of order one; (ii) only one equation is estimated, combining both short- and long-run information; and (iii) cointegration testing is fairly straightforward. To assess the

---

[30] This approach is advocated by Geda (1996:4) as a better reflection of African countries' fiscal process: (i) set expenditure targets based on expected growth and development objectives; and then (ii) finance these expenditures through foreign aid, domestic revenues or borrowing (often in this order).

existence of a long-run relationship amongst the variables (cointegration), we need to test the joint significance of the long-run coefficients in each equation (Wald test). Standard diagnostic tests are also applied to the estimated equations: the Jarque-Bera test (normality), Ramsey RESET test (specification), Breusch-Godfrey LM test (serial correlation) and Breusch-Pagan-Godfrey test (heteroscedasticity). All the specifications passed these tests at the five percent confidence level, unless otherwise specified.

Table 2-7: Long-Run Impact of Aid on Expenditures and Revenue

Variables	Capital Expenditures (Ig)			Current Expenditures (G)				Revenue (T)	
	(1.1)	(1.2)	(1.3)	(2.1)	(2.2)	(2.3)	(2.4)	(3.1)	(3.1b)
Constant	−4.21***	−2.43***	−2.30***	0.83	1.21**	1.27	0.76	1.19	−6.45*
Revenue	1.27***	1.05***	1.02***	0.91***	0.71***	0.77***	0.87***		
Imports								0.44*	1.35***
Exports								0.43	0.39
Total Aid	0.14				0.18**				
Grants		0.00				0.07**	0.06**	−0.09	0.13
Loans		0.19***	0.20***			0.05		0.15*	0.05
D83	1.11***								
D94				0.86***					
R^2	0.70	0.71	0.69	0.46	0.45	0.42	0.40	0.71	0.86
F-stat ⊥	4.91**	5.35**	7.19***	7.47**	5.92**	4.15*	5.49**	3.48	6.16***

Obs.: All variables are in logarithmic form, except the constant and dummy variables. Deterministic time trends were not statistically significant. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels, according to the critical values in Pesaran et al (2001). ⊥ Wald test.

The results in the first three columns correspond to the impact on government capital expenditures. Since the first specification (1.1) failed the normality test on the residuals, a dummy variable had to be added to solve an outlier in 1983.[31] All three specifications have strong R-squares, while the Wald tests indicate that these are valid long-run relationships. The coefficients suggest that both domestic revenues and aid loans have a positive impact on the level of capital expenditures. However, aid grants are not statistically significant.

These results seem robust across time and support the previous findings, namely, that aid loans have a stronger positive impact on investment than grants. The coefficients can be read as long-run elasticities (i.e. percentage changes), since the equations were estimated in log-log form. Therefore, a one percent increase in aid loans will cause an increase of about 0.2 percent in capital expenditure.

The results in the following four columns also seem to suggest that domestic revenue and aid flows have a positive impact on current expenditures. However, it appears that in this case grants have a stronger impact than loans. As before, the first specification failed the

[31] This outlier is possibly due to the economic 'zemetcha' campaigns.

normality test, and a dummy variable for 1994 had to be included.[32] In this case the R-squares are weaker, possibly due to the recurrent behaviour of this type of expenditure, which makes it less responsive to changes in revenue and aid flows. Nonetheless, the Wald tests suggest cointegration amongst the variables.

The final two columns contain the results for the impact of aid flows on domestic revenues. [33] The first specification (3.1) is not a cointegration relationship, and it fails the structural form test (RESET). Alternative specifications were tested but did not improve the results (e.g. including GDP as a proxy for the level of economic activity as well as private consumption). Further analysis through CUSUM and Chow Breakpoint tests suggests that there is a breakpoint around 1992. This is not particularly surprising, as the relationship between domestic revenues and trade variables (imports and exports) is likely to have changed after economic reforms were introduced during this period.

Equation (3.1b) uses a sub-sample (1964-1992), and suggests that imports accounted for a substantial share of domestic revenues during this period. Neither exports nor aid flows seem to be significant. Although most of the coefficients on aid are not statistically significant in the various specifications utilised, there were very few cases in which the coefficients were negative. This might suggest, as highlighted before, that the impact of aid on domestic revenue is not a direct effect since the negative coefficient might hide specific tax and non-tax dynamics.[34]

With regard to domestic borrowing, regressing its value (B) on the fiscal deficit before grants and aid flows would be the same as estimating the budget equation.[35] However, it seems clear from the plots that domestic borrowing is strongly driven by the size of the budget deficit (excluding grants), while both types of aid flows seem to be strong substitutes for domestic financing.

To conclude, the results clearly suggest that domestic revenues are the most important source for funding government expenditures, while foreign aid flows also have a significant positive impact on spending. It is important to note, however, that while aid loans tend to be associated with capital expenditures, grants have a stronger relationship with

[32] This outlier is caused by a sharp increase in government (recurrent) expenditures in that year.

[33] Private consumption and GDP (proxy for the level of economic activity) were not statistically significant.

[34] In fact, since this dynamic specification fails to find a robust effect of aid on domestic revenues it may be that the causality runs from domestic revenues (Derg regime policies) to aid flows (donor withdrawal). We recall that traditional fiscal response models are static, and therefore not able to assess causality.

[35] The coefficient for the deficit (expenditure minus domestic revenue) would be 1, and –1 for foreign aid flows. In the levels form it would be: $B = a + b^*(G + Ig - T) + c^*A$, with $a = 0$, $b = 1$, $c = -1$.

current expenditures. This result does not suggest that loans are more effective than grants, since some recurrent expenditure is essentially developmental, such as salaries for health staff and the purchase of medicines. Moreover, off-budget grants often tend to be directed for investment purposes.

Overall, these results corroborate the conclusions of the fiscal response model. Aid flows have a positive impact on government spending, while the results for domestic revenue do not seem to be robust across time. More research would thus be required into specific tax and non-tax dynamics (e.g. disaggregating by type of revenue or adding omitted explanatory variables) in order to more clearly understand whether aid flows have an impact on domestic revenues.

2.7 Conclusions

The main aim of this paper was to assess the fiscal effects of aid flows in Ethiopia. It started by analysing some of the most important economic developments of the past 40 years, focusing its attention on the government sector. It then provided a brief overview of the recent debates surrounding the fiscal response literature, and applied a fiscal response model to the Ethiopian data (1964-2005).

Two versions of the model were estimated. The first specification included a variable representing total aid inflows, while the other disaggregated this variable into aid grants and foreign loans. The main objective of the model was to examine how an increase in foreign aid inflows affects the recipient country's expenditure decisions, namely, the allocation of public resources and borrowing for such purposes, and the extent to which foreign aid produces negative incentives in relation to revenue collection. The model was then complemented by cointegration analysis.

Overall, the results presented in this study seem encouraging. Foreign aid has a clear positive impact on public spending, with a slight bias towards investment expenditures. The impact on domestic borrowing seems to be robust to different specifications, leading to the conclusion that both aid grants and loans act as substitutes for domestic borrowing.

There are two possible interpretations for this relationship. Since domestic borrowing is often an expensive last resort to balance the budget, an increase in aid flows could be used to reduce such an onerous burden.[36] In this case, the incremental aid is indirectly paying for expenditures that would otherwise be financed through expensive domestic borrowing. Conversely, it might also be the case that, facing a shortfall in aid flows, the government resorts to domestic finance in order to keep expenditure levels stable. This could be one of the negative consequences of aid volatility and unpredictability, which should be addressed by donor countries in the spirit of the Paris Declaration.

Finally, there is some evidence that an increase in aid will have a negative impact on revenue collection. This may fuel concerns about external dependency and the lack of long-term sustainability of such dynamics. However, this result does not seem to be robust across the sample, suggesting that the model does not fully capture the dynamic behaviour of domestic revenues.[37]

These results are not totally surprising. Aid flows (especially loans) are often earmarked to specific investment projects, while governments are likely to use tax revenues to pay for most recurrent costs. According to UNECA (2002:89), roughly half of the Ethiopian government's capital spending was financed through external sources, while most of the recurrent budget was paid for by domestic revenues. It is therefore predictable that an increase in aid inflows would have a stronger positive impact on the capital budget than on current expenditure.

As a result of the analysis undertaken in this paper, it is possible to say that over the 40-year period here considered, the fiscal deficit before grants has clearly widened (aid has been 'spent'), while aid flows have gradually increased, especially post-1992. The empirical model indicates that aid flows have not only induced an increase in government expenditures but also (subject to some caveats) led to a reduction in domestic revenues.

In conclusion, the results suggest that donors and recipients of aid should focus on ensuring greater 'aid additionality.' The priority for donors and governments should be to finance new capital spending, which can contribute not only to enhancing human development but also to expanding the economy's productive capacity. While using foreign aid to pay-off onerous domestic debt can serve a useful purpose at least in the short-run, the

[36] In Ethiopia, the main domestic source of deficit financing has been the banking system (e.g. through central bank 'advances') rather than non-banking sources. This is particularly noticeable during the Derg regime, when the budget deficit was often covered by monetisation (i.e. printing money). More recently, the government has also resorted to a (developing) domestic bond market and loans from commercial banks.

[37] An earlier section provided some insights into the composition of domestic revenues through the years. Its volatility, presence of structural breaks, and the omission of important determinants could justify the lack of robustness.

medium-term purpose of aid should clearly be to expand MDG-related government spending to improve human development. Moreover, donors should channel more aid into building up national capacities to mobilise domestic revenue. Otherwise, developing countries such as Ethiopia will have difficulty in graduating from their reliance on external aid.

2.8 Appendix

Fiscal Response Modelling Debates

(i) Functional form. The linear-quadratic utility function used by Heller (1975) is problematic. As Binh and McGillivray (1993) point out, the inclusion of linear terms in the utility function is not consistent, since the function is not maximised even when the targets are met (instead, it is maximised by overshooting expenditures and undershooting revenues and borrowing). Hence, most of the subsequent work uses a symmetric quadratic form. However, some argue that the symmetry assumption is not realistic – i.e. overshooting a target is not the same as undershooting it. Gang and Khan (1999:126) propose an alternative (asymmetric) utility function to overcome this problem. They specify a loss function that depends on certain assumptions about the type of policy-maker. They derive eight different types based on preferences for overshooting/undershooting specific targets. The downside of this approach is that the results are inherently dependent on choosing the correct type for the country under analysis.

(ii) Constraints. The fiscal response literature tends to avoid the single constraint based on the fact that it allows aid to be completely fungible.[38] However, using a split constraint over-restricts the model,[39] not allowing the utility function to reach its maximum even when aid revenues are enough to meet all targets (McGillivray and Morrissey, 2001a). White (1994) argues that this allocation is usually the result of the utility maximisation problem (policy choice) and should not be predetermined by the specifications on the budget constraint, consequently advising the use of a single budget constraint (which avoids the unnecessary imposition, a priori, of fungibility). Franco-Rodriguez et al (1998)

---

[38] It is argued that the use of a single budget constraint implies no constraints on how revenues are allocated, thus aid is completely fungible (McGillivray and Morrissey, 2001a).

[39] The split budget constraint over-restricts recipient budgetary behaviour since the model predetermines the allocation of the income terms in given proportions (White, 1994). See also Gang and Khan (1994) for further discussion.

try to solve these problems by assuming that external pressures may not allow recipient economies to reach the targets even if revenues satisfy the general budget constraint.

(iii) Targets. The use and estimation of targets in the utility maximisation framework is a contentious issue. Fiscal response models tend to be sensitive to the way targets are modelled. Since in most cases fiscal targets are not readily available from government sources, they need to be approximated (estimated) using actual values. As White (1994) points out, if the 'fits' (R-squared) of the regressions are high, then each regressor in the structural equation will be actually regressed on itself. If the 'fits' are low, then the interpretation of the fitted values as proxies for the targets will become meaningless. Furthermore, White (1994) underlines the need to define targets that are consistent with the budget constraint. He suggests that once all the relevant targets are formed, the target for borrowing should fill the gap.

(iv) Treatment of aid. There is some discussion in relation to whether aid should be formulated as an exogenous or endogenous variable. In theory, this will depend on the recipient government's bargaining power. Most of the recent empirical studies opt for endogenising aid (by incorporating the aid variable in the utility function).

(v) Level of disaggregation. Some authors disaggregate aid flows (e.g. grants and loans) and government expenditure (e.g. developmental versus non-developmental) to provide more detail on fiscal dynamics. However, disaggregation comes at a cost: (a) increased number of parameters to estimate; and (b) greater complexity of the model. Furthermore, it is not always possible to obtain such disaggregated data that is consistent with the remaining sources (e.g. disaggregated aid statistics from DAC or CRS may not be consistent with budget accounts published by recipient governments or the IMF).

(vi) Estimation. The perceived lack of robustness of non-linear least squares estimation (3SLS) may raise questions about the reliability of the results. Furthermore, high correlation between the targets and actual values may lead to a 'near singular' matrix. This multicollinearity problem may render the estimation of the reduced form impossible (see White, 1994). Moreover, behavioural relationships may change over time (parameter instability) as a result of economic policy reforms. Hence, we could expect structural breaks to be present in the estimated relationships. Finally, the properties of the data (e.g. non-stationarity) and potential dynamic effects are not directly addressed – this is a static model in levels.

(vii) Interpretation. The interpretation of the parameters may be problematic. The rho associated with foreign aid (say, ρ_2) is usually interpreted as the extent of fungibility, although that assumes aid inflows are exclusively earmarked for investment purposes. In reality, donors also finance some types of recurrent spending (e.g. health staff salaries or road maintenance costs), which means that ρ_2 is a measure of 'maximum' fungibility (McGillivray and Morrissey, 2001a). Moreover, some studies obtain values for the rhos that are outside the [0, 1] range, which also raises interpretation problems (for example, a negative value).

Unit Root Tests

ADF Unit Root Tests

Variable	Lags	t-stat	Variable	Lags	t-stat
Ig	0	−2.91	ΔIg	0	7.79***
G	0	−2.80	ΔG	0	6. 28***
T	1	−2.74	ΔT	0	4.92***
A	0	−3.49*	ΔA	1	7.14***
A1	0	−1.55	ΔA1	0	6.97***
A2	1	−3.81**	ΔA2	1	6.49***
B	0	−4.67***	ΔB	0	8.37***
Y	0	−1.30	ΔY	0	6.97***
IP	1	−5.48***	ΔIP	0	12.32***
M	0	−2.35	ΔM	0	8.36***
X	0	−1.74	ΔX	0	5.07***

Obs.: All variables are in logarithms. The levels specification includes a time trend. The lag length was chosen according to the Bayesian Criterion (SIC) with a maximum of 9 lags. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

Maximum Likelihood

Total Impacts of Aid (Maximum Likelihood Estimation)

Impact	Ig	G	T	B
Total Aid (A*)	0.281	−0.122	−0.141	−0.325
Grants (A1*)	n/a	n/a	n/a	n/a
Loans (A2*)	n/a	n/a	n/a	n/a

Obs.: Starting parameter values were set to 0.1 since the initial estimation procedure did not find unique solutions. However, we were still not able to obtain unique estimates for the disaggregated model, even when including a time trend in the equations.

2.9 Bibliography

Abegaz, B. (2001). 'Ethiopia'. In S. Devarajan, D. Dollar, and T. Holmgren (eds), Aid and Reform in Africa, Washington D.C.: The World Bank.

Bevan, D. (2001). 'The Fiscal Dimensions of Ethiopia's Transition and Reconstruction'. WIDER Discussion Paper No. 2001/56.

Binh, T.-N. and M. McGillivray (1993), "Foreign Aid, Taxes and Public Investment: A Comment." *Journal of Development Economics*, 41(1), 173-176.

Cashel-Cordo, P. and S. Craig (1990). 'The Public Sector Impact of International Resource Transfers.' *Journal of Development Economics*, 32, 17-42.

Chole, E. (2004). Underdevelopment in Ethiopia. Organisation for Social Science Research in Eastern and Southern Africa, Addis Ababa.

Feyzioglu, T., V. Swaroop and M. Zhu (1998). 'A Panel Data Analysis of the Fungibility of Foreign Aid'. *World Bank Economic Review*, 12(1), 29-58.

Franco-Rodriquez, S. (2000). 'Recent Advances in Fiscal Response Models with an Application to Costa Rica'. *Journal of International Development*, 12(3), 429-442.

Franco-Rodriguez, S., M. McGillivray and O. Morrissey (1998). 'Aid and the Public Sector in Pakistan: Evidence with Endogenous Aid'. *World Development*, 26, 1241-1250.

Gang, I. and H. Khan (1999), "Foreign Aid and Fiscal Behaviour in a Bounded Rationality Model: Different Policy Regimes." *Empirical Economics*, 24, 121-134.

Gang, I. and H. Khan (1994), "Reply to Howard White: Foreign Aid, Taxes and Public Investment: A Further Comment." *Journal of Development Economics*, 45(1), 165-167.

Gang, I. and H. Khan (1991). 'Foreign Aid, Taxes and Public Investment.' *Journal of Development Economics*, 34, 355-369.

Geda, A. (2007). 'The Political Economy of Growth in Ethiopia'. In B. Ndulu, S. O'Connell, J.P. Azam, R. Bates, A. Fosu, J. W. Gunning and D. Njinkeu (eds), *The Political Economy of Economic Growth in Africa, 1960-2000 (Volume 2 – Country Case Studies)*, Cambridge University Press.

Geda, A. (2001). 'Macroeconomic Performance in Post-Derg Ethiopia'. *Northeast African Studies*, 8(1), 159-204.

Geda, A. (1996). 'The Fiscal Response to External Finance in Africa: An Alternative Approach'. Institute of Social Studies Working Paper, Sub-Series on Money, Finance and Development No. 61. The Hague.

Gupta, K. (1993). 'Sectoral Fungibility of Foreign Aid: Evidence from India.' Mimeograph, University of Alberta.

Heller, P. (1975). 'A Model of Public Fiscal Behaviour in Developing Countries: Aid, Investment and Taxation.' *American Economic Review*, 65, 429-445

IMF (2006). 'The Federal Democratic Republic of Ethiopia: Selected Issues and Statistical Appendix'. IMF Country Report No. 06/122. Washington D.C..

IMF (2005). 'The Macroeconomics of Managing Increased Aid Inflows: Experiences of Low-Income Countries and Policy Implications'. SM/05/306 (August 8, 2005), unpublished, Policy Development and Review Department, Washington D.C..

Iqbal, Z. (1997). 'Foreign Aid and the Public Sector: A Model of Fiscal Behaviour in Pakistan.' *Pakistan Development Review*, 36, 115-129.

Khan, H. and E. Hoshino (1992). 'Impact of Foreign Aid on the Fiscal Behaviour of LDC Governments.' *World Development*, 20, 1481-1488.

Khilji, N. and E. Zampelli (1991). 'The Fungibility of US Assistance to Developing Countries and the Impacts on Recipient Expenditures: A Case Study of Pakistan'. *World Development*, 19, 1095-1106.

Machado, R. (2009), 'Aid and Fiscal Policy in Nicaragua: A Fiscal Response Analysis.' *Journal of International Development*, published online (early view).

MacKinnon, J. (2003). 'How Does Aid Affect the Quality of Public Expenditure? What We Know and What We Do Not Know'. Background Paper for the *World Development Report 2004*, Washington D.C., World Bank.

Mavrotas, G. (2005). 'Aid Heterogeneity: Looking At Aid Effectiveness from a Different Angle'. *Journal of International Development*, 17, 1019-1036.

Mavrotas, G. and B. Ouattara (2006). 'Aid Disaggregation and the Public Sector in Aid-Recipient Economies: Some Evidence from Côte d'Ivoire'. *Review of Development Economics*, 10(3), 434-451.

McGillivray, M. (2000). 'Aid and Public Sector Behaviour in Developing Countries'. *Review of Development Economics*, 4(2), 156-163.

McGillivray, M. and B. Ouattara (2005). 'Aid, Debt Burden and Government Fiscal Behaviour in Côte d'Ivoire'. *Journal of African Economies*, 14(2), 247-269

McGillivray, M. and O. Morrissey (2001a). 'A Review of Evidence on the Fiscal Effects of Aid'. CREDIT Research Paper 01/13, University of Nottingham.

McGillivray, M. and O. Morrissey (2001b). 'Aid Illusion and Public Sector Fiscal Behaviour.' *Journal of Development Studies*, 37(6):118–136.

McGillivray, M. and O. Morrissey (2000). 'Aid Fungibility in Assessing Aid: Red Herring or True Concern?' *Journal of International Development*, 12(3):413-428.

McGillivray, M. and A. Ahmed (1999). 'Aid, Adjustment and Public Sector Fiscal Behaviour in the Philippines.' *Journal of the Asia-Pacific Economy*, 4, 381-391.

Mosley, P., J. Hudson and S. Horrell (1987). 'Aid, the Public Sector and the Market in Less Developed Economies'. *The Economic Journal*, 97, 616-641.

Osei, R., O. Morrissey and T. Lloyd (2005). 'The Fiscal Effects of Aid in Ghana'. *Journal of International Development*, 17, 1037-1053

Otim, S. (1996), 'Foreign Aid and Government Behaviour in Low-Income South Asian Countries'. *Applied Economics*, 28(8):927-933.

Ouattara, B. (2006a). 'Aid, Debt and Fiscal Policies in Senegal'. *Journal of International Development*, 18(8), 1105-1122.

Ouattara, B. (2006b), 'Foreign Aid and Government Fiscal Behaviour in Developing Countries: Panel Data Evidence.' *Economic Modelling*, 23(3):506-514.

Pack, H. and J. Pack (1993). 'Foreign Aid and the Question of Fungibility.' *Review of Economics and Statistics*, 75, 258-265.

Pack, H. and J. Pack (1990). 'Is Foreign Aid Fungible: The Case of Indonesia?' *Economic Journal*, 100, 188-194.

Pesaran, H., Y. Shin, and R. Smith (2001). 'Bounds Testing Approaches to the Analysis of Level Relationships'. *Journal of Applied Econometrics*, 16(3), Special Issue in Memory of John Denis Sargan, 1924-1996: Studies in Empirical Macroeconometrics, 289-326.

Pettersson, (2007). 'Foreign Sectoral Aid Fungibility, Growth and Poverty Reduction.' *Journal of International Development*, 19(8):1074-1098.

Rubino, C. (1997). 'Aid, the Public Sector and the Real Exchange Rate: The Case of Indonesia.' Unpublished PhD thesis, Department of Economics, University of Warwick.

Swaroop, V., S. Jha and A. Rajkumar (2000). 'Fiscal Effects of Foreign Aid in a Federal System of Governance: The Case of India'. *Journal of Public Economics*, 77, 307-330.

UNECA (2002). Economic Report on Africa 2003: Tracking Performance and Progress. United Nations Economic Commission for Africa, Addis Ababa.

White, H. (1994), "Foreign Aid, Taxes and Public Investment: A Further Comment." *Journal of Development Economics*, 45(1), 155-163.

White, H. (1992). 'The Macroeconomic Impact of Development Aid: A Critical Survey'. *Journal of Development Studies*, 28(2), 163-240.

World Development Indicators (WDI) (2007). 'World Development Indicators Database'. Washington D.C., World Bank.

World Bank (1998). 'Assessing Aid: What Works, What Doesn't, and Why'. Oxford University Press, New York.

3. Fiscal Dynamics in Ethiopia: The Cointegrated VAR Model with Quarterly Data

Abstract: This paper uses the cointegrated vector autoregressive (CVAR) model to assess the dynamic relationship between foreign aid inflows, public expenditure, government revenue, and domestic borrowing in Ethiopia. It departs from the existing literature in two main ways. Firstly, a unique quarterly fiscal dataset is compiled to study the impact of foreign aid flows on government behaviour. This approach has a number of advantages over its annual counterparts: (i) it allows for important intra-year dynamics; (ii) it provides a larger sample size; and (iii) it minimises the likelihood of structural breaks (especially those arising from regime changes) by focusing on the 1990s and 2000s. Secondly, the paper provides new insights into the formulation of testable hypotheses and the interpretation of structural shocks, while placing a strong focus on model specification. The econometric results suggest the presence of three long-run relationships: the government budget constraint, a donor disbursement rule, and a financing trade-off. While domestic fiscal variables adjust to budget imbalances, foreign aid grants adjust to the level of development spending, which can be seen as an indication of (procyclical) aid conditionality. Since earmarked grants have a contemporaneous impact on spending (by definition), this finding implies that budget support grants are provided after the government makes a strong commitment to poverty reduction. Moreover, domestic borrowing often compensates for lower levels of revenue and grants (substitutes), highlighting the cost of aid unpredictability and revenue volatility. The moving average representation of the CVAR suggests that unanticipated shocks to foreign aid grants do not have permanent effects on the remaining fiscal variables (as expected). There is also evidence of aid heterogeneity. The policy implication is that if foreign aid flows are to be made more effective, they should be provided in a predictable and countercyclical fashion in order to smooth exogenous shocks.

3.1 Introduction

Concerns about aid effectiveness are particularly relevant in the context of the least developed countries in Africa, due to their chronic dependence on foreign aid. Since most foreign aid flows are provided directly to recipient governments, it is natural that the impact of aid on fiscal behaviour emerges as a critical question. The main purpose of this paper is to assess how aid inflows impact on the allocation of public resources, domestic revenues and borrowing requirements in Ethiopia.

Ethiopia was chosen due to the vast aid resources it receives (currently and historically), and also due to the availability and quality of fiscal data. An extensive fiscal dataset was compiled from several Ethiopian official sources, covering the period from 1993Q3 to 2008Q2. To our knowledge, this is the first fiscal response study to use quarterly data, which provides an interesting complement to the existing empirical evidence using annual data. Moreover, this study follows recent developments in the fiscal response literature by using the cointegrated VAR (CVAR) model to capture the rich dynamics of fiscal aggregates. In this context, a number of relevant hypotheses regarding the impact of aid flows can be formulated and tested (see Chapter 2.1):

(a) Additionality. Do aid flows induce an equivalent increase in government expenditure?
(b) Aid illusion. Do aid flows induce expenditure increases larger than the amount of aid?
(c) Fungibility. Do aid flows shift the expenditure pattern towards (non-development) recurrent costs?
(d) Tax displacement. Do aid flows reduce the government's 'tax effort'?
(e) Deficit financing. Are aid flows associated with falls in domestic borrowing?
(f) Spending. Do aid flows widen the non-aid fiscal deficit?
(g) Aid Heterogeneity. Do aid grants and loans induce similar fiscal dynamics?

This paper is divided into eight main sections. After this short introduction, section two provides a brief overview of the literature on the fiscal effects of aid. Section three introduces the methodology – the cointegrated vector autoregressive (CVAR) model. Section four presents the data that will be used for this empirical exercise, with some preliminary analysis of the main fiscal trends. Section five reports the main empirical results, whereas section six addresses the identification issue and performs impulse response functions. Section seven undertakes robustness tests, while section eight concludes the paper.

3.2 Literature Review

In the previous chapter we distinguished between two interrelated but competing methodologies. One is the so called 'categorical' fungibility literature, which is thoroughly surveyed in McGillivray and Morrissey (2001). Studies that fall in this category are restricted to the observation of the impact of aid on the composition of government spending. For this purpose, expenditure data is collected for several sectors such as health and education, and the extent of fungibility is estimated. However, this approach distracts attention from the broader fiscal impacts of aid, such as those on borrowing or taxation, which tend to be a more fundamental issue (McGillivray and Morrissey 2000). Indeed, these studies assume domestic revenue to be a residual, not allowing aid to influence explicitly the tax effort or borrowing, providing only a partial insight into fiscal behaviour.

The 'fiscal response' literature incorporates these concerns by focusing on the broader picture. Empirical work in this vein assesses 'aggregate' fungibility, i.e. the extent to which foreign aid dynamics induce perverse expenditure patterns (recurrent over capital spending), reduce the tax effort, and increase the reliance on domestic borrowing. It is worth outlining the standard theoretical framework that supports 'traditional' fiscal response models, such as those embodied in Heller (1975), Mosley (1987), Franco-Rodriguez et al (1998) and Mavrotas and Ouattara (2006). These studies start by assuming that public sector decision-makers are rational and make well-ordered (consistent) budgetary choices. They allocate state revenue to different expenditure categories by maximising the government's utility function subject to a budget constraint. This utility framework is modelled as deviations from government 'targets', which are defined by the government *ex-ante* for each fiscal variable (revenues and expenditures). Since this utility function is often assumed to be quadratic (symmetric), overshooting or undershooting a target entails the same disutility. Utility is maximised when all targets are achieved, i.e. when there is no *ex-post* difference between actual (observed) and desired (budgeted) levels of revenue, expenditure, borrowing and foreign aid. A set of reduced form equations is then derived from the maximisation problem and estimated simultaneously, often by three-stage least squares (3SLS). The previous chapter summarised and discussed the results of several models in this tradition.

Nonetheless, there have been several criticisms of this methodology, most of them related to its inherently strong assumptions: the specification of the utility function, the specification of the constraints, the use and estimation of targets, the treatment of aid, the interpretability of some parameters, and the robustness of 3SLS estimation.

In order to overcome many of these difficulties, Osei et al (2005), Fägernas and Roberts (2004c), and M'Amanja et al (2005) have used vector autoregressive (VAR) models to analyse the relationship between foreign aid and fiscal aggregates. Moreover, 'impulse response functions' are estimated in order to investigate the dynamic effects of aid on fiscal variables. This is a clear departure from the utility maximisation framework, which presents a number of advantages. The VAR approach is an empirical approach aimed at capturing the 'data generating process' (DGP). Although at first it may seem an atheoretical approach, economic theory is often invoked to choose the variables to include in the analysis, to help identification of the system, and to assist in interpreting the results. Little economic theory is imposed directly as it avoids making strong *a priori* assumptions. Since empirical and theoretical predictions are often ambiguous about how fiscal variables are determined, this seems to be a reasonable starting point.[40]

In the unrestricted VAR specification all variables are assumed to be endogenous (there is one equation for each and every variable), avoiding unnecessary distinctions between endogenous and exogenous variables. The fact that it does not assume an *a priori* direction of causality among the variables is particularly useful for fiscal variables, which are often jointly determined. Instead, the framework allows a number of hypotheses to be tested within the specified model. This framework is often used to help the formulation of realistic models, uncovering facts and describing the characteristics of the data.

Table 3-1: Comparison of Methodologies

Methodology	Strengths	Weaknesses
Utility Maximisation	Provides a theoretical base (structure).	Requires strong *a priori* assumptions; Needs to define and estimate targets; Does not allow for cointegration analysis, structural breaks or outliers.
CVAR model	Avoids strong assumptions (government behaviour) and target estimation; Very informative about the data.	Needs long time series (degrees-of-freedom problems); Robustness concerns (especially with short samples).

Source: Author's compilation.

The VAR model can be easily re-written in an error-correction representation (VECM), which allows the researcher to analyse non-stationary data. The framework handles both I(1) and I(0) variables, so there is no need to test whether individual variables are stationary or not – each stationary variable will be associated with an additional cointegrating vector. It is also possible to detect the presence of I(2) variables within the model. The implications of neglecting non-stationarity are now well known (spurious inference), and this framework takes a strong concern with the properties of the data. It is thus possible to investigate both the long-run static relationships (cointegrating vectors) and the dynamics of the variables in the system. Moreover, it allows the examination of several cointegrating relations. The single-equation error correction models (ECM) are restricted to one long-run relation between the variables of interest, and require strong exogeneity assumptions

[40] The main argument is that it is not feasible to assume that governments' behaviour can be adequately characterised by a utility maximisation framework.

on all but one variable (Juselius, 2007:82). These can be seen as a special case of the more flexible VECM methodology.

Nevertheless, the VAR approach also has some drawbacks. The methodology requires a relatively long time series (demanding data requirements) or the use of a small set of variables, in order to avoid a quick erosion of degrees-of-freedom. Moreover, VAR models are inherently over-parameterised, while results may be sensitive to model specification (variables included), sample size and the choice of the lag length. Nonetheless, there are several tests that can be used as robustness checks and evaluate model stability. Finally, the estimation methodology requires that we exclude one fiscal component from the system of equations (to avoid estimating an identity), with potential loss of information.[41]

The table below presents the cointegrating relations found in this literature. All studies find a single cointegrating relation, except M'Amanja et al (2005), which adds GDP to the fiscal system, and Fägernas and Schurich (2004). The cointegrating relations found often resemble the budget constraint, since the expenditure variables have opposite signs in relation to the revenue and financing variables. However, there are some exceptions, such as the last three models in Fägernas and Roberts (2004a). Moreover, we would expect all long-run coefficients to be (close to) 1.[42] The budget constraint does not clearly emerge as a cointegrating relation possibly due to: (i) omitted variable bias, since one fiscal variable is left out; (ii) misspecification of the deterministic terms (a time trend might be required in the levels); (iii) presence of structural breaks in the data; or (iv) use of inconsistent aid data (e.g. OECD-DAC).

[41] For example, Osei et al (2005) exclude non-tax revenue and external borrowing from the estimation.
[42] We will demonstrate that this is possible if we use all 'observed' fiscal variables and exclude 'net errors and omissions', which are usually derived as a residual from the budget identity. Hence, this relationship can also be seen as a consistency check of the data. 'Net errors and omissions' are likely to be stationary.

Table 3-2: Cointegrating Relations in the Literature

Study	Model (CI#)	Cointegrating Relation(s)
Osei et al (2005)	I	$D + 0.07A + 1.04R - 0.84GT - 0.47 \approx 0$
	II	$D + 0.05A + 1.07R - 0.72GC - 1.20GK - 0.61 \approx 0$
M'Amanja et al (2005)	I(1)	$Y - 0.39GT - 0.02G + 0.10L \approx 0$
	I(2)	$L - 3.90GT + 2.29R + 0.04G \approx 0$
Fägernas & Schurich (2004)	I	$D + 0.39R + 2.68G - 1.57GK - 0.91GC + 1198 \approx 0$
	II(1)	$GC - 1.46R - 0.13D + 0.34GK + 343 \approx 0$
	II(2)	$L - 0.11R + 0.44D - 0.27GK - 163 \approx 0$
	III	$A + 1.75D - 2.16R + GK - 0.1GC + 830 \approx 0$
Fägernas & Roberts (2004a)	I	$G + 0.65D + 0.38R - 0.46GT + 23.1 \approx 0$
	II	$L + 0.35D + 0.63R - 0.53GT - 29.1 \approx 0$
	III	$G - 0.81D - 0.98GK + 0.35GC - 0.38R + 628 \approx 0$
	IV	$L + 1.91D + 0.19GK - 0.61GC + 1.06R - 1802 \approx 0$
	V	$A + 3.19R - 3.1GT - 0.37D + 1843 \approx 0$
Fägernas & Roberts (2004b)	n/a	Stationary variables (no cointegration analysis undertaken)

Obs.: 'D' Domestic Borrowing, 'A' Total Aid, 'G' Grants, 'L' Foreign Loans, 'R' Domestic Revenue, 'GT' Government Total Expenditure, 'GC' Government Current Expenditure, 'GK' Government Capital/Development Expenditure, and 'Y' GDP per capita.
Source: Author's compilation.

In isolation, these cointegrating relations do not provide information about fiscal dynamics. To shed light on the impact of aid inflows, impulse response functions are often estimated (see table below).[43] Osei et al (2005) suggest that foreign aid to Ghana does not have a direct effect on the volume of government spending, but is treated as a substitute for domestic borrowing. Government spending does rise significantly following aid, but this is principally due to an indirect effect arising from higher tax revenues associated with aid inflows. Hence, aid to Ghana has tended to be associated with reduced domestic borrowing and increased tax effort, combining to increase public spending. M'Amanja et al (2005) extend the fiscal response framework by adding growth of per capita income in the analysis. They find that aid grants appear to have a positive effect on long-run growth, while loans seem to substitute for taxes and finance fiscal deficits, hence having a negative effect on growth. Government spending is found to have a positive long-run influence on growth, while tax revenue has no significant direct effect (but may have an indirect effect through expenditure). The authors conclude that foreign aid to Kenya could be more effective if given in the form of grants, and associated with fiscal discipline. Fägernas and Roberts (2004c) summarise fiscal response studies for three African countries (Malawi, Uganda and Zambia). Foreign aid flows seem to have a strong positive correlation with the development budget of the three countries studied. The other fiscal effects vary according to the country under analysis. In Zambia, aid flows displace tax revenues, have a moder-

[43] The VAR methodology does not provide numerical results that are directly comparable with standard fiscal response models (e.g. total effects). Hence, the impact of aid is assessed through impulse response functions, which show the effects of a shock (e.g. increase in aid) on the entire fiscal system. However, the identification and interpretability of these shocks tends to be a contentious issue.

ately positive impact on the recurrent budget, and are associated with higher levels of domestic borrowing. In Malawi, aid is correlated with lower recurrent budget and consequently lower domestic borrowing. Finally, in Uganda, aid raises both development and recurrent spending, with a negligible impact on domestic borrowing.

Table 3-3: Results from VAR-Based Models

Study	Country	Aid	Capital Spending	Recurrent Spending	Domestic Revenue	Domestic Borrowing
Osei et al (2005)	Ghana	ODA	+	++	++	––
M'Amanja et al (2005)	Kenya	Grants	n/a	n/a	n/a	n/a
		Loans		–	––	n/a
Fägernas & Schurich (2004)	Malawi	Grants	++	––	+	––
		Loans	+	?	+	––
		ODA	++	––	+	––
Fägernas & Roberts (2004a)	Uganda	Grants	++	+	+	0
		Loans	++	++	+	0
		ODA		++	+	0
Fägernas & Roberts (2004b)	Zambia	Grants	++	+	––	+
		Loans	+	+	––	0
		ODA	++	+	––	+

Source: Fägernas and Roberts (2004c:33), M'Amanja et al (2005), and Osei et al (2005).
Notes: ++ strongly positive; + moderately positive; ? ambiguous; 0 insignificant; – moderately negative; –– strongly negative. All the results are obtained through the use of generalised impulse response functions (GIRF), as described in Pesaran and Shin (1998).

In conclusion, it is difficult to find a consistent pattern regarding the impact of aid on public fiscal accounts. The empirical evidence (and theoretical predictions) regarding the impact of foreign aid on fiscal policy is ambiguous, which strengthens the argument that results tend to be country-specific, either because economic circumstances are different or simply because governments behave differently.

3.3 Methodology

The vector autoregressive (VAR) model is a multivariate time series specification developed as a generalisation of the univariate autoregressive (AR) model. It was initially proposed by Sims (1980) to avoid the 'incredible identification restrictions' of (large scale) structural econometric models and it has since become an important tool in empirical macroeconometrics. Following Engle and Granger's (1987) seminal work on the non-stationarity of variables, which has dramatically shaped modern time series econometrics, Johansen and Juselius (1990, 1992) extend the VAR model by applying the concepts of cointegration and error-correction to analyse long-run relations amongst non-stationary variables. This methodology is often known as the vector error-correction model (VECM) or as the cointegrated VAR (CVAR) model.[44]

In the usual unrestricted VAR specification, there is one equation for each and every variable. Therefore, all variables are assumed to be endogenous, which avoids unnecessary *a priori* distinctions between endogenous and exogenous variables. Any assumptions regarding endogeneity and causal effects can be tested (and therefore substantiated) within the VAR framework. Moreover, for each endogenous variable there is a set of explanatory variables that comprise its own lags and lags of all the other variables in the model, allowing for rich dynamic effects to be captured. In the unrestricted form, all the variables in the system are treated symmetrically in the sense that they have precisely the same set of regressors. In the following paragraphs we will explain the basic characteristics of the reduced form VAR and the VECM representation. Then we take a step back and present the structural form of the model.

Consider the following reduced form of a k^{th} order vector autoregressive model with p variables, i.e. a p-dimensional VAR(k):

$$x_t = \sum_{i=1}^{k} \Pi_i x_{t-i} + \phi D_t + \varepsilon_t$$

where x_t is a $p \times 1$ vector of endogenous variables with $t = 1, 2, \ldots, T$; Π_i are $p \times p$ matrices of parameters (i.e. coefficients to be estimated) with $i = 1, 2, \ldots, k$; D_t is a vector of deterministic components (e.g. intercept, trend and dummy variables), with a vector of coefficients ϕ; and ε_t is a $p \times 1$ vector of (unobservable) error terms. The VAR(k) model is linear in the parameters and assumes that these are constant over time. Moreover, we assume that the error terms are identically and independently distributed, i.e. they are serially uncorrelated ($E(\varepsilon_t\ \varepsilon'_{t-k})=0$ for $k \neq 0$), have zero mean ($E(\varepsilon_t)=0$), and have a time-invariant positive definite covariance matrix ($E(\varepsilon_t\ \varepsilon'_t)=\Omega$). Hence, the error terms follow a Gaussian (normal) distribution (or white-noise process): $\varepsilon_t \sim$ iid $N_p(0,\Omega)$. The residual covariance matrix (Ω) has dimensions $p \times p$, and contains information about possible

[44] This paper is greatly influenced by Juselius (2007), which presents a methodical empirical approach to the specification, estimation, testing, and interpretation of the CVAR. The book is based on Søren Johansen and Katarina Juselius's extensive work on cointegration analysis for systems of equations. Johansen (1996) provides a detailed mathematical and statistical analysis of the CVAR, including the derivations of the estimators and several test statistics.

contemporaneous effects.[45] These assumptions are "consistent with economic agents who are rational in the sense that they do not make systematic errors when they make plans for time t based on the available information at time $t - 1$" (Juselius, 2007:46). To be able to make reliable statistical (and economic) inference it is important that these properties are satisfied. For this purpose, there are a number of misspecification tests that can be used to evaluate whether these assumptions are reasonably valid.

VAR processes are a suitable class of models for describing the data generating process (DGP) of a small set of variables (Lütkepohl and Krätzig, 2004:86), since they are essentially a reformulation of the covariances of the data (Juselius, 2007:46). Its dynamic and stability properties can be investigated by looking at the roots of the process. These can be obtained in two different ways: (i) by reformulating the VAR into the 'companion AR(1) form' and then solving for the eigenvalue problem; or (ii) by solving the characteristic (polynomial) function. The first method calculates the roots directly as $\rho_1, \rho_{2,\ldots,} \rho_{p\times k}$ (eigenvalue roots), while the second method computes the characteristic roots (z), which are the inverse of the former ($z_j = \rho_j^{-1}$).

In terms of the dynamic and stability behaviour of the process, a real root inside the unit circle ($|\rho_j| < 1$) will generate exponential declining behaviour. If the modulus of a complex pair of roots ($|\rho_j| = |\rho_{\text{real}} \pm i\rho_{\text{complex}}|$) is inside the unit circle, then it will generate exponentially declining cyclical behaviour. Both types of roots are compatible with stationary processes (i.e. all variables are stationary), albeit with different dynamics. In addition, a real root lying on the unit circle ($|\rho_j| = 1$) will generate non-stationary behaviour, while if the modulus of a complex pair of roots is one, it generates non-stationary seasonal behaviour (Juselius, 2007:49). Finally, if any of the roots lies outside the unit circle ($|\rho_j| > 1$), we have an explosive process.

If the process is found to contain non-stationary behaviour (at least one variable is non-stationary), then inference based on the VAR may be invalid and the relationships among the variables spurious. In this case, it will be more appropriate to analyse the data within a cointegration framework. For this purpose, the VAR can be re-written in the general VECM(k-1) form:[46]

---

[45] "The VAR model is often called a 'reduced form' model because it describes the variation in x_t as a function of lagged values of the process, but not of current values. This means that all information about current effects in the data is contained in the residual covariance matrix Ω." (Juselius, 2007:66)

[46] D_t can also include stationary stochastic variables that are weakly exogenous or that can be excluded from the cointegrating space (Dennis, 2006:3).

$$\Delta x_t = \Pi x_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi D_t + \varepsilon_t$$

where,

$$\Pi = \sum_{i=1}^{k} \Pi_i - I_p \; , \quad \Gamma_i = -\sum_{j=i+1}^{k} \Pi_j \quad \text{with} \;\; i = 1, \dots, k-1 \;\; \text{and} \;\; \varepsilon_t \sim IN(0, \Omega)$$

The VECM is a tractable formulation that, due to the clear separation between the long-run (Π) and the short-run (Γ_i) effects, allows for an intuitive interpretation of the estimates. Moreover, the specification reduces the multicollinearity effect, since the first differences of the variables tend to be more 'orthogonal' than the levels. It should be noted that the reformulation of the VAR model as a VECM does not impose any binding restrictions on the original parameters (Juselius, 2007:60-1). Therefore, the value of the maximised likelihood function remains the same and there is a direct correspondence between the estimated parameters of both forms (as shown above).

As before, the dynamic properties of the VECM can be analysed via the roots of the characteristic polynomial (by setting $A(z) = 0$):

$$A(z) = (1 - z)I_p - \Pi z - \sum_{i=1}^{k-1} \Gamma_i (1 - z)$$

If the process in levels (x_t) is non-stationary and integrated of order one, then the first-differenced process (Δx_t) is stationary. However, this reformulation still leaves us with some non-stationary variables (those related to Π). If we have at least one unit root (z = 1), then Π must have reduced rank ($r < p$), because $|A(1)| = |\Pi| = 0$.[47] Consequently, Π can be written as $\Pi = \alpha\beta'$, where α and β are $p \times r$ matrices of full column rank, and the hypothesis of cointegration can be formulated as a reduced rank condition on the Π matrix (Dennis, 2006:3). Thus, we can interpret the relations $\beta' x_t$ as stationary relations between non-stationary variables. It should also be noted that stationary variables are by themselves a cointegrating (stationary) relation, since they are associated with a unit vector in

[47] This means that Π is a singular matrix (i.e. non-invertible).

β.[48] Therefore, adding a stationary variable to the system will increase the cointegrating rank (r) by one. These variables can play an important role for the long-run relations, especially if they have a high degree of autocorrelation (i.e. are near-integrated). If $r = 1$, the (unique) stationary relation can be easily interpreted as the long-run equilibrium for the levels data. If $r > 1$, then we have an identification problem because it is the space spanned by β that is identified (uniquely determined) and not β itself (Dennis, 2006:4-5).

Moreover, the VECM specification is not particularly useful in the following extreme cases: (i) if $r = 0$, the variables are non-stationary but there is no linear combination that is I(0); and (ii) if $r = p$ (i.e. Π has full rank), all the variables in the system are stationary. In the latter case, the researcher can make inference with the levels model (VAR), whilst in the former, a VAR with first-differenced variables can be used to analyse the short-term relations, i.e. setting $\Pi = 0$ (since there are no long-run relations between the variables). Finally, there are other less interesting cases, even when $0 < r < p$. It can be the case that all variables but one are I(0) ($r = p - 1$), or that we have a system with $p - r$ unrelated I(1) variables and r I(0) variables. In these cases, no cointegration in the original sense is present (Lütkepohl and Krätzig, 2004:90).

Nonetheless, a unit root is often a convenient statistical approximation (e.g. of an exact root of 0.9), which enable us to utilise a much richer framework (VECM) that distinguishes between the longer and shorter term dynamic effects.[49] It is therefore useful to consider unit roots for the empirical analysis of macroeconomic relationships. Moreover, neglecting a unit root when there is some non-stationary behaviour may invalidate the empirical analysis.[50]

> "In such cases [variables with significant mean reversion] it is nevertheless often advantageous to use the unit-root approximation for the purpose of modelling (even though the root is statistically significant from unity) as this allows us to use cointegration techniques to structure the data into persistent and less persistent components." (Juselius, 2007:20)

---

[48] In systems where both I(1) and I(0) variables are considered, the original definition of cointegration is extended so that any linear combination that is stationary is called a cointegration relation, even between stationary variables (Lütkepohl and Krätzig, 2004:86).

[49] "In practice, it is useful to classify variables exhibiting a high degree of time persistence (insignificant mean reversion) as non-stationary and variables exhibiting a significant tendency to mean reversion as stationary. However, it is important to stress that the stationarity/non-stationarity or, alternatively, the order of integration of a variable, is not in general *a property of an economic variable* but a convenient statistical approximation to distinguish between the short-run, medium-run, and long-run variation in the data." (Juselius, 2007:18)

[50] The non-stationarity property of a variable (presence of a unit root) is not only dependent on the modelling approach (e.g. AR(k), VAR(k), etc.) but also on the time frame of the analysis. It may be the case that a variable looks stationary over a long time period, while a sub-sample of the same variable exhibits less mean reversion and is therefore considered non-stationary.

The most usual estimation method for the VECM presented above is the maximum likelihood estimator (MLE) proposed by Johansen (1996), which uses the reduced rank regression (RRR). In Johansen's approach, the "parameter estimator $\hat{\beta}$ is made unique by the normalisation of the eigenvectors, and $\hat{\alpha}$ is adjusted accordingly" (Lütkepohl and Krätzig, 2004:98). However, when $r > 1$, only the cointegration space ($\Pi = \alpha\beta'$), and not the cointegration parameters (α and β), is estimated consistently. Therefore, appropriate identifying restrictions need to be imposed. Unfortunately, most software packages (e.g. EViews and JMulTi) automatically normalise the cointegration matrix as:

$$\beta = \begin{bmatrix} I_r \\ \beta(p-r) \end{bmatrix}$$

Although this scheme enables the identification of the system, it is highly restrictive and may invalidate inference. For example, one needs to be careful with the order of the variables. If a variable is not part of the cointegrating relations, normalising on this variable may result in dividing the remaining coefficients by zero. Moreover, the zero restrictions imposed by the identity matrix (I_r) when $r > 1$ may not be acceptable and cannot be tested.

Under some basic assumptions,[51] the VAR model has a moving average (MA) or Granger representation. This means that the process (x_t) can be re-written as a function of the innovations of the system. This is a useful reformulation since it facilitates the investigation of the common stochastic trends, which are responsible for the non-stationarity of the process. The MA representation of the VAR is given by:

$$x_t = C\sum_{i=1}^{t}(\varepsilon_i + \Phi D_i) + C^*(L)(\varepsilon_t + \Phi D_t) + X_0 \qquad t = 1, 2, \ldots, T$$

where,

$$C = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}\alpha'_\perp \quad \text{and} \quad C^*(L) = \sum_{i=0}^{\infty} C_i^* L^i$$

[51] For example, the absence of explosive roots (all roots must be inside, or lie on, the unit circle).

In the representation above, C is the long-run impact matrix, which has reduced rank (p – r) and relates to the stochastic part of the process (cumulation of errors). Therefore, it indicates that only p – r of the linear combinations of the p innovations (ε_t) have permanent effects. C^* is a convergent matrix polynomial in the lag operator (L), and thus relates to the stationary part of the process. X_0 contains the initial values. The common trends are given by:

$$\alpha'_{\perp}\sum_{i=1}^{t}\varepsilon_i$$

These are often called the 'pushing' forces, since they push the system away from the long-run (steady-state) equilibrium. Note that $\alpha_{\perp}$ contains "the vectors that define the space of the common stochastic trends and the slopes of the linear trends in the data" (Dennis, 2006:6). In contrast, the 'pulling' forces relate to the equilibrium correction (with adjustment speed α) that is activated as soon as the process is out of steady-state cointegration relations (Juselius, 2007:88-90).

As it was mentioned before, the VAR model is a powerful tool to summarise the properties of the data. Nonetheless, the estimated parameters cannot be given a meaningful economic interpretation because the estimated VAR is essentially a reduced form model. This means that all right-hand-side variables are either predetermined or exogenous, and any contemporaneous effects present in the data will be captured by the residual covariance matrix (Ω). This reflects the fact that the underlying structural form VAR (from which the reduced form was obtained) had to be solved for the endogenous variables to enable estimation. In many ways, the VAR approach suffers from the same estimation and identification problems facing (traditional) structural macroeconometric models. In fact, several economic models can be seen as a special case of the more general class of models, the unrestricted SVAR.

The challenge, therefore, is to recover the information about the structural parameters, which will enable us to investigate important economic questions. By premultiplying the reduced-form model by a matrix (A_0), we obtain the structural VECM (SVECM):

$$A_0\Delta x_t = \mathrm{a}\beta' x_{t-1} + \sum_{j=1}^{k-1} A_j\Delta x_{t-j} + \widetilde{\Phi} D_t + v_t \qquad t = 1, 2, \ldots, T \qquad v_t \sim IN(0, \Sigma)$$

where A_0 is a non-singular $p \times p$ matrix, which if set to be an identity matrix gives the reduced form model. We can relate these coefficient matrices with the ones from the reduced form in the following way: $a = \alpha A_0^{-1}$, $A_j = \Gamma_j A_0^{-1}$, $\widetilde{\Phi} = \Phi A_0^{-1}$, and $\Omega = A_0^{-1} \sum A_0'^{-1}$. We are now able to isolate the contemporaneous effects in matrix A_0, while the vector v_t contains error terms associated with 'structural' shocks. In order for these shocks to have an economic interpretation they ought to be mutually uncorrelated (i.e. orthogonal) so that we can isolate/identify its dynamic impact through the system. If these shocks are correlated, then we need to take into consideration the relationship between the shocks (Lütkepohl and Krätzig, 2004:161). Structural shocks are related to their reduced form counterparts through matrix A_0. To solve the identification problem, restrictions may need to be imposed on A_0, the long-run structure (α and β), the short-term structure (Γ_j), and/or the covariance matrix (Ω).

3.4 Data

Raw Data

This paper uses a fiscal dataset comprised of 60 quarterly observations, covering the period from 1993Q3 to 2008Q2. The dataset is arranged in a way that the first quarter refers to the first three months of the Gregorian calendar (i.e. January to March), and so on. Since the Ethiopian fiscal year starts in July, it runs from the third quarter to the second in this dataset. The use of quarterly data offers several advantages over studies that employ annual variables. Firstly, it allows for important intra-year dynamics. Fiscal decisions are taken throughout the year and are often based on (preliminary) monthly and quarterly information. For example, the government may need to increase domestic borrowing unexpectedly due to donors failing to disburse committed funds earlier in the year. Therefore, quarterly data will be better suited to capture the rich dynamic pattern of the decision-making process than the aggregate yearly data, which often contain large contemporaneous effects that complicate the analysis and the interpretation of results. However, since the concept of the long-run depends on the frequency and time-span of the dataset, this study may not be directly comparable to its annual counterparts.

Secondly, it tries to mitigate the problem of vanishing degrees of freedom in the VAR model by substantially increasing the sample size. Studies on the fiscal response to aid in

Africa are usually constrained to about 30 (yearly) observations.[52] This can be partly explained by the fact that most countries in Africa did not reach independence before the mid-1960s. Moreover, there are some concerns about the quality of the fiscal data produced during the 1970s and 1980s by most African countries, which may further undermine a rigorous fiscal analysis. These concerns range from weak budget recording capacity, lack of transparency in budget implementation, different conceptual notions on the categorisation of items, etc. To some extent, some of these issues may still be valid today, but it is undeniable that most African countries have made strong progress in improving their public finance management processes since the 1990s.

Table 3-4: Sample Size and Data Sources

Paper	Country	Sample	Obs.	Data Sources
Osei et al (2005)	Ghana	1966-98	33	IMF-IFS and OECD-DAC
M'Amanja et al (2005)	Kenya	1964-02	39	National
Fägernas & Schurich (2004)	Malawi	1970-00	31	Several: IMF-GFS, WDI, OECD-DAC and National.
Fägernas & Roberts (2004a)	Uganda	1974-99	26	
Fägernas & Roberts (2004b)	Zambia	1972-98	27	

Source: Author's compilation.

Thirdly, this approach seeks to improve the robustness of empirical results by focusing on a period that has been relatively stable in terms of economic policy. Thus, the model is less likely to be affected by significant regime changes, which in turn affect the constancy of the estimated parameters (a crucial assumption in most regression models). Most African countries have experienced periods when socialist economic policies were promoted,[53] with clear impacts on public expenditure and revenue dynamics – both in terms of volume and composition. In the particular case of Ethiopia, by focusing on the post-1993 period we avoid dealing with the potential structural breaks in the data arising from the Derg regime period (1974-1991), during which a specific set of policies were promoted that differed in scope and nature from the ones currently implemented by the Ethiopian People's Revolutionary Democratic Front (EPRDF) government since 1991-1992. Hence, we hope that the economic relationships found in the quarterly fiscal data are significantly more robust than its annual counterparts. It is also worth mentioning that most studies fail to report robustness checks such as tests for the constancy of estimated parameters.

Finally, our approach favours the use of local sources, since these form the 'information set' that shapes government decision-making (i.e. fiscal management and planning). Several studies on this topic use OECD-DAC statistics on foreign aid, mainly due the ready

[52] The problem of small (annual) samples is compounded by the lack of dynamic information contained in such aggregate data.
[53] See Ndulu et al (2007:90).

availability of such data and the level of disaggregation. However, this may cause a number of serious problems. This is a donor measure, based on questionnaires filled by DAC member countries,[54] which is likely to be inflated for a number of reasons: inclusion of technical assistance, emergency food-aid, and donor-implemented projects. In most cases, these funds do not pass through the central treasury and are not reported in the fiscal budget (this is why they are sometimes called 'off-budgets'). Hence, the central government does not have information on some of these activities (often undertaken at the sub-national level), which are therefore not likely to influence central fiscal decisions.

> "Aid may or may not go through budget. Certainly our work at the micro level data in the sectors confirms that there are large amounts of aid that do not get recorded. Even when aid is shown in the budget estimates, actual spending is often not reported. In general, loan disbursements are reflected in the sectoral expenditure reports; grant expenditures may in some cases be reported by the donor to the Region, to the sectoral Ministry, or to MoFED, but they are not consistently reported, and even when they are, there is not a systematic process for consolidating those expenditures. This, as much as problems of physical implementation and procedural bottlenecks in using donor funds, appears to account for under-spending of aid resources." (World Bank, 2004:31)

In the previous chapter we suggested that there can be substantial discrepancies between data sources. OECD-DAC grant figures were, on average, 300 percent higher than government/IMF reported data during 1993-2005. This problem does not seem to be particularly worrying for aid loans, but since grants considerably outweigh concessional loans in most African countries, this becomes a very serious problem. Since aid is a crucial variable in the analysis, using an imperfect proxy may compromise the estimation of the long-run relations and the dynamic results of the impulse responses. Moreover, the OECD-DAC reports aid flows in calendar years, which is not consistent with most recipients' fiscal year (e.g. Ethiopia's fiscal year runs from the 8th July to the 7th July). It is not clear how most studies resolve the time-inconsistency of merging these two sources. Finally, most studies also use IMF-reported data (from the IFS or GFS databases). Although less problematic, there might still be some discrepancies between the data provided by local sources (e.g. Ministry of Finance) and the data published by the IMF. The reason is that IMF staff are likely to 'treat' the data originally reported by the country in order to meet international standards and allow cross-country comparability. Moreover, data is often published as a budget identity, where residual items such as 'errors & omissions' are

[54] Non-DAC members are not obliged to fill the questionnaire and therefore there may be some under-reporting of the total ODA provided to a recipient country (as a whole). For example, support from the ex-USSR is often not accounted for in these statistics, for which amounts can be substantial during the cold war years.

incorporated in domestic financing. This approach can undermine statistical inference, and forces researchers to drop one important budget item to avoid estimating an identity.

The data used in this study was provided by the Ministry of Finance and Economic Development (MoFED) in Ethiopia, and is complemented by secondary local sources, namely from the National Bank of Ethiopia (NBE) and the Ethiopian Economic Policy Research Institute (EEPRI).[55] Before analysing the main trends in the data, we need to be aware of the limitations of this dataset. Government fiscal data are produced by the MoFED on a quarterly and yearly basis. The preliminary data are usually released with a short time lag (less than six months), and then updated with revised figures at a later stage.[56] These revisions often arise from the late reporting of items from regional governments and the resolution of discrepancies. However, there might be a significant lag from the time of release of the preliminary figures ('pre-actuals') and the closure of accounts ('actuals'), which in practice means that some revisions will only be incorporated in the yearly data, creating a discrepancy between the quarterly and annual figures. While in theory these discrepancies could be considerable, our consistency checks suggest that the difference between the aggregated quarterly values and the yearly 'actuals' (in particular for revenue and expenditure items) are not significant. However, minor problems may arise for deficit financing, namely, foreign loans and domestic borrowing. The analysis will therefore pay careful attention to these two variables. Moreover, as long as the data (even if preliminary in some periods) capture the broad dynamics of the fiscal sector, our dataset will certainly contain relevant information that is worthwhile to be analysed by the statistical model. One may also suggest that 'real-time' data are likely to be more informative about fiscal behaviour than the final estimates, since policy-makers often base their decisions on preliminary figures.

The table below presents trends of the main fiscal variables as a percentage of GDP (3-year averages). In terms of government receipts, both domestic revenue and foreign grants have consistently grown as a percentage of GDP, except in the last period. This can be partly explained by the strong GDP growth experienced in the last few years. We can also observe that while current expenditure has fallen in the last two periods, capital spending has shown a strong increase. Finally, the financing of the deficit has shifted from a reliance on external borrowing to lending from domestic sources.

---

[55] The latter two sources record older quarterly data (1993Q3 to 1997Q2), which were not directly available from the MoFED.

[56] "Data on public expenditure are subject to revisions at various stages from budget to revised budget, provisional and final" (World Bank, 2004:1).

Table 3-5: Summary of Fiscal Variables (% GDP)

	1993-96	1996-99	1999-02	2002-05	2005-08
Total Revenue and Grants	13.8	17.3	19.0	20.8	17.3
Revenue	11.6	14.7	15.5	15.6	13.4
Grants	2.1	2.6	3.5	5.2	3.8
Total Expenditure	17.2	21.9	26.2	25.6	21.1
Current Expenditure	10.3	14.0	17.9	15.3	10.5
Capital Expenditure	6.8	7.9	7.5	9.8	10.6
Sinking Fund	0.0	0.1	0.0	0.0	0.0
Special Programmes	0.0	0.0	0.7	0.5	0.0
Financing	3.4	4.6	7.1	4.8	3.8
External	3.0	1.9	4.1	3.5	1.1
Domestic	0.7	1.0	2.8	2.7	2.8
Privatisation Proceeds	0.0	0.9	0.6	0.0	0.1
Other and Residual	-1.2	0.8	-0.3	-1.5	-0.3
GDP growth (nominal)	12.8	2.4	3.6	19.1	30.1
GDP growth (real)	7.6	1.8	5.0	6.3	10.5

Source: Author's calculations.

The following tables present detailed trends of government receipts and public expenditure. In terms of the composition of government receipts, the table below indicates that grants have increased their share from 15 percent to over 20 percent of total receipts. A strong move from untied cash to earmarked grants is noticeable from the table. Tax receipts remain the main source of domestic revenue, with a share of about 75 percent (except in the period 1996-99). In terms of the tax composition, direct taxes have been declining since 1999-02, while indirect taxes have reinforced their position as the main source of tax revenue. Within indirect taxes, import duties and taxes have had a strong performance (rising from about 57 percent to 68 percent of total indirect taxes), while export taxes have recently been abolished. Domestic indirect taxes lost some importance in the past few years, as well as the non-tax component of revenue. It can be seen that import duties and taxes have been the main source of domestic revenue in Ethiopia (37 percent in the last period). The persistence of this trend may look surprising given the economic reforms (namely trade liberalisation) introduced by the current government. Nevertheless, the reduction of tariff rates on trade was more than compensated by a sharp increase in trade volumes (in particular imports).

Table 3-6: Compositions of Government Receipts (% of level above)

	1993-96	1996-99	1999-02	2002-05	2005-08
Revenue	84.9	86.7	83.9	77.5	78.6
Tax revenue (inc measures)	73.1	66.7	72.9	77.9	78.3
Direct taxes	33.3	35.6	37.3	33.2	30.1
Indirect taxes	66.7	64.4	62.7	66.8	69.9
Domestic indirect taxes	38.8	35.6	31.9	31.3	31.7
Import duties & taxes	56.5	59.8	66.0	68.7	68.3
Export taxes	4.7	4.6	2.1	0.0	0.0
Non-tax revenue	26.9	33.3	27.1	22.1	21.7
Grants	15.1	13.3	16.1	22.5	21.4
Grants in kind/earmarked	11.9	19.6	62.8	65.6	63.3
Untied cash & CPF/DBS grants	88.1	80.4	37.2	34.4	36.7

Source: Author's calculations

The table below presents the composition of government expenditure. Capital expenditures as a percentage of total government spending show a U-shaped trend during the period under analysis. There has been a clear effort to finance capital purchases from domestic sources (central treasury), while the relative importance of external loans has been declining steadily. With regard to current expenditure, defence spending was particularly high during the Eritrean war (1998-00) and the main factor responsible for its increasing share during that period. The shares for poverty-targeted expenditure and capital expenditure have suffered from the war. Domestic interest payments have increased recently, probably due to the recent reliance on domestic borrowing.

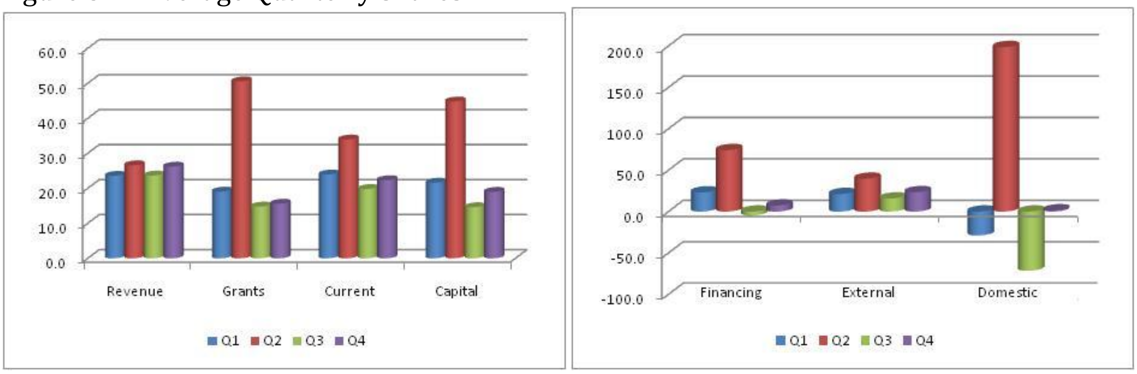
Table 3-7: Composition of Government Expenditure (% of level above)

	1993-96	1996-99	1999-02	2002-05	2005-08
Current expenditure	62.8	66.1	70.4	59.5	50.2
Defence	14.8	25.4	36.3	21.1	17.7
Poverty-targeted expenditure	33.3	30.1	24.0	33.3	44.7
Education	53.2	53.9	56.8	61.5	61.7
Health	18.4	18.2	16.9	13.8	13.2
Agriculture & Nat. Resources	19.7	22.4	22.6	21.4	22.7
Roads (Urban Dev. & Const.)	8.7	5.5	3.8	3.2	2.4
Interest payments	20.7	13.9	9.4	8.5	6.3
Domestic interest	75.9	62.5	57.1	53.5	66.7
External interest	24.1	37.5	42.9	46.5	33.3
External assistance	n/a	n/a	8.9	9.1	2.6
Others	31.2	28.6	21.3	28.0	28.6
Capital expenditure	37.2	33.8	26.9	38.9	49.7
Central Treasury	n/a	n/a	48.3	63.5	74.9
External assistance	n/a	n/a	12.0	13.9	17.1
External loans	n/a	n/a	39.7	22.6	8.0
Sinking fund	0.0	0.1	0.0	0.0	0.0
Special programs	0.0	0.0	2.7	1.7	0.1

Source: Author's calculations

The previous tables presented a summary of various fiscal variables aggregated over three-year periods. However, it is crucial to look at the quarterly fiscal performance to understand the intrinsic intra-year dynamics. The average quarterly shares for the main fiscal variables (below) reveal a significantly skewed pattern. Seasonality can arise for a number of reasons: weather conditions, timing of decisions, donor financial cycles, calendar events (e.g. Christmas), etc. The next few paragraphs will provide an interpretation of the nature and behaviour of seasonality. In our sample, 34 percent of current expenditure and 45 percent of the capital expenditure is carried out in the second quarter of the year (i.e. last quarter of the Ethiopian fiscal year). Perhaps not surprisingly, current expenditure is relatively smoother throughout the year since public wages (an important share of recurrent costs) are usually paid on a regular basis. Nevertheless, the pattern of capital expenditure is particularly concerning, since it may promote inefficient spending and impact on planning and project quality. It can be argued that Ethiopia would benefit from a smoother expenditure pattern.

Figure 3-1: Average Quarterly Shares



Since 1995 Ethiopia has had a federal system with some degree of budget independence, where regions are not allowed to borrow but can collect taxes. The regions receive a regional block (un-earmarked) grant from the federal government. Hence, it may be the case that the regional governments often rush to spend the budgeted funds in the last quarter of the fiscal year to avoid returning the unspent block grant to the federal treasury. Since our data is the consolidated general budget, it is not possible to assess how significant this behaviour is. External factors such as weather conditions (spending 'rush' just before the rainy season) and donor disbursements may also be partly responsible for this pattern. The data suggests that government revenues and its subcomponents are evenly spread throughout the year. However, external grants show a skewed distribution, where 51 percent of the yearly disbursements arrive in the last quarter of the Ethiopian fiscal year. This pattern may have implications for the effectiveness of aid flows, which cannot be captured by annual data. With regard to deficit financing, there is a marked difference between the sources. While external loans follow a similar pattern to expenditures, domestic borrowing mainly takes place in the second quarter, and is partly repaid in the first and third quarters. This volatile behaviour (government massively borrowing in second quarter and repaying large chunks of domestic debt in others), may be induced by foreign aid uncertainty, since domestic revenues seem relatively stable.

Transformed Data

The variables that will be included in the VAR system are quarterly observations for: (i) development expenditures (Development), which include capital expenditures and poverty-targeted recurrent spending (e.g. health and education); (ii) current expenditure minus poverty-targeted recurrent spending and food aid (Current); (iii) domestic revenue (Revenue), which includes tax and non-tax revenue, as well as privatisation proceeds; (iv) foreign grants minus food aid (Grant); (v) foreign loans (Loan); and (vi) government borrowing (Borrow). For the purpose of this study, the original (nominal) variables are deflated by the non-food component of the Consumer Price Index (CPINF), since the GDP deflator is not available on a quarterly basis. The variables are expressed in billion birr.[57]

All fiscal variables included in the system are 'observed', whereas 'other items' is left out (usually calculated as a residual from the budget identity). This variable may include 'check float' (e.g. checks issued to contractors but not cleared by banks), expenditure committed unpaid, 'revenue in transit', (regional) cash balances and any other statistical discrepancies. Other studies have usually dropped a fiscal item (e.g. non-tax revenue or loans), but this is likely to induce an omitted variable bias, arising from neglected dynamics between the main variables and the excluded item. In some cases, this may be strictly necessary because some data sources (e.g. IMF) tend to include any residuals under the domestic borrowing item (which in itself complicates the analysis).

Capital expenditures and poverty-targeted recurrent spending are lumped together since these two items are often interdependent. The construction of a hospital or a primary school needs to be met by increasing recurrent costs such as the wages of health professionals and teachers. Moreover, maintenance and repair of these and other socio-economic infrastructure (e.g. roads) normally fall under the recurrent budget. The correlation coefficient between these two variables is very high, which supports this decision. Since the aim of this study is to evaluate the fiscal impact of foreign aid, it is also natural that these items are pulled together. In fact, the classification between capital and recur-

[57] A similar approach can be found in Osei et al (2005). Taking logarithms of the variables is not appropriate, since it would violate the budget identity.

rent expenditures in some sectors is often blurred and recurrent items often appear in the capital budget anyway. This often follows from donor-implemented projects that, although having a recurrent component, are entirely classified as capital costs.

> "Data on capital expenditure have to be interpreted with caution. As in many countries, all donor-funded operations (and many government-financed initiatives organized in the form of projects) are shown under the capital budget, as a result of which it includes some expenditures which are essentially recurrent in nature (for example provision of textbooks, drugs, or agricultural program implementation). Also, as noted earlier, the capital budget by no means captures all project-related expenditure by donors. Furthermore, actual aid-funded expenditures are systematically under-reported, even for those projects that are included in the budget documents." (World Bank, 2004:11)

Food aid was extracted from both the expenditure side (recurrent budget) and the receipts side (grants) since it had the potential to bias the analysis. There have been a number of droughts in Ethiopia over the past 15 years, which considerably affect our variables due to the scale of these unanticipated events. Since the fiscal data is detailed enough to uniquely identify these flows, we chose to extract them from the data without compromising the analysis. Unfortunately, there is probably more noise in the data due to the war with Eritrea. However, it was not possible to extract the 'war levy' from the non-tax revenue component and feasibly identify what proportion of domestic borrowing was used to pay for these extra defence costs.

> "The fact that the surge in defense spending did not significantly squeeze out other types of expenditure is explained by the fact that it was largely financed by extraordinary, and unsustainable, means, in the form of domestic borrowing from bank and non-bank sources. (…) Hence there was not a peace dividend in terms of freeing up domestic resources over the past two years, but rather a drastic reduction in the internal borrowing by which the additional defense spending had been financed." (World Bank, 2004:3)

Therefore, Current will include defence related expenditures, interest payments, and 'other' recurrent spending.

The following plots show the fiscal variables in real terms[58] and a four-period lagged moving average to facilitate the analysis. The graphs suggest that there is substantial volatility in some of the variables. This might be a problem because the CVAR framework cannot take into account all types of seasonality.

---

[58] The nominal variables were deflated by the non-food CPI, since the GDP deflator is not available on a quarterly basis.

Figure 3-2: Variable Plots and 4-Period Moving Average

Obs.: From left to right we have Development, Current, Revenue, Grant, Loan and Borrow.

As expected, development expenditure shows a strong seasonal component. However, when the seasonality is smoothed we can see a clear positive trend, possibly with a break around 2000. In terms of current expenditure, we notice a sharp increase in defence spending in 1998 due to the war with Eritrea (May 1998 to June 2000). In 1998Q2, defence spending accounted for about half of total recurrent costs. After the war, there was a gradual scaling down of defence spending, although still relatively high in 2001 and 2002. In 2003Q2 there were high interest payments on external debt (probably on 2002 loans), while the fall in 2006Q4 is mainly due to lower than usual defence spending and 'other items'. Domestic revenue shows a strong positive slope, albeit with some variability in the past few years. While privatisation proceeds kept revenue levels high during the war, there might be a potential break around 2003.

The trend of external grants is more complicated to disentangle. These flows have been low during the war with Eritrea, and there has been a marked increase ever since.[59] The peaks of 2004Q1 and 2004Q4 are explained by large amounts of untied cash (mainly IDA and some bilateral budget support). From 2004Q4 until 2006Q1 we notice a sharp decline in the amount of grants, mainly due to donors withholding their commitments after the civil unrest that followed the 2005 elections. Finally, we look at the main two sources of deficit financing. Foreign loans have three main peaks: 2001Q3, 2002Q2, and 2002Q4, which can be mostly explained by large disbursements of IDA loans. Before 2003 most IDA support was provided in the form of concessional loans, while after 2003 there was a shift to aid grants. Domestic borrowing (mostly drawings from the banking system) has been particularly volatile, especially after 2002. It seems clear from the graphs that most of the extraordinary costs related to the Eritrean War were financed by domestic borrowing. Finally, we also note the negative correlation between aid grants and domestic borrowing, especially in 2004Q1 and 2004Q4. These quarters were associated with very high budget support disbursements, suggesting that (directly or indirectly) foreign aid was used to retire domestic debt.

In summary, the initial inspection of the variables suggests the following hypothesis: (i) aid grants seem positively correlated with development expenditures, when accounting for extraordinary events (Eritrean War and the 2005 election); (ii) domestic borrowing is positively correlated with current expenditure (mainly due to the Eritrean War) but its volatility in the end of the sample may be associated with donors' failing to disburse aid grants and (iii) foreign loans seem exogenously determined, since the trends in the remaining fiscal variables do not seem to explain its behaviour.

3.5 Estimation

3.5.1 Unit Root Tests and Seasonality

We could argue that individual unit root tests should not be performed since they are based on the univariate framework, which is not compatible with our assumption about the data generating process (DGP). This stems from the fact that the unit root is not an

---

[59] "This increase has been driven by two factors: the return of bilateral grant donors who had largely withdrawn support during the war with Eritrea, and substantial increases in budget support – particularly from the EU and the World Bank." (World Bank, 2004:28)

intrinsic characteristic of the data, instead it depends on the model that is chosen to represent the dynamic process (usually an AR(p) in the univariate case). Hence, one could proceed immediately to the estimation of the unrestricted VAR, and then test the stationarity of the individual series on the more natural multivariate context. Note that the trace test (for the determination of cointegration rank) is a multivariate version of the Dickey-Fuller test (Dennis, 2006:12), while every stationary variable will be directly associated with a cointegrating relation in the model (itself and perhaps a trend).

Nonetheless, we decide to undertake a preliminary assessment of the presence of unit roots in the data, with a special focus on seasonal unit roots. Hylleberg et al (1990:216) suggest that, "because many economic time series exhibit substantial seasonality, there is a definite possibility that there may be unit roots at other frequencies such as the seasonal." If this is the case, then first differencing may not eliminate all the roots and one may need to apply seasonal differences (and test for seasonal integration). This aspect has often been overlooked in empirical studies, although neglecting seasonal unit roots may give rise to spurious results, in the same way that 'regular' unit roots do. Some researchers try to circumvent these potential problems by 'de-seasonalising' the data. There are a number of dedicated programmes that try to purge the effects of seasonality.[60] Nonetheless, there is a growing perception that these procedures may distort the data rather than reliably remove the seasonal 'noise', as part of the trend and cycle components may also be eliminated. Hylleberg (2006) argues that seasonally adjusted data may not be the most effective use of information due to (i) the use of the wrong seasonal adjustment filter, and (ii) individual seasonal adjustment, which neglects the (multivariate) DGP. Moreover, the seasonal components may not be a 'noise', but rather a source of important information (regarding the underlying DGP) across series.

Consider the conventional (multiplicative)[61] decomposition of a macroeconomic variable into trend (T), cyclical (C), seasonal (S), and irregular (E) components:

$$X = T \times C \times S \times E$$

The main focus of this study is on the trend and cycle components. We could further subdivide the trend component into its deterministic and stochastic parts, and the cycle

[60] Some of the standard filters used in applied work to remove seasonal fluctuations include Census X11, X12-ARIMA, and Tramo/Seats.

[61] An additive formulation could be obtained by taking the logs of the macroeconomic variables. However, some of the variables in the sample take both positive and negative values (e.g. borrowing) and therefore could not be logged.

into long and short duration cycles. Depending on the time perspective of the study, the sub-cycle can be either non-stationary or stationary (Juselius, 2007:21). The stochastic time dependence of the variables will be the main focus of this study, although the deterministic part (linear time trend) will also be important as a measure of non-zero average linear growth rates usually present in economic data (Juselius, 2007:21). Moreover, since most variables plots suggest the presence of a significant seasonal component, a rigorous study should be concerned with modelling seasonality. It is thus important to understand the nature of seasonality, whether deterministic or stochastic, before choosing an appropriate modelling approach. For example, if most capital projects are consistently executed in the second quarter due to, say, the end of the fiscal year or the 'dry season', then seasonality is said to be deterministic and can be approximated by seasonal dummies. However, if after a certain point in time the seasonal pattern is altered (e.g. bulk of implementation moved to the fourth quarter), then we have time-dependent seasonality and need to use a different strategy to account for it.

Seasonality is often modelled by the following three types of models (Hylleberg et al, 1990): (i) purely deterministic seasonal process; (ii) stationary seasonal process; or (iii) integrated seasonal process. The CVAR approach falls in the first category, as it allows the inclusion of seasonal dummies to control for a constant (deterministic) seasonal component, but not a time-varying (stochastic) seasonal component. The presence and treatment of stochastic seasonality is less straightforward and presents substantial modelling challenges (Johansen, 2000).[62] Therefore, the following paragraphs will test for the presence of seasonal unit roots.

Since standard unit roots tests are not adequate to assess seasonal unit roots, we use the seasonal unit root test proposed by Hylleberg et al (1990:216) (HEGY). The authors develop a "testing procedure which will determine what class of seasonal processes is responsible for the seasonality in a univariate process". The HEGY test is especially appropriate to the current context since it was originally formulated for quarterly time series (it has been since extended to monthly). The test is based on the model:

$$\Delta_4 y_t = \pi_1 z_{1,t-1} + \pi_2 z_{2,t-1} + \pi_3 z_{3,t-1} + \pi_4 z_{3,t-2} + \sum_{j=1}^{p} \alpha_j^* \Delta_4 y_{t-j} + u_t$$

[62] "A phenomenon that is not directly covered by the above model [CVAR] is the seasonal variation of time series" (Johansen, 2000:372).

where $z_{1t} = (1 + L + L^2 + L^3)y_t$, $z_{2t} = -(1 - L + L^2 - L^3)y_t$, $z_{3t} = -(1 - L^2)y_t$ and L is the lag operator. The null hypotheses H_0: π_1=0, H_0: π_2=0 and H_0: π_3=π_4=0 correspond to tests for regular, semi-annual and annual unit roots, respectively. These hypotheses are tested by estimating the model above by OLS and using the relevant t-tests and F-tests. The critical values reported are from Franses and Hobijn (1997). Moreover, F-tests may also be used to test the joint null hypothesis H_0: π_2=π_3=π_4=0, or that all π's are jointly zero. It should be noted, however, that the asymptotic distributions of the test statistics under the respective null hypotheses depend on the deterministic terms in the model. This fact is taken into consideration since there is evidence that at least some of the series seem to be trended. The number of lagged seasonal differences is chosen by using standard model selection criteria (or alternatively by testing the significance of the coefficients on the lagged seasonal differences) (Lütkepohl and Krätzig, 2004:67).

The null hypothesis of the HEGY test is that there is a unit root. We include a constant, a deterministic trend, and seasonal dummies in the (levels) test regression. The lag length of the seasonal differences was selected according to the Schwarz Criterion and the Hannan-Quinn Criterion. In case of conflict, both lag lengths are reported. The test statistics are reported in the table below. As expected, the results show that most variables have regular (zero frequency) unit roots (i.e. cannot reject π_1=0). The only exception is Borrow, for which the unit root is rejected at 5 percent. Moreover, a semi-annual unit root (π_2=0) is rejected for all variables except Development and Revenue. In the case of Revenue, there is only weak evidence since the statistic is not far from the 10 percent level, and the joint tests reject seasonal roots. With regard to Development, the joint test is rejected at lag zero. Unit root tests are known to have low power (especially if several lagged differences are included), which may explain the different conclusions for alternative lag lengths. Finally, the annual unit root (π_3=π_4=0) is rejected for all variables except for Development at lag 5 (but rejected for zero lags).

Table 3-8: Seasonal Unit Root Tests (Levels and Regular Differences)

Var.	Lags	H_0	Test	Stat	Var.	Lags	H_0	Test	Stat
DEV	5	$\pi_1=0$	$t_{\pi 1}$	0.32	ΔDEV	7	$\pi_1=0$	$t_{\pi 1}$	−0.87
		$\pi_2=0$	$t_{\pi 2}$	−1.29			$\pi_2=0$	$t_{\pi 2}$	−1.83
		$\pi_3=\pi_4=0$	F_{34}	5.07			$\pi_3=\pi_4=0$	F_{34}	3.44
	0	$\pi_1=0$	$t_{\pi 1}$	0.25		0	$\pi_1=0$	$t_{\pi 1}$	-4.60^{***}
		$\pi_2=0$	$t_{\pi 2}$	−0.27			$\pi_2=0$	$t_{\pi 2}$	0.03
		$\pi_3=\pi_4=0$	F_{34}	7.71^{**}			$\pi_3=\pi_4=0$	F_{34}	5.35
		$\pi_2=\pi_3=\pi_4=0$	F_{234}	5.92^{**}			$\pi_2=\pi_3=\pi_4=0$	F_{234}	3.58
CUR5	0	$\pi_1=0$	$t_{\pi 1}$	−1.67	ΔCUR5	0	$\pi_1=0$	$t_{\pi 1}$	-4.17^{***}
		$\pi_2=0$	$t_{\pi 2}$	-2.90^{**}			$\pi_2=0$	$t_{\pi 2}$	−2.39
		$\pi_3=\pi_4=0$	F_{34}	14.13^{***}			$\pi_3=\pi_4=0$	F_{34}	7.64^{**}
							$\pi_2=\pi_3=\pi_4=0$	F_{234}	6.80^{**}
REV2	3	$\pi_1=0$	$t_{\pi 1}$	−0.09	ΔREV2	2	$\pi_1=0$	$t_{\pi 1}$	-5.33^{***}
		$\pi_2=0$	$t_{\pi 2}$	−2.39			$\pi_2=0$	$t_{\pi 2}$	-2.54^{*}
		$\pi_3=\pi_4=0$	F_{34}	8.48^{**}			$\pi_3=\pi_4=0$	F_{34}	9.00^{***}
		$\pi_2=\pi_3=\pi_4=0$	F_{234}	9.44^{***}					
	0	$\pi_1=0$	$t_{\pi 1}$	−1.60					
		$\pi_2=0$	$t_{\pi 2}$	−2.24					
		$\pi_3=\pi_4=0$	F_{34}	9.88^{***}					
		$\pi_2=\pi_3=\pi_4=0$	F_{234}	8.18^{***}					
GR2	9	$\pi_1=0$	$t_{\pi 1}$	0.09	ΔGR2	7	$\pi_1=0$	$t_{\pi 1}$	−1.83
		$\pi_2=0$	$t_{\pi 2}$	-3.19^{**}			$\pi_2=0$	$t_{\pi 2}$	-2.88^{**}
		$\pi_3=\pi_4=0$	F_{34}	6.19^{*}			$\pi_3=\pi_4=0$	F_{34}	12.48^{***}
	0	$\pi_1=0$	$t_{\pi 1}$	−1.65		0	$\pi_1=0$	$t_{\pi 1}$	-4.94^{***}
		$\pi_2=0$	$t_{\pi 2}$	-5.10^{***}			$\pi_2=0$	$t_{\pi 2}$	-4.26^{***}
		$\pi_3=\pi_4=0$	F_{34}	11.55^{***}			$\pi_3=\pi_4=0$	F_{34}	10.12^{***}
LOA	0	$\pi_1=0$	$t_{\pi 1}$	−1.90	ΔLOA	0	$\pi_1=0$	$t_{\pi 1}$	-5.73^{***}
		$\pi_2=0$	$t_{\pi 2}$	-3.44^{***}			$\pi_2=0$	$t_{\pi 2}$	-3.70^{***}
		$\pi_3=\pi_4=0$	F_{34}	11.14^{***}			$\pi_3=\pi_4=0$	F_{34}	11.66^{***}
BOR	0	$\pi_1=0$	$t_{\pi 1}$	-3.59^{**}	ΔBOR	0	$\pi_1=0$	$t_{\pi 1}$	-7.55^{***}
		$\pi_2=0$	$t_{\pi 2}$	-3.02^{**}			$\pi_2=0$	$t_{\pi 2}$	-3.77^{***}
		$\pi_3=\pi_4=0$	F_{34}	9.12^{***}			$\pi_3=\pi_4=0$	F_{34}	8.22^{**}

Obs.: The Schwarz and the Hannan-Quinn Criterion were used (maximum set at 10 lags). The deterministic components included were: constant, trend (levels) and seasonal dummies. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

We also test whether the (first) differenced variables show any signs of unit roots. We excluded the trend from the deterministic components because it was insignificant when included. The results reported above broadly reject the null hypotheses of unit roots. The only exception is Development, where the hypotheses of seasonal unit roots are again not rejected, even for lag zero. This may suggest that its seasonal pattern may have changed in the sample period. However, we have to bear in mind that these tests are based on the univariate framework, while the presence of seasonal unit roots in the multivariate model seems to be a more pertinent question. Moreover, cointegration analysis with seasonal differences does not seem to be appropriate in this case, since this approach only removes the seasonal unit root but not the regular unit root. Regular differences eliminate most of the unit root symptoms, therefore warranting the use of the CVAR framework. It should also be noted that these problems (especially semi-annual roots) often do not disappear in seasonal differences. In fact, Hassler and Demetrescu (2005) suggest that seasonal differencing may introduce artificial persistence and therefore may create spurious unit roots.

Overall, the results from the HEGY tests do not provide strong evidence of seasonal unit roots. Hence, the seasonal components do not seem to be time-dependent, suggesting that the fiscal policy pattern within the year remained relatively stable throughout the sample (i.e. summer does not become winter). Nonetheless, the potential issue with the Development variable will be further investigated later on.

3.5.2 Specification and Estimation of the Unrestricted VAR

Economic theory does not provide much guidance on the lag order of the system. The empirical literature on the fiscal response to aid has often found (through lag determination tests) that a VAR with two lags is an appropriate representation of the fiscal dynamics. Since these studies have all used annual datasets, one could suggest that a VAR(4) or VAR(8) may be appropriate for a quarterly model. Nonetheless, since the frequency of the data and the time-frame of this study are different from previous empirical models, the notions of long-run equilibrium and short-run dynamics may also be distinct. In fact, it is difficult to conceive that a fiscal shock may still have a significant impact on the remaining fiscal variables 8 quarters later, since part of its impact is likely to be contemporaneous with relatively quick adjustment dynamics. For example, an increase in aid flows is likely to have an immediate impact on expenditure (especially if earmarked, since there will have an exact counterpart in the expenditure side), while the unspent funds will probably affect expenditure (or borrowing) in the next quarter or two. Moreover, since several CVAR studies with quarterly data start from a VAR(2), we will also start with this lag length and then test whether this is a correct representation of the DGP.[63] Moreover, there is a good argument to avoid overfitting the initial model (i.e. including many lags): each lag that is added to the specification will correspond to $p \times p$ additional parameters to be estimated, seriously compromising the number of degrees of freedom available. This may affect the efficiency of the estimates, as it is usually associated with larger variances (Kennedy, 2003:205).

[63] "[i]n practice it is seldom the case that a well-specified model needs more than two lags. Therefore, as a rule of thumb it seems useful to start with a VAR(2) model, search for structural shifts and, if necessary, respecify the model. When the model is well-specified one should test whether the lag length needs to be altered and, in case this is so, one should redo the specification checking in the new model." (Juselius, 2007:72)

In terms of the deterministic components, we have included an unrestricted constant and a deterministic trend restricted to appear in the cointegrating relations. This specification of the deterministic components corresponds to 'case 4' below.

$$\Delta x_t = \alpha[\beta', \beta_0, \beta_1]\begin{bmatrix} x_{t-1} \\ 1 \\ t \end{bmatrix} + \gamma_0 + \gamma_1 t + \varepsilon_t$$

Table 3-9: Specification of the Deterministic Components

Cases	Restrictions	Deterministic Components
1	$\beta_0 = \beta_1 = \gamma_0 = \gamma_1 = 0$	No deterministic terms in the model.
2	$\beta_1 = \gamma_0 = \gamma_1 = 0$	Constant restricted to the CI.
3	$\beta_1 = \gamma_1 = 0$	Constant unrestricted (no linear trends in VAR, but in variables).
4	$\gamma_1 = 0$	Trend restricted to the CI, but constant unrestricted in model.
5	No restrictions	Trend and constant unrestricted in model.

Obs.: 'CI' cointegrating relations
Source: Juselius (2007:99-100)

The advantage of this specification is that, not only does it allow for linear trends in the cointegration space, but also in the variables in levels. This choice seems appropriate since it follows from the observation that the variables (in levels) appear to be trending, and because we are not sure whether these deterministic trends cancel out in the cointegrating space. 'Case 4' is usually considered to be a good starting point for most empirical applications since it is less restrictive than most alternatives, while 'case 5' has the drawback of creating quadratic trends in the levels. Moreover, the hypothesis that some variables (or in fact cointegrating relations) are trend-stationary can be easily tested within this specification.

> "Given linear trends in the data, case 4 is generally the best specification to start with unless we have strong prior beliefs that the linear trends cancel in the cointegration relations (...). This is because case 4 allows for trends both in the stationary and non-stationary directions of the model and hence similarity in the [rank] test procedure. When the rank has been determined, it is always possible to test the hypothesis $\beta_1 = 0$ as a linear hypothesis on the cointegrating relations." (Juselius, 2007: 140)

The initial model specification is not likely to satisfy all the desirable properties (assumptions) under which the model was derived, and therefore will fail some specification tests. In most cases, these issues can be addressed by examining the source of misspecification and modifying the model accordingly. The use of 'intervention' dummy variables is a common way to deal with outlier observations, which are often caused by exceptional (political or institutional) events (i.e. model is not capable to reasonably explain the value of the endogenous variable). Other useful tools are 'shift' and 'level' dummy variables, which account for structural breaks in the variables or relationships. In some cases, a few variables may be responsible for most of the problematic observations, e.g. when they are not reasonably explained by the other variables in the system, thus severely failing the normality tests. In this case, and provided that economic theory and appropriate tests support it,[64] it may be convenient to treat these variables as weakly exogenous. Conditioning on weakly exogenous variables often improves the stability of the model.

The tools summarised above will often be sufficient to obtain a statistically well-behaved model. The diagnostic tests (which will be described below) for the unrestricted 6-dimensional VAR(2) suggest that the model should be re-specified in the following way. Firstly, lag length tests suggest that a VAR(1) is a more efficient representation of the DGP, with no signs of residual autocorrelation. Secondly, the Loan variable seems to be problematic as it generates several outliers. The equation was excluded from the system and the variable assumed to be weakly exogenous. There are a number of reasons to support this decision. Given that multivariate tests suggest that the variable is stationary (i.e. shocks to the system do not have a permanent impact on the variable), its inclusion as an endogenous variable is not crucial for the long-run analysis.[65] The variable is still kept in the cointegrating space. Moreover, graphical evidence suggests that the remaining variables in the system are not likely to explain its behaviour, also demonstrated by the relatively lower R-squared. Thirdly, Current also seems to cause some instability, especially because of the war period.[66] If we exclude defence expenditure from the variable, Current becomes trend-stationary and therefore its path is not likely to be influenced by the remaining variables in the system. Since we know that the causality runs from Current to Borrow (at least in the war years) we also decide to make this variable weakly exogenous. Fourthly, even when estimating a partial model (conditioned on Loan and Current) there are a few outlier observations that need to be corrected with intervention dummies to guarantee Gaussian residuals: 2004Q1, 2004Q4, 2005Q2, 2006Q1, and 2007Q2. The dummy in 2005Q2 accounts for the general election, while 2006Q1 corresponds to the period when donors decided to withhold budget support funds following a period of political and civil unrest. The other three are mainly a result of large disbursements of budget support. In order to reduce the number of coefficients to be estimated, we can try to group dummies with similar impacts. Looking at the estimated coefficients for each

[64] "Including unmodeled stochastic variables may be problematic for inference and analysis purposes unless the variables satisfy exogeneity requirements" (Lütkepohl and Krätzig, 2004:67).
[65] The weak-exogeneity test (for a zero row in alpha) is rejected because the variable is stationary. Stationarity implies that the variable adjusts to its own long-run means (unit vector in beta), which in turn explains why there is strong evidence of a unit vector in alpha (i.e. variable is purely adjusting to one cointegrating relation).
[66] There is some evidence that the variable is weakly exogenous.

equation, the 2004 and 2006 dummies are grouped together (the latter with an opposite sign).

Based on this evidence, the initial model is re-specified as a 4-dimensional VAR(1). This means that the unrestricted 6-equation system was reduced to four endogenous equations, as two variables were set as weakly exogenous. The estimation of the fiscal model was thus conditioned on Loan and Current. This is often known as a 'partial model' since not all variables included in the system are (endogenously) explained by the process, but taken as given. There are a number of advantages in doing so. Firstly, our initial model suffered from several misspecification problems, and a close look at the individual series suggested that some variables might be responsible for the instability of the system. Conditioning the estimation of the model on these 'problematic' variables reduces misspecification issues. Secondly, this strategy reduces the number of estimated parameters in the model, and therefore may improve the efficiency of the estimates. In this particular case, theory and common sense may suggest that some of the variables should, *a priori*, be considered weakly exogenous. Coupling this knowledge with testing procedures provides a powerful tool.[67] The estimates of the new model are presented below in the VECM form.

[67] There is a clear trade-off here. Setting a variable as weakly exogenous can potentially bias our results due to neglected short-run effects, while allowing it to be endogenous can create 'noise' that will be transmitted throughout the system. Since these two variables have been strongly affected by external events (e.g. war and IDA shift to aid grants) it seems reasonable to admit that we are not able to explain its dynamics. In fact, the full 6-dimensional VAR(1) provides similar conclusions but does not pass several consistency checks.

Table 3-10: CVAR Estimates

Sample:	1993:03 to 2008:02 (60 observations)
Effective Sample:	1993:04 to 2008:02 (59 observations)
Obs. - No. of variables:	43
System variables:	DEVELOPMENT REVENUE2 GRANT2 BORROW
Weakly Exogenous/Fixed Variables:	LOAN CURRENT5
Dummy-series:	DUMGRANT{0} DUM072P{0} DUM052P{0}
Constant/Trend:	Restricted Trend
No. of Centered Seasonals:	4
Lags in VAR:	1

The unrestricted estimates:

BETA(transposed)

	DEVELOPMENT	REVENUE2	GRANT2	BORROW	LOAN	CURRENT5	TREND
Beta(1)	0.860	0.082	-2.262	0.199	0.123	-0.282	-0.033
Beta(2)	0.516	-1.802	1.052	0.374	-0.275	0.488	0.052
Beta(3)	-1.259	1.434	0.995	1.059	1.166	-1.197	-0.009
Beta(4)	-0.923	-0.804	-0.239	-0.124	-1.537	-0.992	0.215

ALPHA

	Alpha(1)	Alpha(2)	Alpha(3)	Alpha(4)
DDEVEL	-0.088	-0.215	0.260	0.184
	(-1.620)	(-3.974)	(4.815)	(3.397)
DREVEN	0.108	0.444	-0.174	0.150
	(1.976)	(8.102)	(-3.179)	(2.730)
DGRANT	0.347	-0.110	-0.026	0.099
	(10.432)	(-3.310)	(-0.780)	(2.970)
DBORRO	-0.886	-0.688	-0.541	-0.068
	(-9.435)	(-7.321)	(-5.763)	(-0.728)

PI

		REVENUE2	GRANT2	BORROW	LOAN	CURRENT5	TREND
DDEVEL	-0.684	0.606	0.187	0.155	0.070	-0.574	0.029
	(-6.812)	(4.590)	(1.285)	(2.503)	(0.658)	(-6.420)	(2.399)
DREVEN	0.404	-1.162	0.013	-0.016	-0.542	0.246	0.053
	(3.967)	(-8.683)	(0.088)	(-0.247)	(-5.065)	(2.718)	(4.325)
DGRANT	0.183	0.110	-0.950	-0.012	-0.109	-0.219	0.004
	(2.969)	(1.356)	(-10.594)	(-0.311)	(-1.680)	(-3.974)	(0.579)
DBORRO	-0.373	0.445	0.759	-0.998	-0.445	0.630	-0.016
	(-2.138)	(1.943)	(2.996)	(-9.262)	(-2.429)	(4.055)	(-0.772)

Log-Likelihood = 228.051

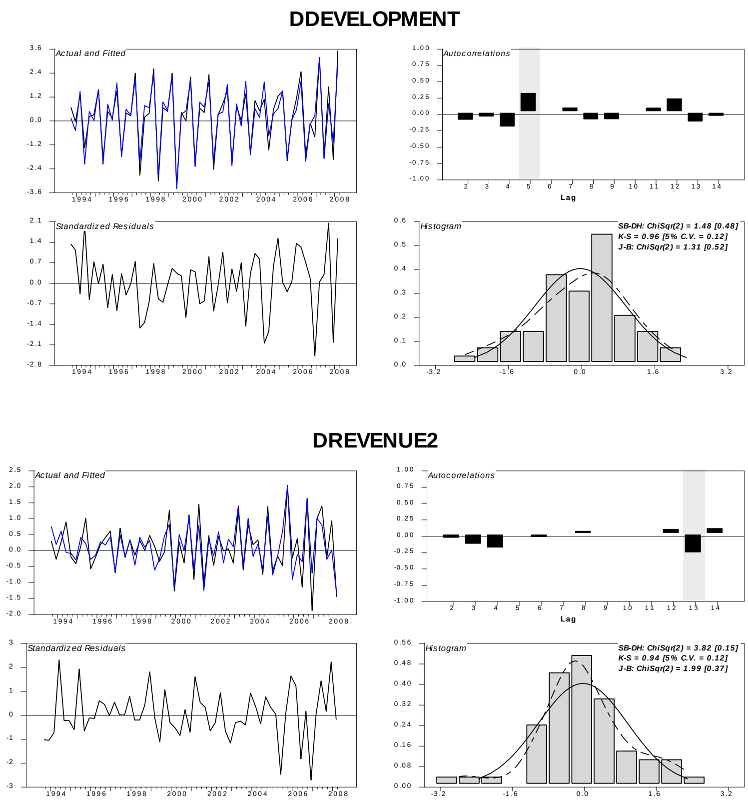
With the information provided above, it is possible to examine whether the cointegrating relations are equilibrium correcting or overshooting in specific equations. For example, if a statistically significant α_{ij} (for i = 1, ..., p and j = 1, ..., r) and the corresponding β_{ij} have opposite signs, then the j^{th} cointegrating relation is equilibrium correcting in equation $\Delta x_{i,t}$. However, if they have the same sign, it suggests that the j^{th} cointegrating relation describes overshooting behaviour in the i^{th} equation (Juselius, 2007:122). Nevertheless, the β vectors are not yet identified and the t statistics reported are only indicative, since our model may have non-stationary variables (invalidating inference). At this stage, it is often not possible to provide a meaningful interpretation of the model estimates.

3.5.3 Misspecification Tests

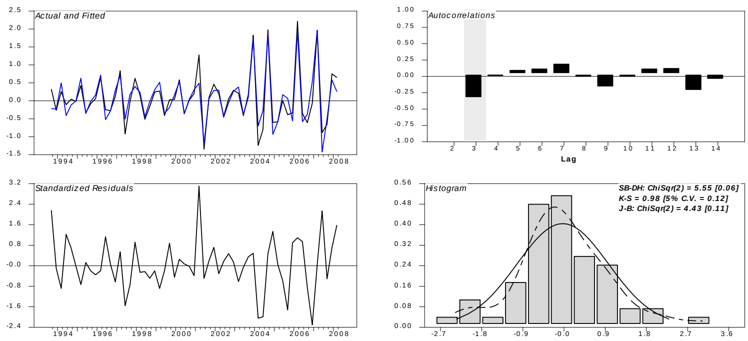
> "Simulation studies have shown that valid statistical inference is sensitive to violation of some of the assumptions, such as parameter non-constancy, autocorrelated residuals (the higher, the worse) and skewed residuals, while quite robust to others, such as excess kurtosis and residual heteroscedasticity." (Juselius, 2007:46)

In this section we report several formal misspecifications tests, most of which focus on the system residuals. These tests are important to assess the validity of the assumptions underlying statistical model (VAR). It is useful to start with a graphical inspection of the residuals, since it can help identify potential problems. The figure below shows, for each equation: (a) the plots of the fitted and actual values of the first-differenced 'endogenous' variable (top left panel); (b) the autocorrelogram of order 14 (top right panel); (c) the empirical and normal distributions (bottom right panel); and (d) the standardised residuals (bottom left panel). The graphs do not suggest any particular problems, especially after the outliers were corrected.

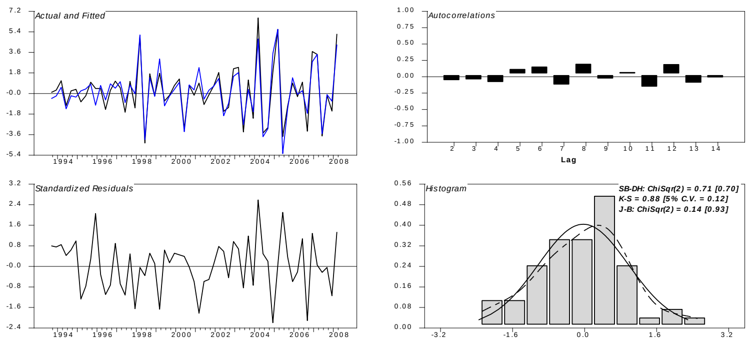
Figure 3-3: Fitted Values, Autocorrelograms, Distributions and Residuals



DGRANT2



DBORROW



Determination of the Lag Length

The appropriate lag length is determined by (sequential) likelihood ratio (LR) tests, where the null hypothesis of k lags (H_k) is tested against the alternative hypothesis of k+1 lags. In the specification below, T represents the size of the effective sample, which is kept constant,[68] and Ω is the residual covariance matrix.

$$-2lnQ\left(\frac{H_k}{H_{k+1}}\right) = T(ln|\Omega_k| - ln|\Omega_{k+1}|)$$

The test statistic is approximately distributed as χ^2 with p^2 degrees of freedom. However, the LR test alone is not going to be particularly informative, since an extra lag will almost always add information and improve the log-likelihood value.[69] Hence, we need to discount the log-likelihood by an appropriate (penalising) factor that represents the loss of degrees of freedom. The Schwarz (SC) and Hannan-Quinn (HQ) information criteria will

[68] The size of the effective sample needs to be the same when testing H_k against H_{k+1}, hence it is determined by the longest lag.
[69] The null hypothesis (H_k) suggests that the VAR model does not have significant coefficients at lag k+1.

serve this purpose, as they apply a penalising factor related to the number of estimated parameters (as a result of increasing the lag length).

$$SC = |\Omega| + (p^2 k)\left(\frac{lnT}{T}\right)$$

$$HQ = |\Omega| + (p^2 k)\left(\frac{2lnlnT}{T}\right)$$

Since we have several regressors in our specification and a relatively small sample, it will not be possible to test large lag-lengths. For example, a 6-dimensional VAR(8) requires the estimation of 288 autoregressive parameters, 6 constant terms, 6 trends, 18 seasonal dummies and 21 parameters in the residual covariance matrix, which adds up to 339 parameters. The estimation of such a model seems prohibitive for the current sample size of only 360 data points. Therefore we provide results up to lag 4.[70]

As it can be seen from the results presented below, the Hannan-Quinn criterion suggests a VAR(4) whereas the Schwarz criterion points to a VAR(1). The disagreement between the two criteria reflects the different ways in which they punish the extra lag. Finally, the LM tests provide a useful indication of 'left-over' residual autocorrelation in each VAR(k) model. In this case, the LM tests suggest that the VAR(1) does not suffer from (first order) autocorrelation. A significant coefficient at lag 4 may suggest the presence of time-varying (stochastic) seasonality not captured by the seasonal dummies. Overall, the results suggest that the VAR(1) provides a good description of the DGP.[71]

Table 3-11: Lag Length Determination

Model	K	T	Regr	Log-Lik	SC	HQ	LM(1)	LM(k)
VAR(4)	4	56	34	322.858	−1.755	−4.767	0.504	0.343
VAR(3)	3	56	28	283.937	−2.090	−4.570	0.573	0.028
VAR(2)	2	56	22	240.195	−2.253	−4.202	0.014	0.046
VAR(1)	1	56	16	219.944	−3.255	−4.672	0.232	0.232

Obs.: Effective Sample from 1994Q3 to 2008Q2. The LR lag reduction tests are not reported as they do not include a penalising factor.

Nonetheless, it is important to note that the tests mentioned above are only valid under the assumption of a correctly specified model. Therefore, there is often a need to go 'back and forth' with other misspecification tests until a satisfactory model specification is

[70] Lütkepohl and Krätzig (2004:110) suggest that an "excessively large value of p_{max} [maximum lags for test] may be problematic" since it affects the overall Type I error of the testing sequence.
[71] Due to the size of our sample, and provided that a VAR(1) does not show evidence of autocorrelated residual, a shorter lag-length seems advisable.

achieved. In fact, a long lag length suggested by test criteria can be interpreted as a sign of model misspecification. However, it is often not straightforward to assess whether residual autocorrelation is due to model misspecification (e.g. omitted variable) or due to neglected dynamics (i.e. model fitted with too few lags).

Moreover, other types of misspecification can also generate autocorrelated residuals (e.g. outliers). Although the assumption of uncorrelated residuals is one of the most crucial in the CVAR model, adding too many lags can also be harmful (overparameterisation).

> "[E]xperience suggests that adding too many lags is more harmful for the results than accepting some moderate residual autocorrelation in the model. This is because regime shifts, non-constant parameters, etc. are often difficult to diagnose in a heavily overparameterized model. Furthermore, experience suggests that residual autocorrelation in a first tentative VAR(2) model is more often associated with structural misspecification, rather than with left-out dynamics." (Juselius, 2007:72)

Residual Autocorrelation

The assumption of uncorrelated residuals is a crucial one in the VAR framework. One reason is that all χ^2 and *F*-tests are derived under the assumption of independent errors. If the model does not have this desired property, then the distribution of the tests may be significantly distorted. The test for residual autocorrelation is a Lagrange Multiplier (LM) test of n^{th}-order correlation with a small sample correction. The test is also asymptotically distributed as χ^2 with p^2 degrees of freedom. We perform the test until order 4 with the aim of detecting potential seasonal autocorrelation left-over in the model. If H_0 is rejected for LM(4), it may be evidence of seasonal unit roots. The centred seasonal dummies in the CVAR model control for deterministic seasonality (i.e. constant seasonal means). However, if there is strong evidence of stochastic seasonality (i.e. changing seasonal pattern), then this may need to be modelled explicitly. The results (reported below) do not suggest any significant left-over autocorrelation, even for orders 2 or 4. Therefore, this test eases concerns raised by the HEGY test regarding the presence of seasonal unit roots (either semi-annual or annual).

Table 3-12: Tests for Autocorrelation

Test	DoF	Statistic	P-Value
LM(1)	ChiSqr(16)	20.432	0.201
LM(2)	ChiSqr(16)	23.383	0.104
LM(3)	ChiSqr(16)	18.774	0.281
LM(4)	ChiSqr(16)	23.839	0.093

Heteroscedasticity

To evaluate whether the residuals have constant variance, we apply an m^{th}-order ARCH test to the residuals of each VAR equation. The test statistic is calculated as $(T + k - m) \times R^2$, where T is the total sample size, k is the VAR lag length, and R^2 is taken from an auxiliary regression. The test is approximately distributed as $\chi^2(m)$, and the H_0 assumes homoscedastic errors. The results for the multivariate LM tests indicate mild ARCH effects. This may not be serious, since Rahbek et al (2002) have demonstrated (through simulations) that "cointegration rank tests are robust against moderate residual ARCH effects" (Juselius, 2007:75). Moreover, the tests do not suggest the presence of significant first order (individual) ARCH effects.

Table 3-13: Tests for ARCH Effects

Test	Equation	DoF	Statistic	P-Value
LM(1)	System	ChiSqr(100)	140.709	0.005
LM(2)	System	ChiSqr(200)	240.116	0.028
ARCH(1)	DDEV		0.434	0.510
ARCH(1)	DREV		2.588	0.108
ARCH(1)	DGR		0.156	0.693
ARCH(1)	DBOR		1.804	0.179

Normality

In order to assess residual normality of the entire system, we report the Doornik-Hansen multivariate test (Doornik and Hansen, 2008). However, it is known that this test works best in large samples, since "in small samples skewness and kurtosis are neither asymptotically normal nor independent" and one needs to use transformations (Juselius, 2007:75). Even then, the test does not reject the hypothesis of multivariate normality. We can further investigate the normality of residuals by looking at univariate tests. Moreover, since "VAR estimates are more sensitive to deviations from normality due to skewness [third moment around the mean] than to excess kurtosis [fourth moment]" (Juselius, 2007:77), it is also useful to report this information. We expect the skewness values to be around 0, while kurtosis tends to be around 3. The results reported below do not seem to suggest serious violations of the normality assumption, and while the residuals for Grants may still seem problematic, this is likely to be due to excess kurtosis rather than significant skewness.

Table 3-14: Normality

Equation	Mean	Std. Dev.	Skewness	Kurtosis	Max	Min	Statistic	P-value
DDEV	0.000	0.416	−0.353	2.818	0.844	−1.057	1.475	0.478
DREV	0.000	0.421	−0.061	3.718	0.960	−1.173	3.817	0.148
DGR	0.000	0.256	0.262	4.028	0.788	−0.607	5.549	0.062
DBOR	0.000	0.721	−0.071	3.069	1.845	−1.633	0.712	0.701

Obs.: Multivariate test ChiSqr(8)= 10.289 [0.245]

Goodness of Fit

Finally, we report the ‘trace correlation’ for the system and the R-Squared (R^2) for each equation. The ‘trace correlation’ is a measure that can be interpreted as an average R^2 in the p VAR equations. The results for both measures are very high, which suggests that our model captures, to a large extent, the correlation among fiscal variables in Ethiopia.

Table 3-15: Goodness of Fit

	DDEV	DREV	DGR	DBOR	System
R^2	0.932	0.719	0.871	0.900	
Trace correlation					0.826

In summary, the battery of tests reported above indicate that there is no significant residual autocorrelation in the VAR(1) model, suggesting that this is an appropriate specification of the DGP. The univariate normality tests do not reject the null of good specification at the 1 percent level (except for Grant). In addition, there are no outliers with a standardised value higher than 3.34, which is acceptable for the size of this sample.[72] Moreover, there is no strong evidence of residual heteroscedasticity. The overall measure of goodness of fit (trace correlation) has a high value, whereas the R^2 of the individual equations is also very high (0.83). Hence, this VAR(1) model seems to be a good representation of the data.

3.5.4 Determination of the Cointegration Rank

> “The cointegration rank divides the data into r relations towards which the process is adjusting and $p - r$ relations which are pushing the process. The former are interpreted as equilibrium errors (deviations from steady-state) and the latter as common driving trends in the system.

[72] This value is computed as $X = \Phi^{-1}(1 - 0.025)^{1/T}$, where Φ is the cumulative normal distribution function.

Hence, the choice of *r* will influence all subsequent econometric analysis and may very well be crucial for whether we accept or reject our prior economic hypotheses." (Juselius, 2007:140)

Once a well-specified statistical model is achieved, we can then test for the presence of unit roots in the multivariate framework. We use the trace test to determine the cointegration rank.[73] It is a likelihood ratio test based on the *R*-form of the VAR model, which means that the short-run dynamics and (some) deterministic components are concentrated out.[74] The test evaluates the log likelihood function of the null hypothesis (H_0: rank = p), versus the value for the alternative where rank = r. The test does not give us the exact number of unit roots. Therefore, the determination of the cointegrating rank (r) relies on a 'top-to-bottom' sequential procedure, which is asymptotically more correct than the 'bottom-to-top' alternative (Juselius, 2007:133). The distribution of the Johansen test statistic is non-standard and has been determined by simulations (Johansen, 1996). It should also be noted that the asymptotic distribution of the test depends on the choice of deterministic components. Moreover, short-run effects were assumed to not matter asymptotically (hence, concentrated out), which may not be the case in small samples such as this one.

The trace test has been often criticised on the grounds that there is a bias to accept too many cointegrating relations (test is over-sized). A number of simulation studies suggest that there can be substantial size and power distortions, mainly because the asymptotic distributions are poor approximations of the true distributions in small samples (Juselius, 2007:140). Hence, we also report the small sample Bartlett correction (see Johansen, 2002), which ensures a correct test size. These corrections can be considerable for small samples like this one. Moreover, the sample size is not exclusively related to the number of observations, but also to the amount of information that the data contains. Short samples can be very informative about a hypothetical long-run relation and have good test properties, if the equilibrium error crosses the mean line several times over the sample period. However, the small sample correction does not necessarily solve the power problem, which can be "very low for relevant alternative hypotheses in the neighbourhood of the unit circle" (Juselius, 2007:141-2). Moreover, the critical values were simulated because the inclusion of innovational dummies and exogenous variables changes the asymptotic distribution of the trace test (Dennis, 2006:8). The results suggest the presence of three cointegrating relations, even when correcting for small sample bias.

---

73 It is also known as the Johansen test or the LR test for cointegrating rank.
74 Deterministic components in the cointegrating vector will influence the distribution of the test.

Table 3-16: Trace Test

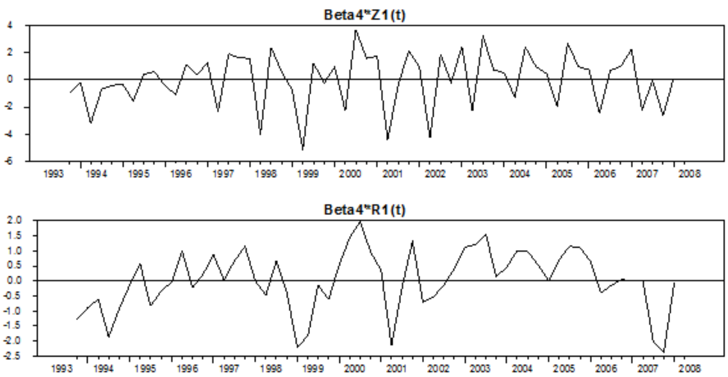
p–r	r	Eigenvalue	Trace	Trace*	Frac95	P-Value	P-Value*
4	0	0.825	256.104	247.230	82.501	0.000	0.000
3	1	0.709	153.342	149.598	57.316	0.000	0.000
2	2	0.672	80.591	79.418	35.956	0.000	0.000
1	3	0.221	14.743	14.669	18.155	0.143	0.147

Obs.: The magnitude of the eigenvalues is an indication of how strongly the linear relations are correlated with the stationary part of the process (Juselius, 2007:132).

The determination of the cointegrating rank is often a difficult choice that can have a significant impact on the analysis. Therefore, and bearing in mind the weaknesses of formal test procedures, Juselius (2007:142) suggests that it is often useful to complement the standard analysis with other available information. Some robustness checks include: (i) examination of the characteristic roots; (ii) significance of the adjustment coefficients; (iii) recursive graphs of the trace statistic; (iv) plots of the cointegrating relations; and (v) economic interpretability of the results.

Therefore, we also present graphs of the companion matrix roots and of the potential cointegrating relations. The graphs of the potential long-run relations also suggest the presence of three cointegrating vectors (see appendix). The bottom-panel (i.e. the concentrated model) of the last relation (below) suggests some persistence. This lack of mean reversion indicates that the last relation is non-stationary and we should accept only three cointegrating relations. Alternatively, we could assume full rank (i.e. all variables are stationary) and estimate a VAR in levels, but there is a greater danger in doing so as this leads to spurious results if variables are in fact non-stationary.

Figure 3-4: Plots of the Fourth Potential Relation



The modulus of the roots of the companion matrix seem relatively far away from the unit circle (0.7), but for such a small sample as this, it may still be possible to statistically accept a unit root. A unit root is often a convenient statistical approximation, which enables us to utilise a much richer framework that distinguishes between the longer and shorter term dynamic effects. It is therefore useful to consider unit roots for the empirical analysis of long- and medium-run macroeconomic relationships. Moreover, neglecting a unit root when there is some non-stationary behaviour may invalidate the empirical analysis.

Figure 3-5: Roots of the Companion Matrix

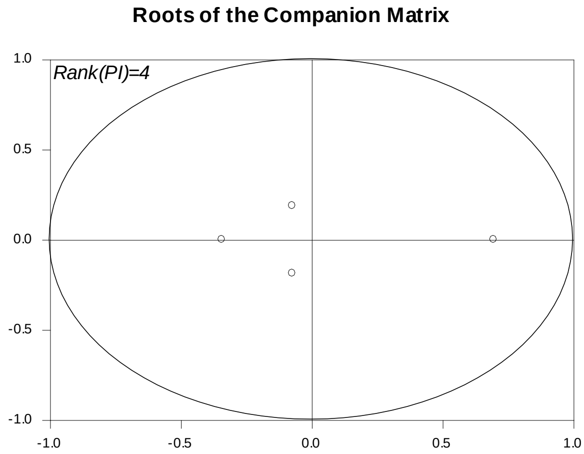


Table 3-17: Roots of the Companion Matrix

	Real	Imaginary	Modulus	Argument
Root1	0.696	0.000	0.696	0.000
Root2	−0.343	0.000	0.343	3.142
Root3	−0.073	0.188	0.201	1.943

Finally, the recursive graph of the trace statistic clearly suggests that there are three cointegrating vectors (see section for stability tests). Hence, these checks support the formal (trace) test in choosing r = 3. We then proceed to normalise the cointegrating relations on the largest coefficients.

The results from the reduced rank regression (RRR) are reported below. At this stage we are not able to provide much economic interpretation to the estimates, since the system is still under-identified. However, the last cointegrating vector resembles the government budget constraint, since sources of financing have an opposite sign to expenditure items. The potentially significant alpha coefficients correspond to a beta with the opposite sign, suggesting that the cointegration relation is equilibrium correcting in all equations, except in Grant. According to the size of the coefficients, the first cointegration may be a relation between Grant and Development, and the second between Revenue and Grant.

Table 3-18: CVAR Model with r = 3 (beta normalised on the highest eigenvalues)

THE MATRICES BASED ON 3 COINTEGRATING VECTORS:

BETA(transposed)

	DEVELOPMENT	REVENUE2	GRANT2	BORROW	LOAN	CURRENT5	TREND
Beta(1)	-0.380	-0.036	1.000	-0.088	-0.054	0.125	0.015
Beta(2)	-0.286	1.000	-0.584	-0.207	0.153	-0.271	-0.029
Beta(3)	-0.878	1.000	0.694	0.738	0.813	-0.834	-0.006

ALPHA

	Alpha(1)	Alpha(2)	Alpha(3)
DDEVEL	0.198	0.387	0.374
	(1.481)	(3.635)	(4.403)
DREVEN	-0.245	-0.800	-0.250
	(-1.862)	(-7.634)	(-2.996)
DGRANT	-0.785	0.198	-0.037
	(-9.730)	(3.088)	(-0.727)
DBORRO	2.004	1.239	-0.776
	(9.393)	(7.289)	(-5.737)

PI

		REVENUE2	GRANT2	BORROW	LOAN	CURRENT5	TREND
DDEVEL	-0.514	0.754	0.231	0.178	0.352	-0.392	-0.010
	(-5.400)	(5.530)	(1.455)	(2.639)	(4.943)	(-5.009)	(-2.859)
DREVEN	0.542	-1.042	0.049	0.003	-0.312	0.395	0.021
	(5.784)	(-7.768)	(0.312)	(0.045)	(-4.456)	(5.130)	(5.794)
DGRANT	0.274	0.190	-0.926	0.000	0.043	-0.121	-0.017
	(4.780)	(2.306)	(-9.674)	(0.009)	(0.996)	(-2.556)	(-7.665)
DBORRO	-0.436	0.390	0.742	-1.006	-0.550	0.562	-0.002
	(-2.869)	(1.796)	(2.930)	(-9.353)	(-4.846)	(4.503)	(-0.260)

Log-Likelihood = 220.679

We ought to be careful with the interpretation of the estimated coefficients, especially the cointegrating vectors. These long-run coefficients cannot be interpreted in the same way as those obtained from a single-equation regression model (e.g. elasticities, if variables were in logs), since the *ceteris paribus* assumption does not hold in the CVAR model. In the CVAR, an increase in (or shock to) one variable will impact all the remaining variables in the system. Johansen (2005) discusses this issue in greater depth (see appendix). Moreover, the cointegration rank does not necessarily correspond to the number of theoretical relationships derived from economic theory. This is because a theoretical equilibrium relation can be a linear combination of several irreducible cointegrating relations (Juselius, 2007:142). In our case, fiscal 'theory' only suggests one equilibrium relation (budget constraint), but that does not mean that the cointegration rank should be one.[75]

[75] "[C]ointegration between variables is a statistical property of the data that only exceptionally can be given a direct interpretation as an economic equilibrium relation" (Juselius, 2007:142). For example, we could have a budget constraint and a cointegrating relation between aid and public investment (r=2), where the latter is not derived from economic theory, but rather an empirical matter.

3.6 Structural Analysis

The main objective of this section is to provide a meaningful economic interpretation of the CVAR estimates. For that purpose, we need to address two different (but interrelated) identification problems: (i) the identification of the long-run relations, and (ii) the identification of the short-run structure. In practice, we need to impose restrictions on the cointegrating relations (β's), as well as imposing a dynamic adjustment structure on the (differenced) equations of the system. These restrictions are required to just-identify the system, which will then enable the researcher to make meaningful inference. In order to achieve just-identification we often need to refer to economic theory since some of the required restrictions may not be testable.[76] Once just-identification is achieved, we can then test any further restrictions (i.e. over-identifying restrictions) that we may wish to impose, for example, based on the (in)significance of the estimated parameters. Until we obtain a just-identified structure, *t*-ratios cannot be used to evaluate whether a parameter is needed in the relation or not. Finally, impulse response functions are often used to trace the dynamic impacts of an unexpected shock through the rest of the system. However, it is often the case that the CVAR residuals are correlated and therefore it is not possible to isolate the impact of a specific shock. The MA representation might help in this regard.

3.6.1 Testing Hypothesis

Johansen (1996) proposes test procedures to assess the validity of different types of restrictions on the cointegrating relations. Although these restrictions may not be identifying by themselves, they can be used to support the identification process. This is because they imply binding restrictions on $\Pi = \alpha\beta'$ (changing the value of the likelihood function) and are therefore testable (Juselius, 2007:173). Here, we will focus on three types of restrictions on β: (i) same restriction on all β: (ii) some β vectors assumed known; and (iii) only some coefficients restricted.

The first tests the same restriction in all cointegrating vectors. As an example, we can test whether a variable can be excluded from the long-run relations ('long-run exclusion'). These restrictions are not identifying, because they impose identical restrictions on all cointegrating relations. The second type of restriction allows us to assess whether a

[76] This scenario arises when the restrictions do not constrain the parameter space and therefore do not change the value of the maximised likelihood function.

hypothetical vector is stationary – all vector coefficients are assumed to be known. For example, we can test whether the 'balanced budget' hypothesis is stationary, i.e. a long-run relation. Moreover, we can also use this test as a multivariate unit root test to investigate whether a variable is stationary, trend-stationary, or a random walk. Finally, the third type of restriction enables us to test hypothesis on one specific vector without imposing any restrictions on the others. This is particularly useful if we want to focus our attention on a hypothetical cointegrating relation, while allowing some of its coefficients to be estimated. The table below presents the results of the different testing procedures. Exclusion tests suggest that all variables should be included in the cointegrating space. Moreover, stationarity tests indicate that none of the variables is trend-stationary.

Table 3-19: Tests of Stationarity and Long-Run Exclusion

Tests	c.v.	DEV	REV	GR	BOR	LOA	CUR	T
Stationarity	7.815	58.915	45.370	57.804	17.042	--	--	--
p-value	--	0.000	0.000	0.000	0.001	--	--	--
Exclusion	7.815	53.111	51.371	79.779	52.990	15.110	27.863	8.764
p-value	--	0.000	0.000	0.002	0.014	0.033	0.000	0.015

Obs.: The stationarity test includes the restricted trend.

The following table presents the results of further testing on β (long-run). First, we test the hypothesis of a 'budget constraint'. This is equivalent to testing whether the estimated expenditure coefficients are not statistically different from 1, while revenue and financing coefficients are not statistically different from –1. As expected, the hypothesis cannot be rejected, suggesting that 'net errors and omissions' (the variable excluded) is a white noise process. Any economic interpretation of this relationship will depend on the adjustment coefficients. For example, if only Borrow is adjusting to disequilibrium, then domestic borrowing can be seen as funding of last resort. Second, the hypothesis that the government tries to meet expenditures exclusively with receipts (revenue and grants) without resorting to deficit financing ('balanced budget') is clearly rejected.[77] This result suggests that, at least for some period, the government has relied on foreign loans and/or domestic borrowing to balance its fiscal accounts.

[77] In the case of a 'balanced budget', deficit financing would only be a short-run tool, since policy-makers aim to keep expenditures in line with revenues.

Table 3-20: Testing Hypothesis on β Vectors

Tests	Hypothesis: Vectors ~ I(0)	dof	Statistic	p-value
'Budget Constraint'	DEV + CUR – REV – GR – LOA – BOR	$\chi^2(6)^\dagger$	0.256	1.000
'Balanced Budget'	DEV + CUR – REV – GR	$\chi^2(6)^\dagger$	36.424	0.000
'Aid Spending'	DEV – REV – a*(GR + LOA)	$\chi^2(5)$	30.042	0.000
'Aid Additionality/Illusion'	DEV + CUR – a*GR – b*LOA	$\chi^2(4)$	47.660	0.000
'Development Funding'	DEV – a*GR + b*T	$\chi^2(4)$	0.805	0.938
'Categorical Fungibility'	CUR – a*GR – b*LOA + c*T	$\chi^2(3)$	40.821	0.000
'Revenue Displacement'	REV + a*GR + b*LOA + c*T	$\chi^2(3)$	16.002	0.001
'Borrowing Substitute'	BOR + a*GR + b*LOA + c*T	$\chi^2(3)$	9.056	0.029
'Financing Rule'	BOR + a*GR + b*REV + c*T	$\chi^2(3)$	0.368	0.947

Obs.: †Bartlett correction. The deterministic trend (T) was included to measure non-zero average linear growth rates. Tests do not include sign restrictions (only indicative). Including foreign loans and domestic revenue in 'development funding' also provides a stationary relation, as does including loans in the 'financing rule'. However, these variables are later found to be insignificant.

The 'aid spending' hypothesis is also rejected. Hence, it is not possible to establish a stable long-run relationship between the budget deficit (excluding grants) and foreign aid (or one of its components) over the period. This result may raise questions about the usefulness of this concept as a means to evaluate aid effectiveness, perhaps due to the presence of structural breaks. The next hypothesis tests the extent to which aid inflows are additional to government expenditure. The coefficients in this relation can offer relevant information (provided that some exogeneity conditions are met): (i) if higher than 1, it may suggest 'aid illusion', (ii) if equal to 1, it implies 'aid additionality', (iii) if lower than 1, then it is a sign of fungibility. Alternative combinations were tested but always rejected. Nonetheless, 'development funding', i.e. a long-run relation between Development and Grant, cannot be rejected. This seemed evident from the plots and is not particularly surprising given that some aid flows are earmarked. We can also construct an approximate test for 'categorical fungibility'. Since aid flows are often targeted to development spending, a relationship between aid flows and the recurrent budget may signal a divergence from donors' intentions. However, this hypothesis is rejected.

The 'revenue displacement' hypothesis tests whether a (negative) relationship can be found between aid flows and domestic revenue. It is strongly rejected, therefore implying that aid flows do not have a pervasive dampening effect on domestic revenue effort. Moreover, the hypothesis that aid flows and domestic borrowing are strong substitutes can also be evaluated. The test does not provide strong support for this hypothesis. Finally, we test whether there is a relationship among sources of revenue and financing, which may be interpreted as a 'financing rule'. This raises an interesting question about how the government reacts when one financing tool is lower than expected. The hypothesis (excluding Loan) is not rejected, but we ought to look at the adjustment coefficients to understand what this implies.

This initial investigation on potential long-run relations among fiscal variables provides interesting insights into the fiscal dynamics in Ethiopia. The tests support the existence of a budget constraint,[78] but not a balanced budget approach. Moreover, typical rules to assess aid effectiveness (aid spending and aid additionality/illusion) are put into question. Aid flows seem to be positively correlated with Development but not recurrent expenditure. The tests do not indicate a propensity of foreign aid to displace domestic revenue. They also suggest that we should take into account aid heterogeneity (since Grant and Loan do not have similar fiscal impacts). Finally the hypothesis that foreign aid flows and domestic borrowing are close substitutes is not strongly supported. This suggests that if these effects take place, they do not seem to be permanent and/or observed for the entire period.

In addition to these tests on the long-run coefficients, we can also carry out tests on α. These tests are closely related to interesting hypotheses about the 'pushing' and 'pulling' forces of the system – i.e. the common trends (or driving forces) and the equilibrium correction of the process, respectively (Juselius, 2007:193). In this context, we can test for: (i) long-run weak exogeneity; and (ii) known vector in α.

The first test assesses whether a variable impacts on the long-run (stochastic) path of the other variables in the system, whilst not being influenced by them (i.e. it is weakly exogenous with respect to the long-run information). This is implemented by testing a zero row in α. The second test relates to imposing the same restriction on each common trend.[79] An important example is the unit vector in α. This test investigates whether a variable is purely adjusting to one cointegrating relation, i.e. shocks to this variable will not have a lasting (permanent) effect on the remaining variables in the system. In this case, the effects will only be transitory. This is the case for grants, and perhaps domestic borrowing. It may seem surprising that aid grants do not have a long-run impact on the remaining fiscal variables, but it may suggest that aid is provided as a reward for the government's commitment to development, rather than pushing development expenditures to a new long-run path.

---

[78] This would not be the case if the excluded variable ('net errors and omissions') was non-stationary.

[79] "[T]he hypothesis of a known α vector is the equivalent of saying that one of the variables is exclusively adjusting to one cointegration relation, whereas the other variables are exclusively adjusting to the remaining r – 1 cointegration relations." (Juselius, 2007:201)

Table 3-21: Testing Hypothesis on α

Tests	c.v.	DEV	REV	GR	BOR
Exogeneity	7.815	23.147	38.758	53.702	66.890
p-value	--	0.000	0.000	0.000	0.000
Unit vector	3.841	3.841	22.845	19.080	2.266
p-value	--	0.000	0.000	0.132	0.084

3.6.2 Identification

This section is concerned with the (generic and empirical) identification of both the long-run and short-run structure of the model. We will impose the valid identifying restrictions found in the previous section to help the identification of the system. We can treat these as two separate statistical problems, which greatly simplifies the identification procedure. This is possible because the long-run parameters are the same in both the reduced and the structural form, and therefore, identification of β can take place in either form (Juselius, 2007:208). Therefore, we will first start with the identification of the long-run, and then proceed with the short-run structure. The latter is facilitated by keeping the identified long-run structure fixed (Juselius, 2007:230).

Long-Run Identification

In order to just-identify the long-run structure, we need to impose at least $r(r - 1)$ restrictions on β, i.e. $(r - 1)$ on each cointegrating relation. Since there are three cointegrating relations, we will require one normalisation and at least two restrictions per cointegrating vector for just-identifying the system. The table below presents the long-run identification scheme previously suggested. We note that the LR (joint) test for over-identifying restrictions is not rejected. All equations are equilibrium correcting and are not overshooting (i.e. significant α's and β's do not have opposite signs).

Table 3-22: (Over-)Identified β Vectors (Transposed)

	DEV	REV	GR	BOR	LOA	CUR	T
Beta(1)	1.000	−1.000	−1.000	−1.000	−1.000	1.000	0.000
	--	--	--	--	--	--	--
Beta(2)	−0.431	0.000	1.000	0.000	0.000	0.000	0.019
	(−13.988)	--	--	--	--	--	(5.555)
Beta(3)	0.000	1.000	−1.127	−0.541	0.000	0.000	−0.046
	--	--	(−9.068)	(−9.664)	--	--	(−9.195)

Obs.: *t*-statistics in brackets. The Log-Likelihood value is 219.903. LR test for over-identifying restrictions: $\chi^2(7) = 1.553$ [0.980]

Table 3-23: Adjustment Coefficients

	Alpha(1)	Alpha(2)	Alpha(3)
DDEV	−0.374	0.295	0.362
	(−4.928)	(1.984)	(4.177)
DREV	0.375	−0.343	−0.695
	(5.148)	(−2.396)	(−8.349)
DGR	−0.069	−0.835	0.122
	(−1.485)	(−9.088)	(2.286)
DBOR	0.515	2.264	0.905
	(4.289)	(9.593)	(6.585)

Obs.: The correct distribution of the *t*-statistic for the adjustment coefficients is somewhere between the Student's *t* and the Dickey-Fuller τ (Juselius, 2007:122). As a rule of thumb, *t*-statistic above 3 are usually considered significant.

As mentioned before, one cointegrating vector represents the budget constraint. Looking at the adjustment coefficients, we notice that they are all significant except for Grant. This suggests that donors are not responsive to budget disequilibria, instead, the government uses their fiscal policy tools to make ends meet. For example, if aid grants are suddenly stopped (causing disequilibria), then there will be excess expenditure over resources. The adjustment process implies that development spending needs to be reduced, revenues increased as well as domestic borrowing. Grants do not contribute to the adjustment.[80]

The second cointegrating vector suggests a positive relation between Grant and Development, which is not surprising. However, it seems that the causality is running from development spending to aid grants, since it is the Grant equation that is strongly adjusting to movements outside equilibrium. Hence, this cointegrating vector may be a proxy for a donor disbursement rule. For example, donors may wait for a clear government commitment to increase poverty-related expenditures before they contribute with funds to the budget. This could be seen as a (donor) 'aid conditionality' relation. Moreover, it may suggest that aid grants are disbursed in a pro-cyclical way. Domestic borrowing also seems to adjust.

The last cointegrating vector illustrates the alternative financing options of the government, given planned expenditures. Since, again, grants do not seem to be adjusting to the long-run disequilibrium, this may imply that revenues and/or domestic borrowing may have to compensate for aid shortfalls. However, revenues seem to follow the path of aid grants, and it is mainly domestic borrowing that compensates – this is similar to 'borrowing substitute'.

[80] A further example: if government expenditure [tax collection] is higher [lower] than planned, the budget imbalance is likely to be covered by domestic borrowing or higher [lower] revenues [expenditure] rather than foreign aid.

Short-Run Identification

With the long-run structure identified, we can now proceed to the short-run structure. However, the identification of the short-run tends to be a controversial issue. The estimation of the reduced-form model means that potential contemporaneous effects between the variables are captured in the residual covariance matrix – unlike single-equation models, these are not explicitly modelled.[81] Hence, there is a strong possibility that the residuals will be correlated, even though this study uses a quarterly dataset.[82] Since a straightforward identification of the short-run effects requires uncorrelated residuals, a careful analysis of the residual dependency structure will be crucial.

We start by analysing the results from the ECM representation. The table below shows the adjustment coefficients associated with each cointegrating vector (discussed above) and the deterministic components: constant, orthogonal seasonal dummies, innovation dummies, and weakly exogenous variables. The contemporaneous effects associated with the weakly exogenous variables are identified, since they were explicitly modelled. The results suggest that current spending has a positive (and proportional) impact on domestic borrowing, which is mostly due to defence spending, while foreign loans have a negative impact. Therefore, domestic borrowing seems to operate as a substitute of other resource options: foreign loans, aid grants and domestic revenues.

[81] The short-run is generically identified by the zero restrictions on the contemporaneous matrix A_0. The challenge is to recover the structural parameters.

[82] This problem is likely to be less significant in a quarterly dataset (compared to an annual dataset) due to richer dynamics and lower contemporaneous impacts.

Table 3-24: Error Correction Formulation

	DDEV	DREV	DGR	DBOR
CI_BC	−0.374	0.375	−0.069	0.515
	(−4.928)	(5.148)	(−1.485)	(4.289)
CI_DEV	0.295	−0.343	−0.835	2.264
	(1.984)	(−2.396)	(−9.088)	(9.593)
CI_REC	0.362	−0.695	0.122	0.905
	(4.177)	(−8.349)	(2.286)	(6.585)
Constant	−0.230	0.663	−0.124	−0.685
	(−2.375)	(7.136)	(−2.075)	(−4.470)
SEAS1	1.742	1.669	0.607	−1.078
	(6.955)	(6.938)	(3.927)	(−2.717)
SEAS2	1.289	0.943	0.657	−1.044
	(4.632)	(3.532)	(3.827)	(−2.368)
SEAS3	3.255	1.005	1.174	0.589
	(11.516)	(3.703)	(6.733)	(1.315)
dumGrant	−0.026	−0.003	1.415	−1.460
	(−0.094)	(−0.013)	(8.412)	(−3.378)
dum072p	1.729	−0.096	1.518	3.101
	(3.508)	(−0.203)	(4.991)	(3.971)
dum052p	−0.130	−0.755	−0.882	3.598
	(−0.269)	(−1.624)	(−2.956)	(4.694)
DLoan	0.194	−0.157	−0.002	−0.626
	(1.642)	(−1.386)	(−0.024)	(−3.343)
DCurrent	0.195	0.270	0.001	1.008
	(1.548)	(2.239)	(0.008)	(5.054)

The table below reports the error correlation matrix. The matrix suggests that there are significant current effects between grants and development expenditure, which is partly due to 'earmarked' grants. There is also a positive correlation between grants and revenue, which may arise from the fact aid flows provide foreign exchange to purchase imports, therefore generating higher trade revenues. This can be particularly true in the later part of the sample, when imports increased significantly. Domestic borrowing is negatively correlated with both revenue and grants, supporting the idea that these are substitutes. The coefficient for development is not significant. These results corroborate the findings from the long-run relations. However, it should also be noted that these contemporaneous effects may affect the interpretation of the adjustment coefficients. This is a tentative interpretation of the short-run impacts.

Table 3-25: Error Correlation Matrix

Correlation	DDEV	DREV	DGR	DBOR
DDEV	1.000			
DREV	0.288	1.000		
DGR	0.673	0.462	1.000	
DBOR	−0.153	−0.456	−0.295	1.000

Obs.: Significant correlations are those above $r_{ij} = 2\times T^{-1/2} = 0.258$.

There are several alternative identification schemes that could be applied for when the residuals are correlated. Juselius (2007:236-52) proposes the following short-run identification schemes: (i) impose restrictions on the short-run parameters when A_0 = I; (ii) impose (just-identifying) zero restrictions on the off-diagonal elements of Σ; (iii) impose general restrictions on A_0 without imposing restrictions on Σ; or (iv) re-specifying the full system model as a partial model based on weak exogeneity test results.

The first scheme allows a more parsimonious specification of the ECM by imposing (over-identifying) restrictions on the short-term coefficients. However, this approach does not solve the problem induced by correlated residuals. The second scheme achieves identification by imposing restrictions on the residual covariance matrix of the structural form (Σ). In practice, the innovations are orthogonalised through a Choleski decomposition, which means that a recursive structure is forced on the model. Since the results from this triangular system depend on the ordering of the variables, we need to have credible assumptions about the causal chain of events. This does not seem to be appropriate, since fiscal theory provides little guidance on this. The third scheme requires identifying restrictions on the matrix of contemporaneous effects (A_0). The objective is to account directly for any significant current effects in order to reduce or even eliminate (high) residual correlation coefficients. In practice, we include current effects in individual equations at the cost of zero restrictions elsewhere. However, we were not able to successfully introduce current effects in the model.[83] Finally, the fourth strategy is redundant for our case, since we already have a partial model but still have significant residual correlations.

3.6.3 Structural Moving Average Model

This section tries to provide further evidence on the dynamic effects of fiscal shocks. For this purpose we analyse the moving average (MA) representation of the CVAR and use impulse response functions to simulate the impact of shocks on our fiscal system of equations. In the table below, the 'alpha orthogonal (transposed)' allows us to identify the common trends in the model with the cumulated disturbances (Dennis, 2006:86-7). The 'loadings to the common trends' show us how the variables react to the common trends. The 'long-run impact matrix' shows how each variable is influenced by the cumulated disturbances. Also reported are the long-run covariance and the slopes of the common trends.

[83] Since we have a VAR(1), the individual ECM equations do not have lagged first-differenced variables to help identification. Hence, we were not able to avoid the violation of rank conditions when including current effects.

Table 3-26: MA Representation and Decomposition of the Trend

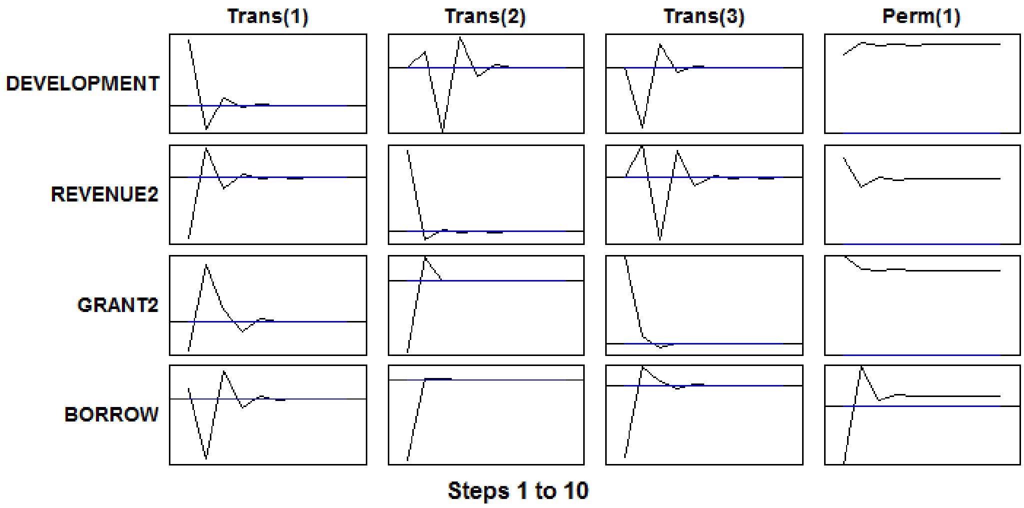
```
The Coefficients of the Common Trends:

RE-NORMALIZATION OF ALPHA Orthogonal:

ALPHA Orthogonal (transposed)
      DEVELOPMENT REVENUE2 GRANT2 BORROW
CT(1)      -0.709   -0.597 -0.351 -0.127

ALPHA Orthogonal (transposed)
      DEVELOPMENT REVENUE2 GRANT2  BORROW
CT(1)       1.000   0.842   0.495   0.180
            (.NA)  (3.493) (1.502) (1.642)

The Loadings to the Common Trends, BETA_ORT(tilde):
         CT1
DEVELO  0.604
       (7.366)
REVENU  0.311
       (7.366)
GRANT2  0.260
       (7.366)
BORROW  0.033
       (7.366)

The Long-Run Impact Matrix, C
       DEVELOPMENT REVENUE2 GRANT2  BORROW
DEVELO      0.604    0.508   0.299   0.108
           (7.366)  (5.130) (1.803) (1.924)
REVENU      0.311    0.262   0.154   0.056
           (7.366)  (5.130) (1.803) (1.924)
GRANT2      0.260    0.219   0.129   0.047
           (7.366)  (5.130) (1.803) (1.924)
BORROW      0.033    0.028   0.016   0.006
           (7.366)  (5.130) (1.803) (1.924)

The Linear Trends in the Levels, C*MJU
 DEVELOPMENT REVENUE2 GRANT2 BORROW
       0.081    0.065  0.016  0.000
```

Obs.: Stationary variables have zero rows while unit vectors have zero columns (C Matrix).

The analysis of the common trends (see long-run impact matrix *C*) suggest that shocks to grants and borrowing do not have substantial effects on the remaining variables of the system – they are purely adjusting (which suggests that these are transitory shocks). On the other hand, there seem to be two 'pushing' variables: development expenditures and domestic revenue. These are common driving trends that push the process away from its long-run equilibrium. There will be at least $p - r$ shocks with permanent effects (stochastic trends) and at most r transitory shocks (no long-run impact). Our model, with $p = 4$ and $r = 3$, will have one permanent shock and three transitory shocks. The permanent shock is already identified, while the structure of the temporary shocks requires some restrictions. Since there are no strong theoretical arguments to support specific restrictions, we will focus on the (identified) permanent shock.

The figure below presents the impulse response functions (based on the moving average representation of the CVAR). The permanent shock, which could be interpreted as a 'higher commitment to development', pushes development spending, domestic revenue, and aid grants to a higher level – i.e. the fiscal values converge to a new long-run equilibrium.[84] Domestic borrowing is substantially reduced in the first period, probably due to grants and revenue overshooting their long-run values, but bounce back before settling in its new (slightly higher) long-run path.[85] These results corroborate the co-movement of the three first variables. The transitional shocks are less straightforward to interpret, since they require further assumptions about the short-run structure.

Figure 3-6: Impulse Response Functions



The table below presents the corresponding values for the permanent shock.

Table 3-27: Structural MA Model

	Impact	DEV	REV	GR	BOR
C-tilde (normalised)	Final	1.000	0.515	0.413	0.055
Contemporaneous	Initial	0.398	0.303	0.223	−0.148
Contemporaneous (normalised)	Initial	1.000	0.761	0.560	−0.372

3.7 Stability Tests

The constancy of the estimated parameters is a crucial assumption for the validity of inference in the CVAR model. The stability tests presented here check whether there are any significant structural breaks that may undermine our conclusions. While this is the essence of most tests, they tend to assess parameter constancy from different angles. Most of these tests are explained in detail in Hansen and Johansen (1999). In practice, there will always be some degree of deviation from constancy, but the question is whether these

[84] The impulse responses are not upward sloping (instead converge to a constant mean) because the deterministic trend accounts for the slope. Moreover, these estimates represent period averages.
[85] The lack of significance intervals does not allow us to infer whether this is significantly different from zero.

deviations are serious enough to undermine inference and therefore require a modification of the model, usually with the inclusion of a shift or level dummy variable to account for changing means (Juselius, 2007:150).

Before starting stability testing, it is important to recall the two versions of the CVAR model: the full model (*X*-form) and the concentrated version (*R*-form). The graphs for the two models may suggest contradictory results, for example, when there is short-run instability, in which case the *R*-form graph should look more stable than the full model. This happens because in the full model all parameters are re-estimated in each step, while in the later only the long-run parameters are re-estimated (Dennis, 2006:95). Moreover, the tests can be calculated as forward recursive (base sample in the beginning of the sample) and backward recursive (base sample is the end of the sample).

Juselius (2007:150) groups recursive tests into four main categories: (i) recursive tests of the full model – e.g. test of the likelihood function; (ii) recursive tests based on the (transformed) eigenvalues – e.g. recursively calculated trace tests, the eigenvalues, the log-transformed eigenvalues, and the fluctuation test; (iii) recursive tests of the constancy of the cointegration space – e.g. 'max test of a constant beta' and the test of a 'known beta'; and (iv) recursive tests of predictive failure.

Recursive tests can be executed in two different ways. One can estimate the model for a reasonably sized sub-sample (base sample) from $t, ..., t_n$, and then add one further observation at a time (t_{n+1}) and re-estimate the model at every step until the end of the sample. The tests then assess whether the new estimates (using the additional information) are significantly different from the original coefficients from the base sample model. This is the usual procedure of a 'forward' recursive test. However, this test is not able to capture non-constancy (parameter instability) within the selected base sample, and therefore, it will also be useful to perform 'backward' recursive tests. These tests use a base sample at the end of the sample ($t_n, ..., T$) and add one observation at each step until the beginning of the sample.

We mainly report forward recursive tests, since the backward recursive tests did not reveal any problems. While it is often believed that the initial part of a sample is the main source of instability (e.g. due to stronger market regulations and poorer data quality), in the case of Ethiopia this seems not to be the case. The volatility in the latter part of the sample (clear from a graphical inspection of the variables) suggests that this is the most problematic part. Therefore, the more stable backward recursive tests are not reported in the Appendix.[86]

Since the current sample is not particularly long, it seems reasonable to estimate the baseline model with two-thirds of the sample and then look for signs of parameter instability in the other third. In this case, the base sample is 1993Q4-2003Q3, providing 40 observations for the estimation. Due to the size of our model, it is not desirable to estimate the baseline model with fewer observations, since it could potentially lead to inaccurate parameter estimates and compromise the recursive analysis – the graphs would show large variability in the beginning of the recursive sample. Nevertheless, we also want to keep the recursive sample large enough so that we can identify potential structural breaks.

Eigenvalues

The eigenvalue plots show the time paths of the *r* largest eigenvalues of the unrestricted VAR model and their respective 95 percent confidence intervals. The potential non-constancy of the long-run coefficients (alpha and beta) would be reflected in their respective eigenvalues (Dennis, 2006:96). If the base sample is relatively short, it is not unusual to find some fluctuation in the beginning [end] of the graph, if it is a forward [backward] recursion. If the long-run coefficients are stable, the eigenvalues should not show time dependency, i.e. there should be a straight line. Moreover, since the distribution of an eigenvalue close to 0 tends to be asymmetrical, we may also plot the 'transformed' eigenvalues, which are given by:

$$\xi_i = \log(\lambda_i) - \log(1 - \lambda_i) \qquad i = 1, \ldots, r$$

Both sets of results seem to suggest that the long-run coefficients are time-invariant, except perhaps the third. The volatility in the beginning of the graph is probably due to the short sample.

[86] These tests suggest that the data compiled from secondary sources (1993Q3-1998Q2) is a valid extension of the original sample.

Figure 3-7: Eigenvalues

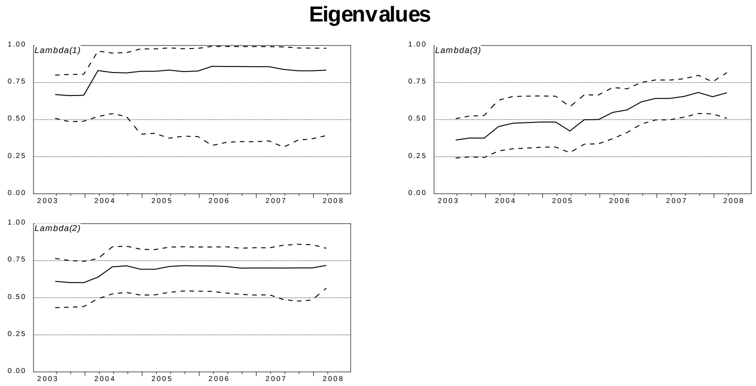
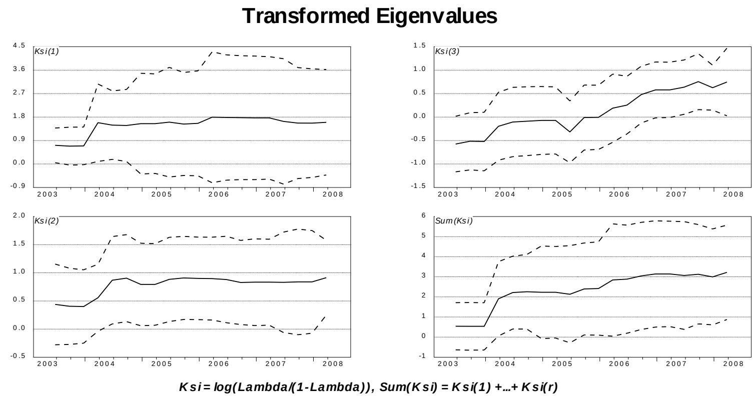
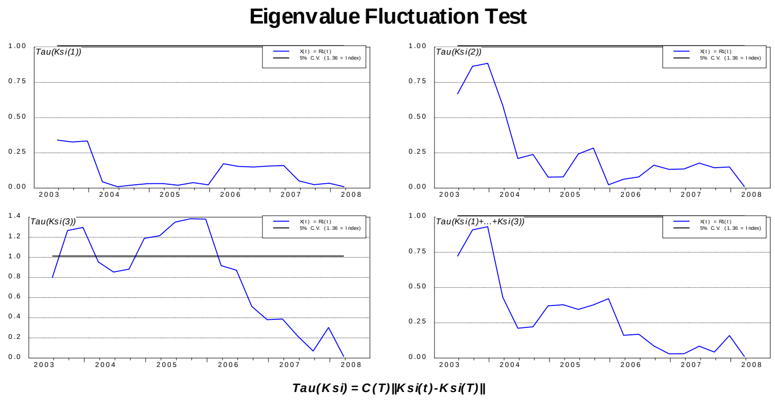


Figure 3-8: Transformed Eigenvalues



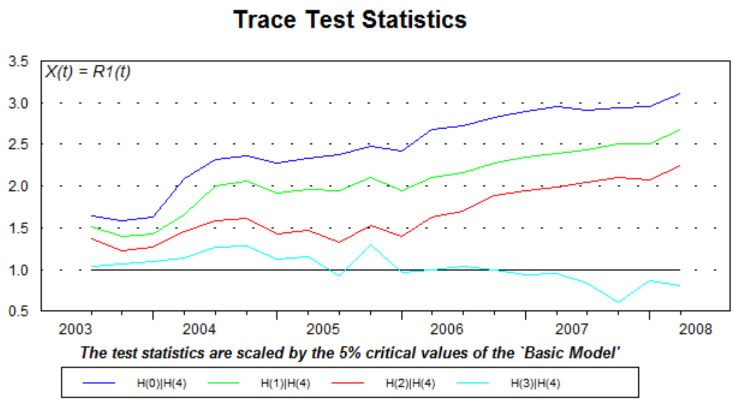
The following graph is the 'fluctuation test of the (transformed) eigenvalues'. The test is a *supremum* test and often regarded as rather conservative (Dennis, 2006:97-8). This means that if the test is rejected, there is strong evidence of non-constancy of the eigenvalues. The test statistic is scaled by the critical value. This means that if the value of the test statistic is above unity, then it lies outside the 95 percent confidence interval of the predicted value. Therefore, the statistic should stay low and not cross the normalised critical value line. The results confirm that the third eigenvector may be problematic, although further tests are required to understand the source of instability. We will now proceed with the analysis, and focus on the stability of both alpha and beta later in this section.

Figure 3-9: Eigenvalue Fluctuation Test



Finally, and as mentioned before, we present the recursively calculated 'trace test statistics'. These are also scaled by the critical value of the trace test distribution derived for a model without exogenous variables, shifts or dummies – 'basic model' (Dennis, 2006:100). The main point of interest in this graph is to observe the time path of the trace statistics. In our case (model with dummies and exogenous variables), the number of trace test statistics above unity may not correspond to the number of cointegrating relations at the 5 percent significance level (critical value should be higher). Nevertheless, we expect the number of upward sloping trace statistics to equal *r*. The results suggest that our choice of rank was correct, since there are three trace tests clearly upward sloping. Therefore, we have three long-run relations among fiscal variables.

Figure 3-10: Recursive Trace Test Statistic



Constancy of Beta

The 'max test of beta constancy' is usually seen as a conservative test, i.e. if it rejects, there are likely to be large deviations from the null (Dennis, 2006:99). Moreover, the test of 'β_t

equals a known β' assesses whether a fixed value of β (usually the full sample estimate or a constant regime period) is contained in the space spanned by $\beta^{(n)}$. The plots do not show any sign of parameter instability. Again, the mild rejection in the beginning of the second graph is likely to be due to the small sample.

Figure 3-11: Test Beta Constancy

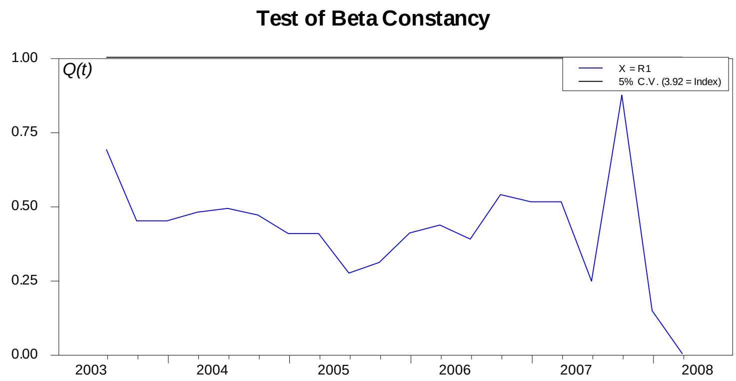
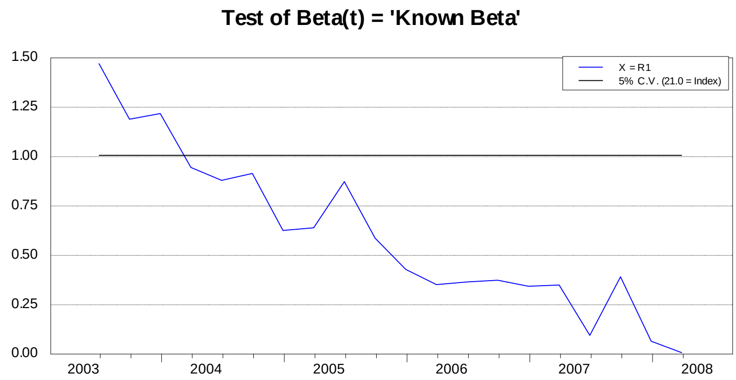


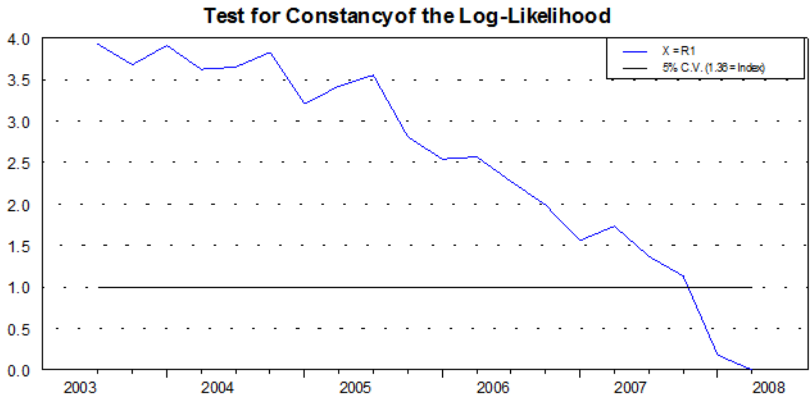
Figure 3-12: Test of Beta(t) = 'Known Beta'



Constancy of the (Maximised) Log-Likelihood Function

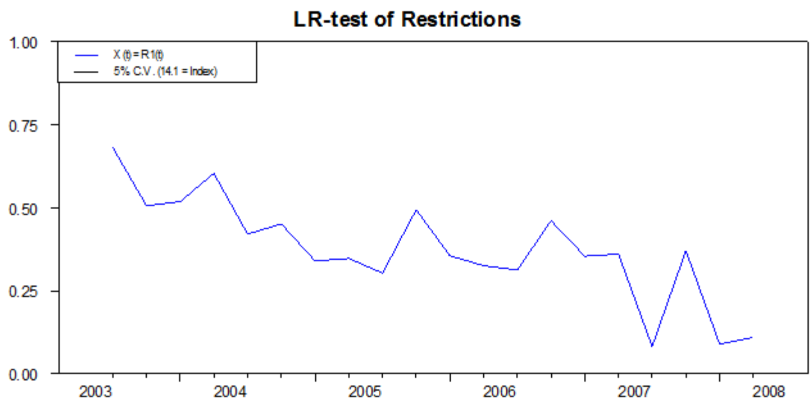
This is "essentially a test on the variances that measures the distance of the subsample and full sample estimates of the covariance matrix Ω" (Dennis, 2006:101). Again, the test statistic is scaled by the critical value. The result seems to suggest that there may be problems in our model. While this is likely to be due to the volatility later in the sample, the origin of this instability can be investigated through constancy tests on the alphas and betas.

Figure 3-13: Test for Constancy of the Log-Likelihood



The LR test for over-identifying restrictions is presented below. The fact that it is not rejected suggests that the (log-run) identification scheme proposed is valid.

Figure 3-14: LR-test of Restrictions



Finally, the graphs in the appendix test the constancy of the adjustment coefficients (alphas) and cointegrating vectors (betas) in the identified model. There is some instability in the estimated parameters, but it does not seem to be crucial as to question the conclusions of the model (e.g. parameters do not change signs). Overall, the stability tests do not seem to provide strong evidence of model misspecification, since there were no strong rejections.

3.8 Conclusion

The objective of this paper was to assess the fiscal effects of foreign aid flows in Ethiopia. For that purpose, a cointegrating VAR (CVAR) model was estimated with a recently collected quarterly fiscal dataset. A number of interesting hypotheses were tested, namely, whether aid flows induce an equivalent increase in public spending (additionality), a reduction in the 'tax effort', and/or a fall in domestic borrowing. The econometric results suggest the presence of three (empirical) long-run relations: (i) the government budget constraint; (ii) a possible donor disbursement rule; and (iii) a financing (trade-off) rule. While domestic fiscal variables adjust to budget imbalances, foreign aid grants seem to adjust to the level of development spending, which can be seen as an indication of (procyclical) aid conditionality. Given that 'earmarked' grants have a contemporaneous impact (there is a budget counterpart under capital expenditure), this finding implies that budget support grants are provided after the government makes a strong commitment to poverty reduction. Moreover, domestic borrowing often compensates for lower levels of revenue and grants (substitutes), highlighting the cost of aid unpredictability and revenue volatility. The moving average representation of the CVAR suggests that unanticipated shocks to foreign aid grants do not have permanent effects on the remaining fiscal variables (purely adjusting). There is also evidence of aid heterogeneity, since aid grants and loans entail different dynamics.

These results suggest some policy implications. It was shown that domestic borrowing has become more volatile in recent years, mainly due to the uncertainty associated with aid inflows and revenue volatility. Interest payments on domestic borrowing have consequently increased to unprecedented levels. Therefore, we argue that donors should make aid inflows more predictable in order to improve medium-term fiscal planning and reduce the need to resort to costly domestic borrowing. Moreover, countercyclical aid inflows have the potential to compensate for revenue shortfalls, avoid domestic indebtedness and help smooth public spending in order to support Ethiopia's development prospects.

3.9 Appendix

An Overview on Cointegration (Johansen, 2005)

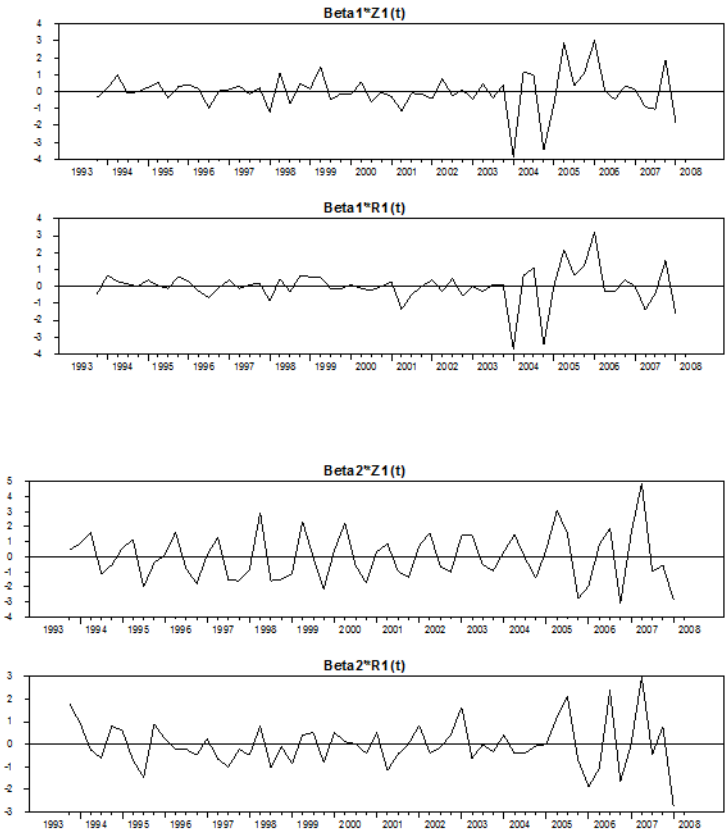
In a single-equation regression context, the estimated coefficients are usually interpreted as total impacts (elasticities if the variables are in logs) via a counterfactual or thought experiment. In the regression below, γ_1 can be thought as the effect of a change in X_t on Y_t, while keeping Z_t constant (*ceteris paribus* assumption).

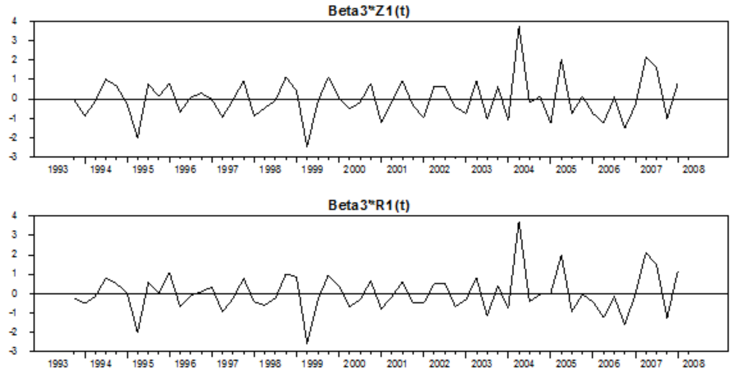
$$Y_t = \gamma_1 X_t + \gamma_2 Z_t + u_t$$

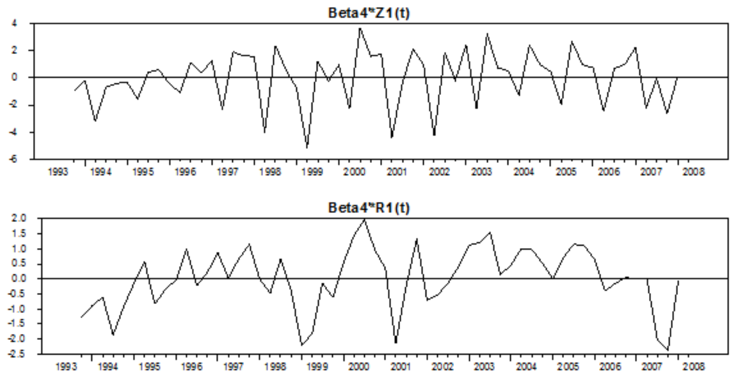
In the CVAR context, the cointegrating relations are interpreted as long-run relations. This means that these relations ‘have been there all the time’ and influence the movement of the endogenous variables, pulling the process towards its mean every time it deviates from the expected value (mean). In a sense, a coefficient in an identified cointegrating relation can be interpreted as the effect of a long-run change to one variable on another, keeping all others fixed. The difference with the usual interpretation of a regression coefficient is that because the relation is a long-run relation, that is, a relation between long-run values, the counterfactual experiment should involve a long-run change in the variables.

Cointegrating Relations

Figure 3-15: Cointegrating Relations

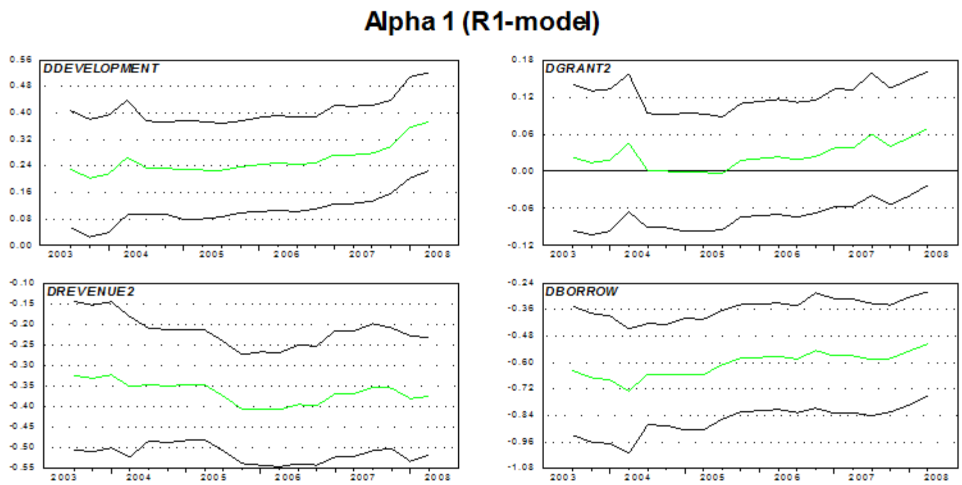




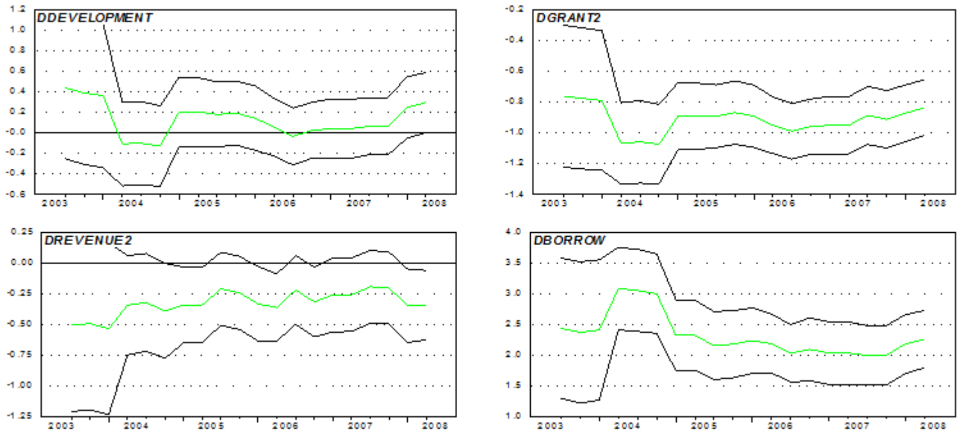


Parameter Constancy

Figure 3-16: Constancy of the Coefficients of Adjustment



Alpha 2 (R1-model)



Alpha 3 (R1-model)

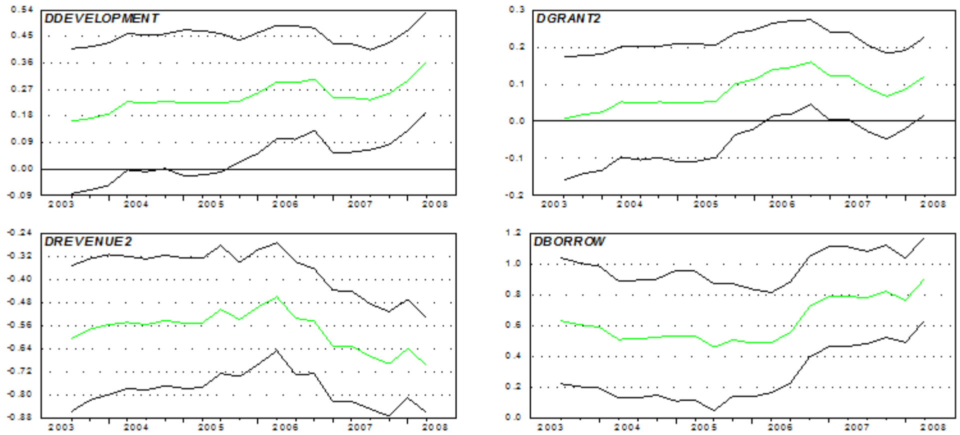
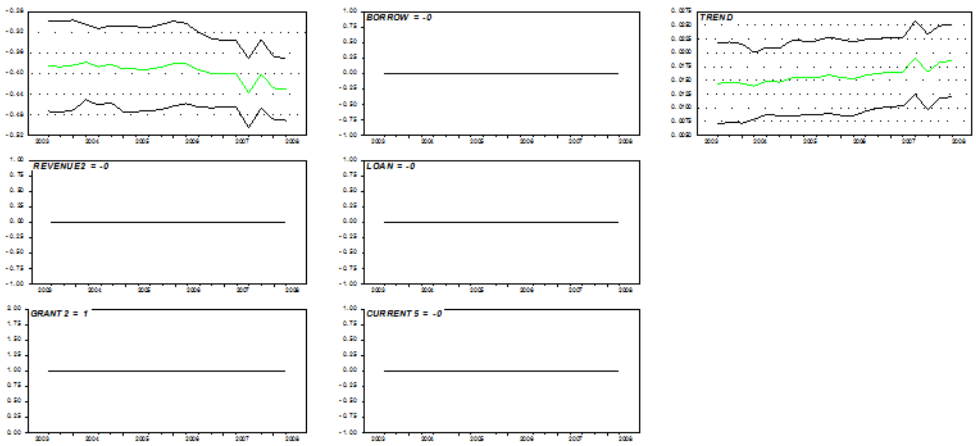
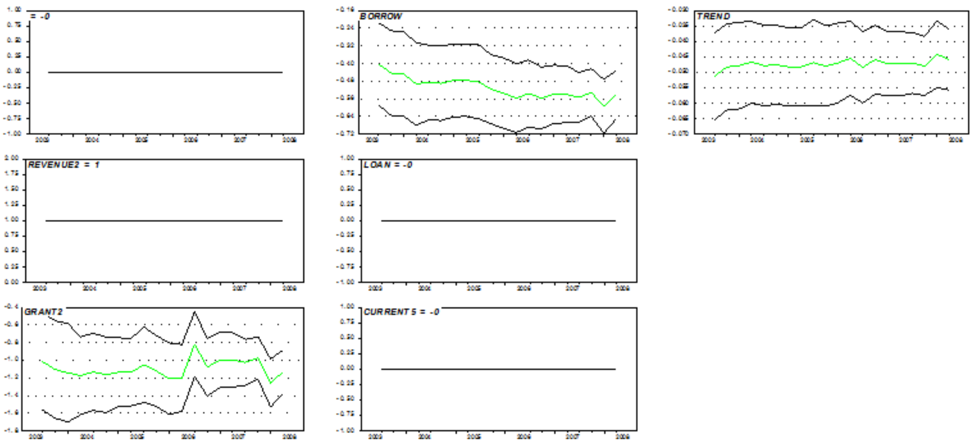


Figure 3-17: Constancy of the Long-Run Coefficients

Beta 2 (R1-model)



Beta 3 (R1-model)



3.10 Bibliography

Engle, R. and C. Granger (1987), “Co-integration and Error-Correction: Representation, Estimation and Testing.” *Econometrica*, 55, 251-276.

Dennis, J. (2006), CATS in RATS: Cointegration Analysis of Time Series (Version 2), manual, Evanston, IL: Estima.

Doornik, J. and H. Hansen, (2008). “An Omnibus Test for Univariate and Multivariate Normality.” *Oxford Bulletin of Economics and Statistics*, 70, 927-939. [or Nuffield Economics Working Paper W4&91 (1994).]

Fagernäs, S. and C. Schurich (2004), “The Fiscal Effects of Aid in Malawi.” ODI/DFID, Economic and Statistics Analysis Unit (ESAU), Working Paper 7.

Fagernäs, S. and J. Roberts (2004a), “The Fiscal Effects of Aid in Uganda.” ODI/DFID, Economic and Statistics Analysis Unit (ESAU), Working Paper 9.

Fagernäs, S. and J. Roberts (2004b), “The Fiscal Effects of Aid in Zambia.” ODI/DFID, Economic and Statistics Analysis Unit (ESAU), Working Paper 10.

Fagernäs, S. and J. Roberts (2004c), “Fiscal Impact of Aid: A Survey of Issues and Synthesis of Country Studies of Malawi, Uganda and Zambia.” ODI/DFID, Economic and Statistics Analysis Unit (ESAU), Working Paper 11.

Franses, P. and Hobijn, B. (1997), “Critical Values for Unit Root Tests in Seasonal Time Series.” *Journal of Applied Statistics*, 24, 25-46.

Franco-Rodriguez, S., M. McGillivray and O. Morrissey (1998), “Aid and the Public Sector in Pakistan: Evidence with Endogenous Aid.” *World Development*, 26, 1241-1250.

Hansen, H. and S. Johansen (1999), “Some Tests for Parameter Constancy in Cointegrated VAR-Models.” *The Econometrics Journal*, 2(2), 306-333.

Hassler, U. and M. Demetrescu (2005), “Spurious Persistence and Unit Roots due to Seasonal Differencing: The Case of Inflation Rates.” *Journal of Economics and Statistics*, 225, 413-426.

Heller, P. (1975), “A Model of Public Fiscal Behaviour in Developing Countries: Aid, Investment and Taxation.” *American Economic Review*, 65, 429-445.

Hylleberg, S. (2006), "Seasonal Adjustment." Economics Working Papers. 2006-04, School of Economics and Management, University of Aarhus.

Hylleberg, S., R. Engle, C. Granger and B. Yoo. (1990), “Seasonal Integration and Cointegration.” *Journal of Econometrics*, 44, 215-238.

Johansen, S (2005), "The Interpretation of Cointegrating Coefficients in the Cointegrated Vector Autoregressive Model." *Oxford Bulletin of Economics and Statistics*, 67, 93-104.

Johansen, S (2002), "A Small Sample Correction for Tests of Hypotheses on the Cointegrating Vectors." *Journal of Econometrics*, 111(2), 195-221.

Johansen, S. (2000), “Modelling of Cointegration in the Vector Autoregressive Model.” *Economic Modelling*, 17, 359-373.

Johansen, S. (1996), Likelihood-based Inference in Cointegrated Vector Autoregressive Models, Advanced Texts in Econometrics, Oxford University Press.

Johansen, S., and K. Juselius (1992). “Structural Hypotheses in a Multivariate Cointegration Analysis of the PPP and UIP for UK.” *Journal of Econometrics*, 53, 211-244.

Johansen, S., and K. Juselius, (1990), “Maximum Likelihood Estimation and Inference on Cointegration – With Applications to the Demand for Money.” *Oxford Bulletin of Economics and Statistics*, 52, 169-210.

Johansen, S. (1988), “Statistical Analysis of Cointegration Vectors.” *Journal of Economic Dynamics and Control*, 12(2-3), 231-254.

Juselius , K. (2007), The Cointegrated VAR Model: Methodology and Applications. Advanced Texts in Econometrics. Oxford University Press.

Kennedy, P. (2003). A Guide to Econometrics, 5th edition. Cambridge, MA: The MIT Press.

Lütkepohl, H. and M. Krätzig (2004), Applied Time Series Econometrics. Themes in Modern Econometrics. Cambridge University Press.

M’Amanja, D., T. Lloyd and O. Morrissey (2005), “Fiscal Aggregates, Aid and Growth in Kenya: A Vector Autoregressive (VAR) Analysis.” CREDIT Research Paper 05/7, University of Nottingham.

Mavrotas, G. and B. Ouattara (2006), “Aid Disaggregation and the Public Sector in Aid-Recipient Economies: Some Evidence from Côte d’Ivoire.“ *Review of Development Economics*, 10(3), 434-451.

McGillivray, M., and O. Morrissey (2001), “A Review of Evidence on the Fiscal Effects of Aid.” CREDIT Research Paper 01/13, University of Nottingham.

McGillivray, M. and O. Morrissey (2000). 'Aid Fungibility in Assessing Aid: Red Herring or True Concern?' *Journal of International Development*, 12(3):413-428.

Mosley, P., J. Hudson and S. Horrell (1987), "Aid, the Public Sector and the Market in Less Developed Economies." *The Economic Journal*, 97, 616-641.

Ndulu, B., S. O'Connell, R. Bates, P. Collier and C. Soludo, eds. (2007), The Political Economy of Economic Growth in Africa, 1960-2000, Cambridge University Press.

Osei, R., O. Morrissey and T. Lloyd (2005), "The Fiscal Effects of Aid in Ghana." *Journal of International Development*, 17, 1037-1053.

Pesaran, H. and Y. Shin (1998), "Generalised Impulse Response Analysis in Linear Multivariate Models", *Economics Letters*, 58, 17-29.

Rahbek, A., E. Hansen and J. Dennis (2002), "ARCH Innovations and their Impact on Cointegration Rank Testing." Department of Applied Mathematics and Statistics, Working Paper 22, University of Copenhagen.

Sims, C. (1980), "Macroeconomics and Reality." *Econometrica*, 48, 1-48.

World Bank (2004), "Ethiopia: Public Expenditure Review: The Emerging Challenge, Volume II, Medium-Term Trends and Recent Developments in Public Spending." Report No. 29338-ET, World Bank: Washington, DC.

4. Do Capital Inflows Hinder Competitiveness? The Real Exchange Rate in Ethiopia

Abstract: African countries have undertaken extensive economic policy reforms since the 1990s. The liberalisation of foreign exchange markets, in particular, has contributed to a growing literature on the determinants of the real exchange rate (RER). This is an important area of research, since the RER is a common measure of a country's external competitiveness. In this study, we are mainly interested in the impact of capital inflows, namely, foreign aid and workers' remittances. Several studies have argued that large capital inflows undermine economic growth through the appreciation of the RER ('Dutch disease'). In this paper, we assess whether this is a valid concern for Ethiopia, one of the largest aid recipients in the world. Moreover, this empirical exercise tries to improve the current literature in a number of ways: (i) the use of quarterly data provides a larger sample size and enables the modelling of important intra-year dynamics, which should lead to better model specifications; (ii) the use of several cointegration approaches enables interesting methodological comparisons; and (iii) the use of a time series model (Unobserved Components) provides a new empirical approach and a robustness check on the econometric models. Overall, the results suggest that there are two main long-run determinants of the RER in Ethiopia: trade openness is found to be correlated with RER depreciations, while a positive shock to the terms of trade tends to appreciate the RER. Foreign aid is not found to have a statistically significant impact, but workers' remittances (proxied by private transfers) could be weakly associated with RER appreciations. The lack of empirical support for the Dutch disease hypothesis suggests that Ethiopia has been able to effectively manage large capital inflows, thus avoiding major episodes of macroeconomic instability. A prudent approach from the Central Bank and aid flows targeted at alleviating supply-side constraints (mostly through public investment) may have played an important role. This research highlights that strong macroeconomic management can reduce the potential negative impact of large aid flows, with clear policy implications for donors and aid recipients.

4.1 Introduction

The term 'Dutch disease' is commonly used to describe the potential negative effects of large inflows of foreign currency on the recipient economy.[87] This 'disease' usually manifests itself through the appreciation of the real exchange rate and the consequent loss of export competitiveness. The surge in foreign exchange often takes the form of higher export receipts (e.g. following an increase in natural resource prices), foreign direct investment, workers' remittances or foreign aid inflows. The main focus of this paper will be on the latter two.

The real exchange rate is one important channel through which foreign aid inflows can affect the recipient economy. Concerns about 'Dutch disease' have been recently revived due to the commitment of the international development community to scale up aid flows to developing countries, and in particular to double the resources to Africa. Evidence that foreign aid has had a detrimental effect on the growth of the export sector could offer an explanation for the difficulty of finding robust evidence that aid fosters economic growth. For example, Rajan and Subramanian (2005) argue that aid flows are responsible for the decline in the share of labour intensive and tradable industries in the manufacturing sector – through its contribution to real exchange rate overvaluation.[88] However, the empirical evidence is mixed, with several studies even suggesting that foreign aid leads to the depreciation of the local currency, potentially through supply side effects or aid tied to imports (Li and Rowe, 2007:17). Moreover, the impact of foreign aid on the composition of (public) expenditure seems to be crucial to the overall effect on the exchange rate. If aid inflows are used to purchase capital goods from abroad (e.g. import support), then they are not likely to have a significant impact on the local currency. However, if the inflows are significantly biased towards the purchase of (non-tradable) local goods, and if there are significant supply-side constraints, then rising domestic inflation will erode the real exchange rate, affecting the competitiveness of the country's exports. These are some of the effects that this empirical exercise will try to uncover in order to improve our understanding of how large aid inflows impact economic performance.

The paper is organised in seven sections. After this short introduction, the theoretical underpinnings of this study will be presented. Section 3 reviews and summarises the

---

[87] The term was originally coined by *The Economist* to reflect the paradoxical impact of the discovery of natural gas deposits in the North Sea on the Dutch manufacturing sector, through the appreciation of the Dutch real exchange rate.

[88] The authors do not find similar effects from remittance flows.

empirical evidence from the ‘Dutch disease’ literature. Section 4 introduces the methodologies to be used in this study, while section 5 draws some considerations about the data. Section 6 presents the empirical results from the econometric models and the structural time series model. Section 7 concludes the paper.

4.2 Theoretical Background

4.2.1 Core Dutch Disease Model

This section deals with the theoretical arguments and predictions of the Dutch disease literature. The core model is described in Corden and Neary (1982:826). Their framework assumes a small open economy with three sectors: (i) the booming export sector (e.g. energy); (ii) the lagging export sector (e.g. manufacturing); and (iii) the non-traded goods sector (e.g. services), which supplies the domestic economy. The price of the non-traded good adjusts to equal supply and demand, in contrast to the export sectors where exogenous world prices prevail. The authors describe two main effects of an export sector boom: (i) spending effect; and (ii) resource movement effect. To better understand the *spending effect*, assume that the energy sector does not use any labour. The energy boom will entail higher export earnings and thus higher foreign exchange inflows. It is unlikely that all this extra income will be spent on imports, thus the boom will have an impact on the domestic economy. Provided that the demand for non-tradables rises with income, the boom will create excess demand in the non-tradable sector, increasing their price and leading to real exchange rate appreciation.[89] Turning to the *resource movement effect*, assume that the income-elasticity of demand for non-tradables is zero. The increased marginal productivity of labour in the booming export sector raises profitability and the demand for labour in the energy sector. This will push wages up and eventually induce a relocation of labour to the booming sector at the expense of the other two. The fall in employment in the manufacturing sector will contract output (‘direct deindustrialization’) and the higher price of non-traded goods leads to real exchange rate appreciation. Through these two effects the authors suggest that the traditional export sector is crowded-out by the other two sectors (Ebrahim-zadeh, 2003).

[89] The price of traded goods in the world market remains unchanged (small economy assumption).

This core model can be adapted to understand the potential impact of a surge in aid inflows, rather than an energy boom (Nkusu, 2004:9).[90] Foreign aid can be seen as a real income transfer that will raise the demand for both tradable and non-tradable goods produced in the economy. In the context of a small open economy, the increase in the demand of tradable goods will not affect their prices, since these are exogenously determined in world markets. However, the increased demand for non-tradable goods will place an upward pressure on prices, hence, the real exchange rate appreciates (*spending effect*). The profits of exporters are squeezed as the prices for domestic inputs (services, labour, etc.) increase, therefore discouraging the production of tradable goods. Meanwhile, the relative incomes of producers of non-tradable goods are increased.

The appreciation of the RER will materialise regardless of the exchange rate regime. In a *flexible exchange rate regime*, an increase in aid inflows will appreciate the nominal exchange rate as foreign currency is swapped for domestic currency. This appreciation will reduce the value of exports in local currency, which will reduce profitability if production costs such as wages are not adjusted downwards. In a *fixed exchange rate regime*, aid flows will push up the price of domestically produced goods that are in short supply. In both cases, the aid surge will induce an appreciation of the real exchange rate, thus hindering the development of the tradable sector of the economy.

Furthermore, there is also a *resource movement effect*. The increased demand for non-traded goods will push wages up in the sector, bidding labour out of the tradable goods sector – where wages cannot be increased without squeezing the profits. As the marginal product of labour and wages in the non-tradable sector increase, there will be an incentive to relocate labour from the (lagging) tradable sector to the non-tradable sector. The final effect will depend on the factor intensity of the sectors.

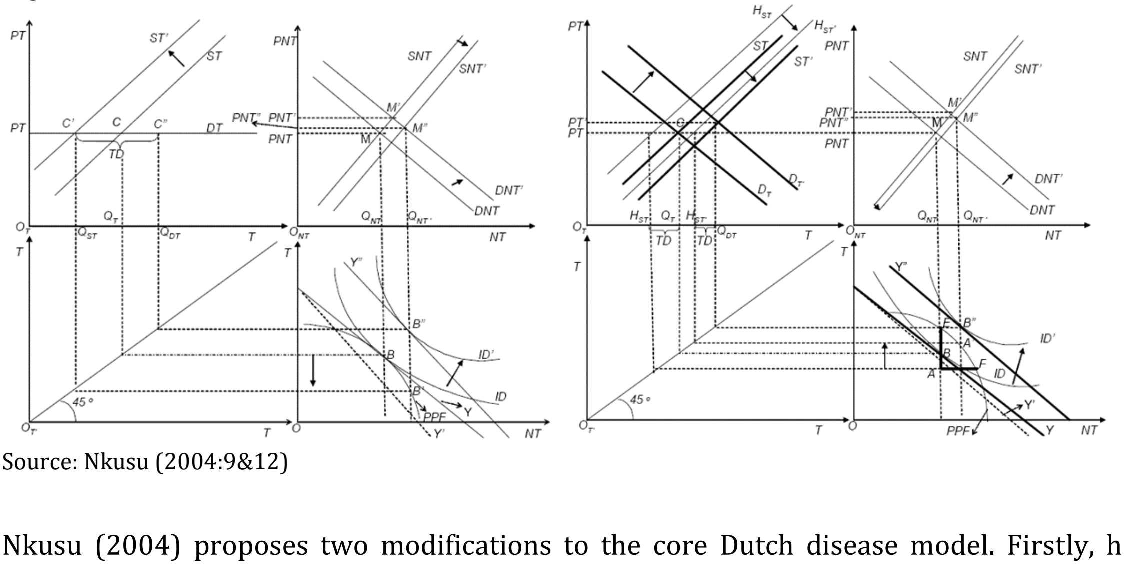
Despite the popularity of the arguments presented above, the empirical support for Dutch disease effects has been mixed, especially with regard to aid inflows (as will be demonstrated in section 3). This apparent disagreement between theory and empirical evidence is bridged by Nkusu (2004). The author provides a modified theoretical framework to show that, under certain circumstances, Dutch disease symptoms may not materialise. He uses a Salter-Swan framework with two sectors to illustrate the importance of two crucial assumptions of the core Dutch disease model: (i) full and efficient employment of produc-

[90] Michaely (1981) is an earlier reference using a two-sector small-country model.

tion factors (i.e. countries produce on their production possibility frontier) and, (ii) perfectly elastic demand for tradables (small-country assumption).

The figure below presents the core Dutch disease model (left panel) and Nkusu's modifications (right panel). The top-left quadrant of each panel represents the tradables market – *PT* is the price and *T* the quantity – while the top-right quadrant illustrates the market for non-tradables. Moreover, the bottom-right quadrant presents the production possibility frontier (PPF). The left panel provides the setting for the core Dutch disease model. The small-country assumption is reflected on the perfectly elastic demand for tradables (horizontal *DT*), thereby making its price exogenous. The trade balance is in equilibrium (*C*). In addition, the country produces and consumes on the PPF (*B*). Hence, a large inflow of foreign aid will increase the demand for non-tradables (*DNT* to *DNT'*) and its price (*PNT* to *PNT'*). Since the price of tradables (*PT*) remains at the same level, the RER will increase and discourage the production of tradables. This is usually referred to as the *spending effect*. Furthermore, the consequent increase in the supply of non-tradables (*SNT* to *SNT'*) will occur at the cost of the tradables sector, which contracts from *ST* to *ST'*. This is the *resource movement effect*, which is mainly due to the relocation of labour to the non-tradables sector. Nkusu (2004) also makes reference to an *expenditure-switching effect*, which is a disincentive to buy non-tradables due to the RER appreciation. The increase in income (*Y* to *Y''*) and the shift from indifference curve *ID* to *ID'* will be consistent with a higher demand for tradables, which will create a trade deficit of magnitude *C'C''*.

Figure 4-1: The Core Dutch Disease and Modified Models



Source: Nkusu (2004:9&12)

Nkusu (2004) proposes two modifications to the core Dutch disease model. Firstly, he argues that many low-income countries (LICs) produce below their potential, and not on the PPF. On the right panel, the country produces at *A* (within the PPF) and consumes at *B*.

The PPF assumption is synonymous of full-employment and efficient use of the production factors, which is clearly not the case for many LICs suffering from supply-side constraints – high structural unemployment and inefficient use of production factors seems a more plausible assumption. By relaxing this assumption, the extra demand for non-tradable goods brought by aid flows can be accommodated without creating inflationary pressures (especially if aid flows are used to improve productivity).

Secondly, he argues that the small-country assumption is not realistic with regard to many domestically produced importables in LICs. The assumption states that the price of tradable goods is exogenously determined, and hence a surge in aid will place upward pressures on wages, squeezing out profits. However, the author argues that the threat of de-industrialisation is mitigated in LICs where there is imperfect substitutability between domestically produced manufactured goods and imported ones. The imperfect substitutability allows domestic manufacturers to raise prices and increase supply in response to domestic market conditions, such as increased demand, regardless of whether they use imported or domestically produced inputs. Therefore, the demand of tradable goods is represented by a downward-sloping curve (*DT*). The total supply of tradables (*ST*) includes home supply (H_{st}) and imports equivalent to the trade deficit (*TD*). The initial dynamic impact of an increase in aid inflows is similar to the core model. The increase in the demand for non-tradables (*DNT*), and its respective price (*PNT'*), leads to an increase in supply (*SNT*). The country now produces at *A'* (on the PPF) and consumes at *B''*. In the modified version, however, the supply of tradables is actually increased (no 'de-industrialisation'), the RER may not appreciate, and the trade balance can be improved.

In summary, Nkusu (2004) suggests that if low-income countries (LICs) can "draw on their idle productive capacity to satisfy the aid-induced increased demand", the appreciation of the exchange rate will be negligible. His theoretical contribution seems to explain why Dutch disease symptoms are not necessarily present in episodes of resource boom and aid surges.[91] This line of investigation has been taken up by several calibrated general equilibrium models, as described in section 3.

[91] Moreover, Torvik (2001) uses a learning-by-doing model (with spillovers between the sectors) to show that, in the long-run, a 'foreign exchange gift' induces a real exchange rate depreciation.

4.2.2 Determinants of the Real Exchange Rate

The exchange rate is a critical price for an open economy. The exchange rate affects the volume of both imports and exports (by changing their relative prices), as well as the stock of foreign debt in domestic currency terms. In fact, all transactions with the rest of the world can be potentially affected by the level of the exchange rate. A depreciation of the exchange rate is often associated with competitiveness gains, in the sense that the relative price of exports will fall, therefore becoming more attractive to foreign importers. Since imports become relatively more expensive, we usually observe an improvement of the trade balance.[92] However, since the stock of foreign debt becomes more expensive in local currency, currency depreciations usually worsen a country's debt position and increase interest payments. Moreover, foreign direct investments may benefit from relatively cheaper domestic goods, but revenues in local currency will translate into less foreign currency to repatriate. Finally, foreign aid flows will be able to purchase more domestic goods than before. In addition to the large literature on the merits and shortcomings of devaluations, there is an understanding that excess exchange rate volatility and misalignment can have substantial negative welfare effects. It is therefore not surprising that there is a vast literature trying to uncover the main determinants of the real exchange rate, in order to improve exchange rate policies.

There are two main approaches to understand the behaviour of the real exchange rate (RER). The traditional approach is based on the Purchasing Power Parity (PPP) theory, which defines the RER as the nominal exchange rate corrected by the ratio of the foreign price level to the domestic price level (purchasing power):

$$RER_{PPP} = NER \times \frac{P^*}{P}$$

where P^* and P are foreign and domestic prices, respectively, and the NER is expressed as units of domestic currency per unit of foreign currency. Therefore, a rise [fall] in the RER represents a depreciation [appreciation] of the domestic currency. Depending on the price indices used in the computations, the RER can be seen as the relative price of foreign to domestic consumption or production baskets (Edwards, 1989:5).

[92] Provided that the Marshall-Lerner condition is met, i.e. the sum of the absolute export and import price elasticities is greater than unity.

The absolute PPP equilibrium condition for a pair of currencies is achieved when the nominal rate ensures that the domestic purchasing powers of the currencies are equivalent. This implies that the RER_{PPP} is equal to 1.

$$NER = \frac{P}{P^*}$$

Hence, the PPP theory is grounded on the law of one price, which suggests that market forces acting on either the prices or the NER (depending on the exchange rate regime) will ensure that relative prices equalise across countries or at least that we have negligible price differentials (Hallwood and MacDonald, 2000:122). Empirically, it is assumed that this PPP relation holds in the long-run with strong mean reversion (stationarity). Another consequence of this approach is that the equilibrium real exchange rate (ERER) is given by a scalar (constant), which is assumed to hold for the entire period. The equilibrium value is often found by looking at the RER for a relative stable period, when external equilibrium was likely to hold. One way to empirically test this hypothesis is to use regression analysis on the following equation:[93]

$$\ln NER = \alpha + \beta_1 \ln P + \beta_2 \ln P^* + \varepsilon_t$$

If absolute PPP holds, then β_1=1, β_2=–1 and α=0. However, empirical work has often found little evidence of the validity of this relationship. Movements outside the 'equilibrium' appear to be persistent, indicating that mean reversion is slow (PPP relation is non-stationary). Hence, the effects of a shock take too long to disappear, which is not compatible with the 'arbitrage condition' (i.e. law of one price). This 'PPP puzzle' (Rogoff, 1996) can be explained by a number of reasons. Trade barriers (in the form of quantitative restrictions or taxes), transaction costs and productivity gains are amongst the factors that are likely to influence the path of the RER, thus it is unlikely that a single value of the RER can be seen as 'equilibrium'.

The weaker version of the PPP (relative PPP) accepts that due to market imperfections the relationship may not hold. Instead, this hypothesis states that the percentage change in the NER will equal the inflation differential (in the equation below, β_3=1 and β_4=–1). Hence, an increase in relative prices will force the NER to depreciate (law of one price). However, it should be noted that this theory requires that factors such as trade barriers and transac-

[93] Equations from Hallwood and MacDonald (2000:136).

tion costs remain constant through time, an assumption that is not likely to hold. In fact, there has been only weak support for this hypothesis.

$$\Delta \ln NER = \beta_3 \Delta \ln P + \beta_4 \Delta \ln P^* + \upsilon_t$$

This paper uses an alternative approach to the PPP theory. The equilibrium real exchange rate (ERER) is defined by Edwards (1989:5&8) as the domestic relative price of tradable goods to non-tradable goods that simultaneously attains internal and external equilibrium:

$$RER = \frac{P_T}{P_{NT}}$$

where P_T is the price of tradables (expressed in local currency) and P_{NT} the price of non-tradables. Internal equilibrium is defined as the clearing of the non-tradable goods market, hence with employment at the 'natural' level. External equilibrium is achieved when current account balances are compatible with long-run sustainable capital flows. This definition implies that the ERER is not a constant number, as it depends on a number of real and nominal determinants. It is also important to distinguish between the short-run and the long-run, since some determinants may only have a temporary impact on the ERER. Misalignment is defined as "sustained departures of the actual real exchange rate from its [long-run] equilibrium level" (Edwards, 1989:15). For example, during the 1980s several developing countries had overvalued real exchange rates.

Edwards (1989) constructs a benchmark intertemporal general equilibrium model to analyse the theoretical impact of a number of real disturbances on the ERER. The assessment suggests that an increase in tariffs (trade protectionism) will usually generate an equilibrium real appreciation, while a relaxation of exchange controls (capital account liberalisation) will induce an initial equilibrium real depreciation. Transfers from abroad (e.g. foreign aid) will always result in ERER appreciation. The total effect of other determinants is ambiguous. For example, a worsening of the terms of trade will result in ERER depreciation if the income effect dominates the substitution effect, while the impact of government consumption will depend on its composition – if mainly tradables, then depreciation will ensue. Finally, the impact of technological progress will depend on how the demand (income) and supply effects play out.

4.3 Literature Review

Notwithstanding the theoretical arguments put forward by Corden and Neary (1982), Corden (1984), van Wijnbergen (1984, 1986)[94] and Edwards (1989), it is has been difficult to establish a robust association between increased aid inflows and the appreciation of the real exchange rate. This section provides an overview of the evidence on the Dutch disease, with special reference to foreign aid inflows. It starts with a brief presentation of the findings of computable general equilibrium (CGE) studies, followed by a survey of the empirical evidence from both cross-country and time series econometrics.[95]

Computable General Equilibrium

Computable general equilibrium (CGE) models are a useful tool to explore the different dynamics and transmission channels surrounding the Dutch disease hypothesis.[96] Vos (1998) finds that foreign aid inflows induce strong Dutch disease effects in Pakistan, although these symptoms can be mitigated if the flows are used to alleviate investment constraints in the production of traded goods. He also argues that, based on similar models for Mexico, the Philippines and Thailand, the findings for Pakistan are not "easily generalisible and depend strongly on the existing economic structure, investment and savings behaviour of institutional agents, and the allocation of additional capital flows among public and private sector agents." Bandara (1991:92) also highlights the fact that the standard Dutch disease results depend, to some extent, on the model assumptions, parameter values, and model closure. Moreover, Benjamin (1990) demonstrates that despite the RER appreciation, the tradable sector (manufacturing) only contracts in the short-run, consequently expanding via new investments. Laplagne et al (2001) calibrate a CGE for a typical South Pacific microstate and also find evidence of Dutch disease effects, in particular of the relative contraction of the tradables sector. Collier and Gunning (1992) suggest that foreign aid inflows tend to inhibit the export supply response to liberalisation through a spending effect (relocation of investment and labour into non-tradables).

---

[94] van Wijnbergen (1984:53) argues that Dutch disease is a 'disease' because "corrective medicine is needed".
[95] CGE simulations do not provide empirical tests of the Dutch Disease hypothesis. However, they can offer insights into the potential size of such effects, provided that the model is a reliable representation of the economy under consideration.
[96] Some studies focus on mitigating policy responses rather than testing the empirical validity of such effects, which are often built-in by construction or dependent on crucial assumptions about the structure of the economy.

Over the last couple of years there has been a renewed interest in the subject, partly due to donor commitments to scale up aid resources to developing economies. Adam and Bevan (2006) develop a small dynamic CGE model (calibrated for Uganda) with learning-by-doing externalities where investment in public infrastructure creates an intertemporal productivity spillover. Their simulations show that the supply-side effects of aid-financed public expenditure can outweigh the short-run (demand-side) Dutch disease symptoms, especially if these investments are biased towards the non-tradable sector – therefore increasing the productivity of private factors in the production of domestic goods. Devarajan et al (2008) use a standard neoclassical growth model, based on the Salter-Swan open-economy framework.[97] They suggest that concerns of Dutch disease do not materialise if recipients are able to optimally plan consumption and investment over time, a result that is valid for permanent and temporary aid shocks, while not requiring "extreme assumptions or additional productivity story." However, Dutch disease might become a concern if aid flows are disbursed in an unpredictable or volatile fashion (affecting intertemporal smoothing). Hence, they conclude that "any unfavourable macroeconomic dynamics of scaled-up aid are the result of donor behaviour rather than the functioning of recipient economies." Cerra et al (2008) use a dynamic dependent-economy model to show that while untied aid causes a temporary (and small) appreciation of the real exchange rate in the short-run, it does not induce Dutch disease effects in the long-run (i.e. no relative price effects), as long as capital is perfectly mobile between sectors.[98] In contrast, tied aid will cause permanent relative price changes, but the effects will depend on the sectoral allocation of aid: aid flows directed to enhancing productivity in the traded sector will lead to an appreciation of the real exchange rate (Balassa-Samuelson effect), while if directed to the non-traded sector they will lead to a depreciation.

Finally, some studies focus their attention on macroeconomic management policies that have the potential to mitigate the undesirable effects of aid inflows, such as RER appreciation and volatile expenditure patterns. Prati and Tressel (2006) investigate optimal fiscal and monetary policy responses in the context of a model with a closed capital account and learning-by-doing externalities, where foreign aid inflows impact productivity growth through (positive) public expenditure and (negative) Dutch disease effects. They argue that although foreign aid inflows have a negative impact on exports, this effect can be undone by tighter macroeconomic policies. This is achieved through the reduction of the net domestic assets of the central bank ('sterilisation'), which reflects both monetary and

[97] The model uses data from Madagascar, Mozambique, and the Philippines.
[98] The model is calibrated for a representative economy.

fiscal policies. This policy can be effective in preventing RER appreciation when aid is excessively 'front-loaded' (when Dutch disease costs are higher than expenditure and productivity benefits), and thus the economy is better-off saving part of the aid flows for future use. Nonetheless, when foreign aid is excessively 'back-loaded' expansionary policies can improve welfare if the stock of international reserves is large enough. Similarly, Buffie et al (2008, 2004) use an intertemporal optimising model to investigate the monetary and exchange rate policy options available to countries facing persistent shocks to aid flows. They conclude that a 'pure float' regime performs very poorly (RER overshoots), while in a 'crawling peg' regime inflation increases in the short-run. Bond sterilisation will control inflation, but the interest burden will rise. The preferred scenario is a 'managed float', where the central bank "uses unsterilised foreign exchange intervention to target the modest real appreciation needed to absorb the aid inflow." This scenario ensures low real interest rates and a quick macroeconomic adjustment, challenging suggestions that the expansion of domestic liquidity brought by aid flows requires a combination of sterilisation and a free float.

In conclusion, the CGE literature seems to suggest that while Dutch disease remains a concern for macroeconomic management in the short-run, these negative effects can be mitigated, or even reversed, if foreign aid flows induce positive supply-side effects. Moreover, there are a number of policy tools that can be deployed in order to mitigate the potentially pervasive effects of a windfall of foreign aid. The challenge for empirical econometric studies is to assess whether there is robust evidence of Dutch disease effects, or if these are undone (or at least mitigated) through appropriate macroeconomic policies (e.g. fiscal and exchange rate).

Cross-Country

Most of the empirical studies reviewed in this sub-section use single-equation panel data regressions. Recent works have used dynamic panel data techniques, while one study has employed cointegration analysis. Adenauer and Vagassky (1998) use a panel of four CFA Franc Zone countries (Burkina Faso, Côte d'Ivoire, Senegal and Togo) to estimate the impact of foreign aid flows on the RER. Their GLS results suggest that both foreign aid and the terms of trade are correlated with RER appreciations. Moreover, real GDP and a proxy for technological differentials[99] were not statistically significant. Ouattara and Strobl

[99] They use a measure capturing the growth rate differences between the country and industrial countries (the Balassa-Samuelson effect).

(2004) expand the previous analysis to 12 countries of the CFA Franc Zone and obtain quite different results. Their application of a dynamic panel data (DPD) estimator indicates that foreign aid flows (ODA) and the ratio of exports and imports to GDP (proxy for openness) tend to cause the RER to depreciate. In contrast, the ratio of government consumption to GDP, the terms of trade, and the ratio of domestic credit to GDP (proxy for monetary policy) induce RER appreciation. Lartey (2007) also utilises a DPD setting for 16 sub-Saharan African countries over the period 1980-2000. He finds that while FDI and ODA flows lead to a RER appreciation, other capital flows do not have a significant impact. Moreover, government expenditure also seems to cause the RER to appreciate, while openness has the opposite effect. Excess money growth (difference between the growth rates of M2 and GDP) is not significant in most specifications.

Table 4-1: Results from Selected Cross-Country 'Dutch Disease' Studies

Main Studies	Sample	Aid	Methodology	RER
Mongardini & Rayner (2009)	36 SSA (1980-06)	DAC	Panel (PMG)	–
Lartey (2007)	16 SSA (1980-00)	WB	Panel (DPD/GMM)	+
Ouattara & Strobl (2004)	12 CFA Franc Zone (1980-00)	DAC	Panel (DPD/GMM)	–
Elbadawi (1999)	62 developing (1990 & 95)	DAC	Panel (RE, FE, IV)	+
Yano & Nugent (1999)	44 aid-dependent (1970-90)	DAC	Mixed (TS, CS)	+
Adenauer & Vagassky (1998)	4 CFA Franc Zone (1980-92)	WB+	Panel (GLS)	+

Obs.: '+' appreciation, '–' depreciation, PMG Pooled Mean Group, DPD Dynamic Panel Data, GMM Generalised Method of Moments, RE Random Effects, IV Instrumental Variables, TS Time Series, CS Cross-Section.

Elbadawi (1999) presents results from a panel of 62 developing countries (28 of them from Africa). He finds that, apart from the degree of openness, all other variables contribute to the overvaluation of the RER, namely: ratio of ODA to GNP, ratio of net foreign income to GNP, index of productivity, terms of trade, and the ratio of government consumption to GDP. Yano and Nugent (1999) argue that aid flows are associated with an expansion of the non-traded sector, thus explaining the transfer paradox. Mongardini and Rayner (2009) use the pooled mean group (PMG) estimator for a panel of 36 sub-Saharan African countries. Their results suggest that grants are associated with RER depreciations, while remittances do not have a statistically significant effect. Moreover, terms of trade and openness are correlated with appreciation and depreciation, respectively. Finally, Prati et al (2003) estimate the impact of aid flows on the real black market effective exchange rate for several developing countries for the period 1960-1998. They use fixed-effects as well as Arellano-Bond's GMM estimator. Other variables include a commodity price index, and the central bank's net domestic assets. ODA is assumed to be stationary on 'economic grounds'. The authors find that aid flows tend to appreciate the black-market RER.

Time Series

This section provides a summary of empirical studies using time series methods. Following recent developments in time series econometrics, most studies reported in this section use cointegration analysis to avoid inference based on spurious relations. Moreover, this strategy has the advantage of separating the long-run (steady-state) information from the short-run dynamics. The empirical results (for the long-run) with regard to the impact of aid inflows on the RER are reported in the table below. Unless otherwise stated, 'openness' is measured by the ratio of total trade (export plus imports) to GDP and 'terms of trade' by the ratio of export to import price indices. It should also be noted that, perhaps surprisingly, no study has reported an insignificant impact of aid flows on the RER.[100]

Table 4-2: Results from Selected Time Series 'Dutch Disease' Studies

Main Studies	Sample	Aid	Methodology	RER
Issa & Ouattara (2008)	Syria (1965-1997)	DAC	UECM (OLS)	–
Li & Rowe (2007)	Tanzania (1970-05)	DAC	EG (FM-OLS)	–
Bourdet & Falck (2006)	Cape Verde (1980-00)	WB	EG (OLS)	+
Opoku-Afari et al (2004)	Ghana (1966-00)	DAC	VECM (MLE)	+
Sackey (2001)	Ghana (1962-96)	DAC	UECM (OLS)	–
Nyoni (1998)	Tanzania (1967-93)	DAC	UECM (OLS)	–
Ogun (1998)	Nigeria (1965-90)	n/a †	Differences (OLS)	(+)
White & Wignaraja (1992)	Sri Lanka (1974-88)	Local ‡	Levels (OLS)	+

Obs.: '+' appreciation, '–' depreciation, ARDL Auto Regressive Distributed Lag, EG Engle-Granger Two-Step Approach, OLS Ordinary Least Squares, VECM Vector Error Correction Model, MLE Maximum Likelihood Estimator, UECM Unrestricted Error Correction Model, FMOLS Fully-Modified OLS. † Capital Flows, ‡ Includes Remittances. Younger (1992) does not undertake an econometric exercise (Ghana).

The results from Bourdet and Falck (2006) for Cape Verde, Opoku-Afari et al (2004) for Ghana, Ogun (1998) for Nigeria, and White and Wignaraja (1992) for Sri Lanka seem to suggest that foreign aid inflows are associated with appreciations of the real exchange rate. Bourdet and Falck (2006) investigate the determinants of the Cape Verdean real exchange rate. Their results suggest that remittances (ratio of private transfers to GDP), aid inflows (ratio of net ODA to GDP), the terms of trade, growth of excess credit[101] and technological change (time trend) are the factors that contribute to RER appreciations, while openness has the opposite effect. Opoku-Afari et al (2004) reach similar conclusions for Ghana. However, their study goes beyond the estimation of a single-equation and uses a system-based methodology. They use three alternative measures of capital flows: (i) inflows that require repayment (e.g. aid loans, foreign debt, etc.); (ii) 'permanent' flows (i.e. grants, remittances and FDI); and (iii) aid grants and loans. The capital inflow measure

[100] This could be seen as a symptom of 'publication bias'.
[101] Rate of growth of domestic credit minus rate of growth of real GDP lagged one period.

and the terms of trade are shown to appreciate the RER, while openness[102] and technological change (measured by total factor productivity) tend to depreciate the RER.[103]

Ogun (1998) also attempts to estimate the long-run determinants of the RER, but fails to find a cointegrating relation amongst the variables. He suggests that, in the case of Nigeria, the 'fundamentals' (real variables) only have short-run effects on the RER. His results indicate that net capital inflows (proxied by the capital account balance), government expenditure on non-tradables and excess credit[104] tend to appreciate the RER, while openness,[105] technological progress (proxied by the growth rate of real income) and nominal devaluations of the NER induce a RER depreciation. The (income) terms of trade have no significant impact. Finally, White and Wignaraja (1992) investigate the failure of nominal devaluations to translate into a competitive RER in Sri Lanka, but do not use cointegration analysis. The authors find that the growing divergence between the real and nominal exchange rates can be attributed to the increase in total transfers (aid and remittances), which have created such upward pressures on the RER (via high domestic inflation) that the devaluation strategy pursued had little effect on competitiveness. Their results also suggest that the terms of trade and changes in the nominal exchange rate tend to depreciate the RER, while total borrowing, investment, and excess credit creation have no significant impact.

In contrast to the evidence presented above on Dutch disease, the findings from Issa and Ouattara (2008) for Syria, Li and Rowe (2007) for Tanzania, Sackey (2001) for Ghana, and Nyoni (1998) for Tanzania suggest that foreign aid flows are associated with RER depreciation, rather than appreciation. Issa and Ouattara (2008:137) base their model on Edwards (1989) to analyse the determinants of the RER for Syria. Their results show that net ODA flows are associated with RER depreciation, while openness and GDP per capita have the opposite effect. The terms of trade, government expenditure, and growth of M2 (proxy for expansionary monetary policies) are not significant in their (long-run) specification. Li and Rowe (2007) argue that in Tanzania foreign aid and openness create downward pressures in the RER, while an improvement in the terms of trade (real com-

---

[102] They also use the imports-to-GDP ratio and the ratio of imports to domestic absorption (i.e. GDP plus imports minus exports) as alternative measures.

[103] We ought to be careful with the interpretation of the results, since inference from a VAR is not straightforward.

[104] Ratio of domestic credit to M2 minus rate of growth of real income minus log NER minus log US WPI (used as an index of macroeconomic imbalances).

[105] Here measured by the parallel market exchange rate premium, which is used as an index of the severity of trade restrictions and capital controls. Ogun (1998:11) suggests that this variable is preferred to the usual trade ratio because of the dominating influence of oil exports in the country's total exports.

modity price index) seems to cause the appreciation of the RER. Moreover, the ratio of real GDP per capita in Tanzania relative to its major trading partners (to measure relative productivity differentials, and capturing Balassa-Samuelson effects), and government investment as a percentage of GDP were not significant. Central bank reserves were included in the error correction specification (short-run) and found to depreciate the RER. Moreover, Sackey (2001) shows that the terms of trade and net ODA create downward pressures on the RER, while government consumption as a share of GDP, the parallel market premium (proxy for commercial policy stance), and the index of agricultural production (technological progress) entail RER appreciation. The result for the terms of trade implies that the substitution effect dominates the income effect. Nominal devaluations lead to (short-run) RER depreciation. Finally, Nyoni (1998) finds that net ODA and increased openness induce RER depreciation, while increased government expenditure as percentage of GDP caused the appreciation of the RER. External terms of trade and technological change (proxied by a time trend) do not have a long-run impact on the RER, while devaluations of the local currency have an expected (short-run) impact on the RER.

Overall, these studies seem to point to important empirical regularities: the terms of trade and government expenditure seem to appreciate the RER, while openness has an opposite effect. However, with regard to the impact of aid inflows, the results seem to be mixed (even when the same country is under scrutiny, e.g. Ghana). These apparent contradictions of the (long-run) impact of aid inflows on the RER may be explained by a number of factors: (i) the impact on competitiveness depends on the structure of the economy and/or country-specific aid dynamics; (ii) some studies may ignore other aspects that affect the RER (other flows, trade policies, etc); and (iii) the use of different methodologies (e.g. cointegration).

Table 4-3: Main RER Determinants

Main Studies	Aid	ToT	Open	ExM	GE	TP	RES	dNER
Issa & Ouattara (2008)	– (–)	0 (0)	+ (+)	0 (0)	0 (–)	+ (+)		
Bourdet & Falck (2006)	+ (•)	+ (•)	– (•)	+ (•)		+ (•)		
Li and Rowe (2007)	– (–)	+ (0)	– (–)		0(0)	0 (0)	• (–)	
Opoku-Afari et al (2004)	+	+	–		0	–		
Sackey (2001)	– (–)	– (0)			+ (+)	+ (+)		• (–)
Nyoni (1998)	– (–)	0 (0)	– (–)	• (0)	+ (+)	0 (•)		• (–)
Ogun (1998)	• (+)†	• (0)	• (–)	• (+)	• (+)	• (–)		• (–)
White & Wignaraja (1992)	+	–		0				–
Mongardini & Rayner (2009)	–	+	–			+		
Lartey (2007)	+		–	0	+			
Ouattara & Strobl (2004)	–	+	–	+	+			
Elbadawi (1999)	+	+	–		+	+		
Yano & Nugent (1999)								
Adenauer & Vagassky (1998)	+	+				0		

Obs.: †Net capital inflows, '+' appreciation, '–' depreciation. '0' negligible, '•' not applicable. Short-run effects are in brackets. ExM Excess money or domestic credit, GE Government expenditure (composition depends on the paper), TP Technological change or productivity, RES change in reserves.

However, there are a number of other related concerns that can affect the model estimates: (i) the scarce number of observations; (ii) potential structural breaks; (iii) the composition and timing for the aid variable; and (iv) endogeneity. We take these in turn. Amongst the studies surveyed here, the largest sample contains 35 yearly observations, which can be a problem if the model includes several regressors and a long lag structure. Moreover, most samples are likely to contain structural breaks, since they include periods where exchange rate markets were highly regulated and the macroeconomic policies pursued were rather different (mainly 1970s and 1980s). In this regard, the use of quarterly data will enable us to analyse the behaviour of the RER over a shorter time period (1995-2008), and therefore avoid major structural breaks. In terms of the aid variable, data is usually taken from OECD-DAC. The problem, however, is that data reported from donors is likely to include items that do not have an impact on the exchange rate. For example, aid in kind (food aid) is not likely to have a significant impact on the real exchange rate, while a substantial share of technical assistance payments do not even leave the donor country. Another issue relates to the timing of transactions, since donors may record disbursements in a different period from the recipient country. Ideally, we should recover data on grants and concessional loans from the central bank's balance of payments statistics. Finally, the single-equation approach may impose strong exogeneity conditions on the regressors. The estimates can be significantly biased in case there are unmodelled feedback effects from the RER to other variables. The only study analysed here that uses a system approach is Opoku-Afari et al (2004). However, potential misspecification errors in one equation of the system would be propagated to the entire model,[106] while its finite-sample properties may be undesirable (Greene, 2003:413). The decision to use single-equation frameworks is based on two main premises: (i) the argument for endogeneity of most explanatory variables used in this paper is not particularly strong (e.g. remittances or aid flows are not likely to be responsive to the RER level); and (ii) some of the cointegration methods used in this paper provide corrections for endogeneity.

4.4 Methodology

While the core Dutch disease model developed by Corden and Neary (1982) is an important reference point for analytical assessments of the impact of capital inflows on the RER, empirical investigations have traditionally used the equilibrium real exchange rate (ERER) approach proposed by Edwards (1989). Moreover, the methodologies used in this paper follow recent developments in time series econometrics and thus incorporate cointegration analysis. The problem of spurious regressions is highlighted in Granger and Newbold (1974) and Davidson and MacKinnon (1993). They provide Monte Carlo simulations to illustrate how independent variables can very often have statistically significant t-ratios. This happens because the t-statistic does not have a limiting standard normal distribution (and the behaviour of R^2 is non-standard) in the presence of non-stationarity (Wooldridge, 2003). Hence, most modern econometric analysis starts by assessing the properties of the variables, namely the order of integration of the each series. This is usually done through the use of augmented Dickey-Fuller (ADF) tests on univariate AR(p) processes with appropriate deterministic components. Provided that some variables are found to be non-stationary, cointegration analysis should be undertaken. Another advantage of using a cointegration framework is that we may be able to separate the long-run effects from the short-run. This is crucial to disentangle the ERER from the disequilibrium RER.

Edwards (1989:133-7) suggests that the dynamic behaviour of the RER can be captured by:

$$\Delta \ln e_t = \theta(\ln e_t^* - \ln e_{t-1}) - \lambda(Z_t - Z_t^*) + \phi(\ln E_t - \ln E_{t-1})$$

where e_t is the actual RER, e_t^* is the ERER, Z_t is an index of macroeconomic policies, Z_t^* is the sustainable level of macroeconomic policies, E_t is the nominal exchange rate, θ is the

[106] For example, if the aid variable cannot be satisfactorily modelled by the remaining variables in the system, the misspecification errors on the aid equation will affect the entire model.

adjustment coefficient of the self-correcting term,[107] λ reflects pressures associated with unsustainable macroeconomic policies (e.g. excess credit), and ϕ provides information about the impact of nominal devaluations. The long-run determinants ('fundamentals') of the ERER are described by:

$$\ln e_t^* = \beta_0 + \beta_1 \mathrm{lnTOT_t} + \beta_2 \mathrm{lnGCN_t} + \beta_3 \mathrm{lnCAP_t} + \beta_4 \mathrm{lnEXC_t} + \beta_5 \mathrm{lnTEC_t} + \beta_6 \mathrm{lnINV_t} + \varepsilon_t$$

where TOT is the external terms of trade, GCN government consumption of nontradables, CAP controls on capital flows, EXC index of severity of trade restrictions and exchange controls (proxied by the spread), TEC measure of technological progress, and INV ratio of investment to GDP. Finally, he defines the index of macroeconomic policies by excess supply of domestic credit (CRE) and the ratio of fiscal deficit to lagged high-powered money (DEH). Thus the typical equation to be estimated is:[108]

$$\ln e_t^* = \gamma_0 + \gamma_1 \mathrm{lnTOT_t} + \gamma_2 \mathrm{lnGCN_t} + \gamma_3 \mathrm{lnCAP_t} + \gamma_4 \mathrm{lnEXC_t} + \gamma_5 \mathrm{lnTEC_t} + \gamma_6 \mathrm{lnINV_t} + (1-\theta)\ln e_{t-1} - \lambda_1 \mathrm{CRE_t} - \lambda_2 \mathrm{DEH_t} + \phi \mathrm{DEV_t} + \varepsilon_t$$

where DEV is the nominal devaluation defined before ($\ln E_t - \ln E_t^*$), and the γ's are combinations of the β and θ. The specific variables to be included in this study, along with their expected signs, will be presented in section 5. For now, we continue with a description of alternative estimation methods.

4.4.1 Econometric Models

Engle-Granger Approach (EG)

Engle and Granger (1987) propose a two-step approach to cointegration when variables are I(1). The first step entails the estimation of the long-run (static) equation by OLS and, subsequently, testing the stationarity of the estimated residuals (ε_t) through an augmented Dickey-Fuller (ADF) test. The critical values are usually taken from MacKinnon (1996).

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$$

---

[107] Under pegged nominal rates, this adjustment takes place via changes in the price of nontradable goods.

[108] This is obtained by replacing the equation of the RER determinants and the index of macroeconomic policies into the dynamic equation.

OLS estimates of β_1 are super-consistent in the presence of cointegration,[109] even though the usual standard errors are not reliable. If the residuals are found to be non-stationary, then the variables are not cointegrated and the results obtained potentially spurious. However, if the residuals are stationary, then there is a meaningful long-run relationship between the variables. In the second step, the estimated residuals are used to estimate an error-correction model (ECM) to analyse the dynamic effects:

$$A(L)\Delta y_t = B(L)\Delta x_t + \alpha(\hat{\varepsilon}_{t-1}) + v_t$$

where A and B are polynomials in the lag operator (L), and α the speed of adjustment. However, this procedure has some well-known weaknesses. Although β is super-consistent, Banerjee et al (1986) prove that the omission of dynamic terms (in the static equation) in small samples can generate considerable bias in the estimation of β (Maddala and Kim, 1998:162). Moreover, the results may depend on the order of the variables,[110] the step-wise procedure means that potential errors in the first step are transmitted to the second (compounding), and cointegration relies on (low power) unit root tests. Finally, because the standard errors are not reliable, it is not possible to apply statistical tests on β_1 – estimates are consistent but not fully efficient.

Unrestricted Error Correction Model (ECM)

An alternative to the Engle-Granger approach is to estimate directly the ECM presented above, transforming it into a one-step procedure. The estimated error term in the EG procedure is replaced by the potential cointegrating vector.

$$A(L)\Delta y_t = B(L)\Delta x_t + \alpha(y_{t-1} - \beta_0 - \beta_1 x_{t-1}) + v_t$$

It can be shown that this is a reformulation of the autoregressive distributed lag (ARDL) model. Start with the bivariate ARDL(p,q):

$$y_t = \mu + \sum_{i=1}^{p} \delta_i y_{t-i} + \sum_{j=0}^{q} \gamma_j x_{t-j} + \varepsilon_t$$

---

[109] The estimated (cointegrating) coefficients converge quickly to their true value (at a speed T^{-1} instead of $T^{-1/2}$), allowing the dynamic terms to be ignored.

[110] Our decision about endogeneity is likely to have an impact on the results (Asteriou and Hall, 2006:317). For example, if we regress x on y, we may find a cointegrating vector that may not emerge if we regress y on x.

where μ includes deterministic components, δ_i are coefficients of the autoregressive components (i.e. lags of the endogenous variable), γ_j are coefficients for the current and lagged values of the exogenous variable. It can be easily shown that a reparametrisation[111] yields the following ECM:

$$\Delta y_t = \mu + \sum_{i=1}^{p-1} \delta_i \Delta y_{t-i} + \sum_{j=0}^{q-1} \gamma_j \Delta x_{t-j} + \alpha y_{t-1} + \omega x_{t-1} + \varepsilon_t$$

$$\alpha = -\left(1 - \sum_{i=1}^{p} \delta_i\right) \quad \text{and} \quad \omega = \sum_{i=0}^{q} \gamma_i$$

where α and ω contain information on the long-run impacts and adjustment to equilibrium. The long-run information (β_1) is recovered by dividing the estimated coefficient on x_{t-1} by the estimated coefficient on y_{t-1} (α). This approach has the advantage of not imposing restrictions on the short-run terms (therefore the name 'unrestricted' ECM), while combining information about the long-run and short-run effects in the same equation. In fact, we obtain precise estimates of the long-run coefficients and we have valid *t*-ratios for the short-run coefficients.

The ECM equation can be estimated by OLS, whilst there are two main approaches to test for cointegration: (i) an ECM or *t*-test on the coefficient of the lagged dependent variable (α), or (ii) a Wald or *F*-test for the joint significance of the long-run coefficients (α and ω). However, since these tests have a non-standard distribution, the asymptotic critical values have to be obtained by simulation. Banerjee et al (1998) propose a version of the *t*-test and report the respective critical values. Pesaran et al (2001) suggest the use of an *F*-test, which is particularly suitable when we are not certain about the order of integration of the variables. This means that we are able to combine both I(0) and I(1) variables (i.e. trend- and first-difference stationary), which eases problems related to pre-testing such as the low power of unit root tests. However, the major drawback is that the simulations only provide lower and upper bounds for the critical region. Therefore, this 'bounds testing approach' creates an inconclusive region where no inference about cointegration can be made.

[111] Replacing y_t by ($\Delta y_t + y_{t-1}$), etc. (Asteriou and Hall, 2006:311-4).

Dynamic OLS (DOLS)

The DOLS approach, originally proposed by Saikkonen (1991), improves the asymptotic efficiency of the OLS estimator for the static relation by using available stationary information. In practice, it takes into account potential endogeneity biases by adding leads and lags of the first-differenced explanatory variables to the steady-state specification:

$$y_t = \beta_0 + \sum_{i=1}^{n} \beta_i x_{i,t} + \sum_{i=1}^{n} \sum_{j=-k1}^{k2} \gamma_{i,j} \Delta x_{i,t-j} + \varepsilon_t$$

where x_i is a vector of exogenous variables, while k_1 and k_2 denote the order of leads and lags, respectively. These are often chosen by standard information criteria. The OLS estimates of β_i (long-run coefficients) are super-consistent and the respective *t*-ratios valid, although the standard errors need to be corrected for serial correlation. Stock and Watson (1993) propose a GLS correction (which they call DGLS) to obtain heteroscedasticity and autocorrelation consistent (HAC) estimators for the covariance matrix. Alternatively, we can use the Newey-West (1987) procedure (Zivot and Wang, 2006:451).[112] Cointegration can be tested through a standard ADF test on the regression residuals.

Fully-Modified OLS (FMOLS)

The unrestricted ECM method aims to obtain asymptotically efficient estimates of the cointegrating vector through the inclusion of an ECM term, whereas the DOLS approach involves adding leads and lags of the first-differenced regressors. The FMOLS approach, however, applies semi-parametric corrections to deal with endogeneity and serial correlation in the OLS estimator (Maddala and Kim, 1998:161). This method was originally suggested by Phillips and Hansen (1990) and allows for both deterministic and stochastic regressors. Although it does not require a modification of the original (static) specification, it requires two steps, which may affect its performance.[113] In fact, Pesaran and Shin (1999:403) use Monte Carlo simulations to suggest that the ARDL approach performs better than the FMOLS in small samples. The presence of a cointegrating relationship can

[112] “The Newey-West estimator provides a way to calculate consistent covariance matrices in the presence of both serial correlation and heteroscedasticity” (Johnston and DiNardo, 1997:333).

[113] The first step entails the estimation of the static long-run relation, and in the second step, we re-estimate the long-run residual covariance matrix (bias adjusted).

be evaluated through an ADF test on the regression residuals, while the Wald statistic for testing coefficient restrictions can be shown to have an asymptotic Chi-square distribution.

4.4.2 Structural Time Series Model

The main strength of time series models lies in their capacity to summarise the relevant properties of the data. In contrast to econometric models, a pure time series model ignores the role of explanatory variables and does not attempt to uncover economic behavioural relationships. Instead, the focus is on modelling the time series behaviour in terms of sophisticated extrapolation mechanisms to produce efficient forecasts (Kennedy, 2003:319). In recent times, the methodological gap between econometrics and time series analysis has been curbed by a number of factors. The finding that time series models tend to outperform forecasts produced by classic econometric models was taken as a strong indication that the latter were misspecified – they usually lacked a dynamic structure. Moreover, the increasing evidence of 'spurious regressions' in the context of non-stationary data also forced a rethink of econometric models. In practice, this led to the rise of vector autoregressive and error correction models.

Meanwhile, time series researchers were confronted with the lack of economic interpretation of their models. This led to some modelling developments, namely the combination of univariate time series analysis and econometric regressions. Two main strategies have successfully emerged: (a) mixed models, where a time series model is extended to incorporate current and/or lagged values of explanatory variables; and (b) multivariate time series models, where a set of variables is jointly analysed.

The rationale behind mixed models is that explanatory variables will only partly account for the behaviour of the variable of interest (y_t), with some degree of non-stationarity likely to remain in the system. Hence, while dynamic regression models are assumed to provide a full behavioural explanation of the process (disturbance term assumed to be stationary), a mixed model will allow a time series component to capture any left-over non-stationarity (Harvey, 1993:152-4).[114] This is particularly useful for the analysis of the long-run, where it is often difficult to find cointegration between a set of variables pro-

[114] Moreover, we avoid problems of size and power usually found in unit root and cointegration tests.

posed by economic theory. In this case, we can specify a dynamic model with both explanatory variables and a stochastic trend to fully account for the movements in y_t.

State Space Form

The state space form is often a useful way to specify a wide range of time series models. The application of the Kalman filter can then provide algorithms for smoothing and prediction, as well as a means to constructing the likelihood function (Harvey, 1993:82&181). The main concepts are now briefly explained for the univariate case, but these can be easily extended to a multivariate context. The observed variable y_t is related to the *state vector* α_t via the following *measurement equation* (Lutkepohl, 2005:611):

$$y_t = Z_t \alpha_t + \varepsilon_t$$

where Z_t is a matrix of coefficients that may depend on time, and ε_t is the observation error (usually taken as a white noise process). The elements of α_t are usually not observable, but are known to follow a first-order Markov process (Harvey, 1993:83). This can be expressed by the following *transition equation*:

$$\alpha_t = T_t \alpha_{t-1} + \eta_t$$

where T_t is a matrix of coefficients, which again can be time-dependent, and η_t is a white noise error process (uncorrelated to ε_t). A state space model will necessarily comprise both measurement and transition equations.

Unobserved Components (UC) Model

The exposition here follows Koopman et al (2007:171). The univariate structural time series model can be represented by the following measurement equation:

$$y_t = \mu_t + \psi_t + \gamma_t + \varepsilon_t$$

where y_t is the observed variable, μ_t is the trend, ψ_t the cycle, γ_t the seasonal, and ε_t the irregular component. All the components are assumed to be stochastic, but reduce to deterministic components as a limiting case. The stochastic trend is specified by the following transition equations:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \beta_{t-1} + \zeta_t$$

where β_t is the slope of the trend, η_t (level disturbance) and ζ_t (slope disturbance) are independent white noise processes, therefore uncorrelated with the irregular component. The table below presents alternative specifications of the trend.

Table 4-4: Level and Trend Specifications

Level	σ_ε	σ_η	
Constant term	*	0	
Local level	*	*	
Random walk	0	*	
Trend	σ_ε	σ_η	σ_ζ
Deterministic	*	0	0
Local level with fixed slope	*	*	0
Random walk with fixed drift	0	*	0
Local linear	*	*	*
Smooth trend	*	0	*
Second differencing	0	0	*
Hodrick-Prescott	*	0	$0.025\sigma_\varepsilon$

Obs.: '*' indicates any positive number.
Source: Koopman et al (2007, Table 9.1)

The seasonal component is specified by the trigonometric seasonal form:

$$\gamma_t = \sum_{j=1}^{[s/2]} \gamma_{j,t}$$

where each $\gamma_{j,t}$ is generated by:

$$\begin{bmatrix} \gamma_{j,t} \\ \gamma_{j,t}^* \end{bmatrix} = \begin{bmatrix} \cos\lambda_j & \sin\lambda_j \\ -\sin & \cos\lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{j,t} \\ \omega_{j,t}^* \end{bmatrix}, \qquad j = 1, \ldots, [s/2], \qquad t = 1, \ldots, T$$

where $\lambda_j = 2\pi j/s$ is the frequency in radians, and the seasonal disturbances (ω_t and ω^*_t) are mutually uncorrelated white noise processes with common variance. Finally, the cycle is specified as:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho_\psi \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \qquad t = 1, \ldots, T$$

Where ρ_ψ is a damping factor (with a range $0 < \rho_\psi \leq 1$), λ_c is the frequency in radians (with a range $0 \leq \lambda_c \leq \pi$), and the cycle disturbances (κ_t and κ^*_t) are mutually uncorrelated white noise processes with common variance.

Harvey's (1993:142) basic structural model (BSM) is often a good starting point for the analysis of time series data. The model is similar to the general univariate case specified above, except for the cycle component, which is excluded. The BSM can thus be written in the following compact form:

$$y_t = [1\ 0\ 1\ 0\ 0]\alpha_t + \varepsilon_t$$

$$\alpha_t = \begin{bmatrix} \mu_t \\ \beta_t \\ \gamma_t \\ \gamma_{t-1} \\ \gamma_{t-2} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mu_{t-1} \\ \beta_{t-1} \\ \gamma_{t-1} \\ \gamma_{t-2} \\ \gamma_{t-3} \end{bmatrix} + \begin{bmatrix} \eta_t \\ \zeta_t \\ \omega_t \\ 0 \\ 0 \end{bmatrix}$$

Explanatory Variables and Interventions

The model presented above can be extended to include current and lagged values of explanatory variable, lags of the endogenous variable, as well as intervention dummies. The model can then be written as:

$$y_t = \mu_t + \psi_t + \gamma_t + \varepsilon_t + \sum_{\tau=1}^{p} \phi_\tau y_{t-\tau} + \sum_{i=1}^{k}\sum_{\tau=0}^{q} \alpha_{i\tau} x_{i,t-\tau} + \sum_{j=1}^{h} \lambda_j \varpi_{j,t}$$

where x_{it} are exogenous variables, ϖ_{jt} are intervention dummy variables (e.g. impulse, level or slope), while ϕ_τ, $\alpha_{i\tau}$ and λ_j are unknown matrices.

This 'mixed model' is a valuable complement to traditional econometric analysis. Since explanatory variables are often not able to account for all the variation in y_t, we allow the unobserved components to capture 'left over' stochastic behaviour – trend or seasonal (Harvey, 1993:152).

4.5 Data

4.5.1 Data Construction

Nominal Exchange Rate

The Ethiopian birr was pegged to the United States dollar (USD) from its inception in 1945 until the early 1990s. The birr was valued at 2.50 per USD before the collapse of the Bretton Woods system in 1971, which forced an initial revaluation to 2.30 and then in 1973 to 2.07 per USD. The macroeconomic policies of the Derg regime (1974-1991) contributed to a significant overvaluation of the birr. In 1992, the transitional government devalued the birr to 5.00 per USD, and several (smaller) devaluations followed. A foreign exchange 'Dutch auction' system was introduced in 1993, mainly supported by foreign aid and, to some extent, by export earnings (Geda 2006:28). For two years the official and auction-based (marginal) rates co-existed in a dual exchange rate system before they were unified in 1995. In October 2001, a foreign exchange interbank market was established. The current exchange rate system is classified as a (de facto) crawling peg to the USD, i.e. a managed (or dirty) float.[115]

Table 4-5: Chronology of Main Events and Exchange Rates (Birr per USD)

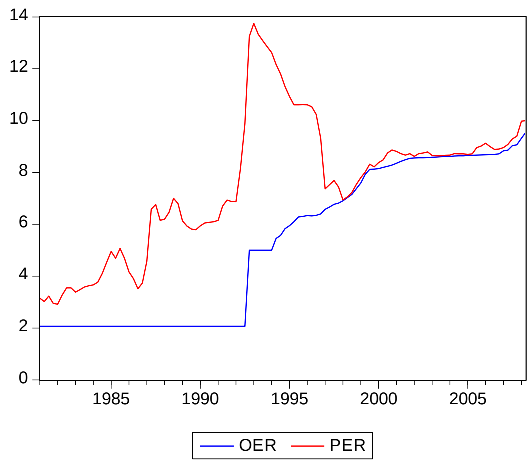
Period	Event	Official	Parallel	Premium (%)
1945:07	Currency Proclamation	2.48	n/a	n/a
1964:01	Devaluation	2.50	n/a	n/a
1971:12	Revaluation (Collapse of Gold Standard)	2.30	2.44	6.1
1973:02	Revaluation	2.07	2.16	4.3
1992:10	Devaluation	5.00	12.75	155.0
1993:05	Introduction of an Auction System (fortnight)	5.00	13.30	166.0
1995:07	Unification of the Official and Auction rates	6.25	10.55	68.8
2001:10	Introduction of an Interbank Market (daily)	8.53	8.70	2.0

Source: National Bank of Ethiopia and Pick's Currency Yearbook.
Obs.: From 1945 to 1976 the currency was called Ethiopian Dollar (E$). The Bretton Woods Agreement established 1E$ = 0.36g of fine gold.

One consequence of the gradual liberalisation of the exchange rate was the significant reduction in the parallel exchange rate premium (Degefa, 2001). The graph below plots the official and parallel exchange rates to the USD.

[115] See www.imf.org/external/np/mfd/er/2008/eng/0408.htm. This means that the central bank (actively) intervenes in the foreign exchange market (by buying or selling foreign exchange) to allow a gradual devaluation of the birr.

Figure 4-2: Official and Parallel Exchange Rates (ETB per USD): 1981-2008



Real Effective Exchange Rate Index

In this paper, the real effective exchange rate is synonymous with the multilateral RER. In the traditional sense, however, an 'effective' exchange rate is one that accounts for the structure of protection. The drawback is that data requirements for its construction are 'usually prohibitive', while there is some evidence that it may not be significantly different from the usual measure (White and Wignaraja, 1992:1472). Therefore, this strategy is not pursued. A multilateral rate (basket of foreign currencies) is usually preferred to a bilateral rate since it tends to be a better representation of overall competitiveness.

There are several alternative methods to compute the real 'effective' exchange rate index (RER). Some important decisions need to be taken, which may or may not have an influence on the analysis. Therefore, we will justify each of these decisions in light of the purpose of this study.[116] The key choices include:

(a) Nominal rate: Official or Parallel;
(b) Averaging method: Geometric vs. Arithmetic;
(c) Weights: Time-varying vs. Fixed, and Total trade vs. Export vs. Import;
(d) Trading partners: Criteria to choose which ones and how many;
(e) Price indices: CPI vs. PPI vs. WPI.

[116] Li and Rowe (2007:33-45) provide a detailed explanation of the concepts and issues involved in the construction a real exchange rate measure.

The first dilemma arises in countries with a dual exchange rate system (White and Wignaraja, 1992:1472). In the case of Ethiopia, the parallel premium is only significant until 1997, after which it tracks fairly closely the official rate. Hence, the use of the official exchange rate seems to be appropriate, although robustness tests will inform us if there is a significant departure from the long-run relationship for the period pre-1997.

The geometric average is often seen as a superior method over the arithmetic averaging procedure, mainly due to its desirable symmetry and consistency properties (Hinkle and Montiel, 1999:49-50). The arithmetic RER index is easier to compute, but it crucially depends on the choice of the base year. This is problematic to the extent that the base year is assumed to represent equilibrium in the assessment of exchange rate misalignment (Opoku-Afari, 2004). Moreover, the rate of change of the RER index (e.g. percent appreciation) will be sensitive to a shift of the reference period (re-basing). Finally, the arithmetic average gives larger weights to currencies that have significantly appreciated or depreciated in relation to the domestic currency.

The weighting scheme should reflect the relative importance of each country's foreign currency and prices in the context of Ethiopia's trade patterns (Hinkle and Montiel, 1999:97-8). Therefore, time-varying weights are often preferred to a fixed weighting scheme, since the former take into account changing trading patterns. For example, exchange rate fluctuations of the renminbi or the rupee (in relation to the birr) in the 1980s were not a significant determinant of Ethiopia's external competitiveness, given the little volume of trade with China and India during that period. Nonetheless, China and India are increasingly important trade partners and this should be reflected in the weighting scheme.[117] Fixed weights may therefore misrepresent the structure of trade in some given period. The weights can be updated every period (e.g. moving average) or we could consider period averages. Various plots of the RER index suggest that there is very little variation with regard to alternative weighting schemes. Li and Rowe (2007:39-41) reach the same conclusion for Tanzania, which suggests that the changes in weights are dominated by changes in relative prices and/or nominal exchange rates. In terms of the determination of the weights, we could use total trade shares or export/import data alone. The choice should be guided by the objective of the study. Since we are interested in both export and import competitiveness, we use total trade shares to capture overall external competitiveness.

---

[117] China and India currently account for about 25 percent of Ethiopia's trade (mostly imports).

We have tried to include a large number of trading partners in order to obtain a representative sample. Although we started with the major 26 trading partners, three countries had to be dropped due to poor quarterly price data (Djibouti, USSR/Russia, and Yemen). These 23 partners represented 73 percent of total trade flows during the period 1981-2008. However, the RER index does not seem to be sensitive to the number of trading partners selected, perhaps because most currencies co-move with hard currencies (e.g. USD, EUR, GBP, YEN, etc.). The RER index for the 6 major trading partners during the 1981-2008 period (Germany, Italy, Japan, Saudi Arabia, UK, and US – which still represent over 50 percent of trade flows) is not significantly different from the main index. The plot shows very little variation, which is corroborated by the very high correlation coefficient between the two indices.

With regard to the price indices, it is not possible to find a precise empirical equivalent to the price of tradables and non-tradables (Edwards, 1989:87). Hence, it is common to proxy (domestic) non-tradable prices by the country's consumer price index (CPI), while the producer price index (PPI) or wholesale price index (WPI) are usually used to proxy (foreign) tradable prices. The argument is that the PPI and WPI are a better representation of the price of intermediate goods, since they mainly contain tradable goods. Despite the fact that the CPI contains some tradable goods, it is greatly influenced by the non-tradable activities, thus it is a good proxy. Some authors suggest the use of GDP deflators and unit labour costs, but these were not available for most of the countries on a quarterly basis.

Edwards (1989:90&126) suggests that the weighting scheme, the choice of trading partners, and the choice of price indices does not seem to have a significant impact on the construction of the RER index. The crucial decision is between bilateral and multilateral rates, which show considerable differences in behaviour (sometimes moving in opposite directions). A multilateral real exchange rate index is preferred. Bearing in mind the discussion presented above, the RER index used in this study is computed as the geometric trade-weighted average of a basket of bilateral real exchange rates,

$$RER_t = \prod_{i=1}^{n} \left(NER_{i,t} \times \frac{P_t^d}{P_{i,t}^f} \right)^{w_{i,t}} \qquad t = 1, \ldots, T \qquad i = 1, \ldots, n$$

where *NER* is the bilateral nominal exchange rate index expressed in foreign currency per birr, while P^d and P^f are domestic and foreign price indices, respectively (proxied by the

CPI and PPI/WPI, as discussed above).[118] Both exchange rate and price indices are period averages, with base 2000=100. The subscript i identifies the trading partner, and t the time period. A total of 23 trading partners (n) were included in the construction of the REER index. Finally, w_i corresponds to the weight of each trading partner, which is allowed to vary with time (8 quarter moving average) to capture changes in trade patterns (e.g. the rise of China and India in the later part of the sample). The weights are computed as the share of each partner's trade (exports plus imports) in the total volume of Ethiopia's trade with the group of 23,

$$w_{it} = \frac{X_{it} + Z_{it}}{\sum_{j=1}^{n}(X_{jt} + Z_{jt})} \quad \text{with} \quad \sum_{i=1}^{n} w_{it} = 1 \qquad 0 < w_{it} < 1$$

Long-Run Determinants

Several studies use net capital inflows as a determinant of the RER. For example, Edwards (1989:136) uses the (lagged) ratio of net capital flows to GDP as a proxy for capital controls. However, there is a growing interest in the impact of specific capital inflows such as foreign aid, remittances and FDI. In fact, some level of disaggregation may be required, since it is unlikely that all inflows will induce the same impact, either in terms of magnitude or even sign. For example, remittances are likely to be biased towards the purchase of non-tradable (domestic) goods, while other flows may be predominantly used to purchase tradable goods.

The focus of this empirical exercise is on foreign aid inflows, both in the form of grants and concessional loans. For that purpose, most studies use net ODA flows (DAC) to proxy for foreign aid inflows. However, this paper argues that this is not a good measure for the reasons explained in previous chapters. Therefore we use data from the balance of payments. Foreign aid grants are listed as 'public transfers' in the current account, while foreign loans are a sub-item in the capital account. Due to data scarcity, remittances are proxied by 'private transfers' from the current account.[119] Data on foreign direct investment is very limited, and therefore it will not be used in this study. If the empirical results suggest that aid inflows lead to long-run RER appreciation, then concerns about Dutch disease might be justified. An insignificant coefficient would suggest that although aid flows may induce short-run adjustments, these are not permanent.

[118] In this case, an increase [fall] in the REER represents an appreciation [depreciation].
[119] This item includes 'remittances' and 'other private transfers'.

The terms of trade describes the effects of external demand and supply in the tradable goods sector (Opoku-Afari et al, 2004). The variable is usually measured by the ratio of export to import price (unit value) indices. However, trade price indices for Ethiopia are extremely difficult to obtain, even for annual data. The National Bank of Ethiopia (NBE) used to publish these until the 1980s, but then the series were discontinued. Therefore, we decided to construct the terms of trade index from secondary data. As a proxy for export prices we use the international price of coffee. Alternatively, we combine coffee and agricultural raw materials with a change in the weight structure around 2001, as coffee averaged 60 percent of total exports during 1981-2000, but only 35 percent in 2001-2008. With regard to import prices, we use the unit value of exports for two groups of countries with equal weights: advanced and developing economies. We also use changing weights to account for the steady rise of developing countries' share in total trade, from about 30 percent in 1981 to 75 percent in 2008.[120] The expected impact of the terms of trade on the RER will depend on how income and substitution effects play out. A deterioration of the terms of trade where the income effect is predominant will tend to depreciate the RER.

The degree of openness is measured by the ratio of total trade (exports plus imports) to GDP, and is used to capture the impact of trade policy (e.g. liberalisation) on the RER.[121] Alternative measures in the literature include ratios of exports to GDP and imports to GDP. In the case of Ethiopia, the ratio of total trade to GDP is more appropriate since export restrictions (e.g. taxes) were significant for most of the sample. We could also use the black market premium (BMP), i.e. the spread between the official and parallel exchange rates. However, the BMP is also likely to capture factors other than trade, exchange and capital controls (Edwards, 1989:136).[122] In terms of its expected impact, openness is likely to contribute to a depreciation of the RER. The rationale is that trade liberalisation measures (e.g. reduction in import tariffs and abolition of non-tariff barriers) will stimulate the demand for imports, leading to the depreciation of the local currency.

Government consumption of non-tradable goods is proxied by government consumption[123] as a share of GDP, since it is not possible to distinguish between expenditures on tradable and non-tradable goods. The quality of this proxy will depend on the share of non-tradable goods in total government consumption, which we think is likely to be high

---

[120] The impact of changing the weights used in the construction of the variable is negligible.
[121] Average tariff rates would constitute a better proxy, but we were not able to obtain such data.
[122] Moreover, Edwards (1989:136) suggests that the RER and BMP are jointly determined.
[123] Constructed from fiscal data as current expenditure minus food aid and interest payments.

in the case of Ethiopia. However, the inclusion of this variable in the econometric specification may cause double-counting, since aid flows are likely to finance some government consumption, such as teachers' and nurses' wages.[124] This variable is expected to be positively correlated with the RER, especially if government consumption is biased toward non-tradable goods, therefore increasing its demand.

Gross Domestic Product (GDP) is a problematic variable since the government of Ethiopia does not compile National Accounts data on a quarterly basis. Since GDP is often an important variable in this literature, we decided to construct a quarterly GDP variable by interpolating annual observations. There are several ways to do this. One is to simply distribute the yearly values across each quarter. However, this method neglects seasonal patterns and may complicate the analysis since GDP changes are equal to zero for the last three quarters. A second approach relies on a local quadratic polynomial with matched sums, i.e. the sum of the quarters equals the respective observed annual observation (QMS, 2007:108). Finally, we can use the multivariate GLS regression method proposed by Chow and Lin (1971). This method allows us to use quarterly variables that are correlated with GDP to shape the intra-year behaviour of the constructed series. Potential explanatory variables include rainfall (measured in millimetres), domestic credit, and indirect taxes (e.g. VAT). Agriculture is responsible for a substantial share of Ethiopia's output and therefore quarterly rainfall can be used as a proxy (possibly with lags). The variable is constructed as an average of three meteorological stations (Addis Ababa, Jimma and Dire Dawa). Economic activity in the remainder of the economy can be proxied by domestic credit. Indirect taxes did not turn out to be a reliable proxy, mainly due to the impact of changes in the tax regime on overall collection. Since there is a real danger of introducing spurious intra-year dynamics in the GDP ratios described above (e.g. aid-to-GDP ratio), the preferred GDP variable is computed through the second approach, which produces a series that is less volatile (reduced denominator variance) and more appropriate for the purpose of this study.

Technological progress will be approximated by a deterministic time trend, as in Bourdet and Falck (2006) and Nyoni (1998), with the objective of capturing the Ricardo-Balassa (or Balassa-Samuelson) effect. Edwards (1989:136) and other authors have proxied technological progress by the growth rate of real GDP, but our quarterly GDP growth series would be very sensitive to the interpolation approach. We expect technological

---

[124] Some authors also use total investment as a share of GDP, which is also likely to suffer from this problem.

progress to appreciate the RER if productivity increases are stronger in the tradable sector.

Short-Run Determinants

With regard to the short-run, expansionary macroeconomic policies may tend to appreciate the RER. A common proxy in the literature is excess money growth in the economy, measured by the growth of broad money (M2) minus real GDP growth. In our case, this variable (EXM2) is dominated by the behaviour of broad money. Moreover, nominal exchange rate devaluations will contribute to RER depreciation. Finally, international reserves (IRES) are likely to be negatively correlated with the RER. As opposed to the long-run determinants, these variables are assumed to only affect the RER in the short-run, and are often not included in the long-run specification.

4.5.2 Data Plots

This section provides an initial analysis of the dataset. Starting with the behaviour of the real exchange rate (LRER), the period from 1984Q2 to 1985Q4 represents the (temporary) overvaluation of the USD in relation to the other major currencies such as the German mark, British pound, and Japanese yen.[125] The quick appreciation of the LRER in early 1991 is mainly due to high inflation observed during the political transition from the Derg regime. The sharp drop in 1992Q4 is caused by the nominal devaluation of the birr in relation to the USD. The exchange rate regime was then changed from a 'hard peg' to a 'crawling peg' to the USD, and from 1993 to 2004 the LRER showed a clear depreciating trend. However, since 2005 the LRER has been appreciating, which is partly explained by the global increase in food prices, which strongly affected Ethiopia's CPI (due to the large weight in the domestic basket) but not its trading partners' WPI or PPI.[126]

[125] An increase of the RER means appreciation.
[126] This effect would only be captured by the terms of trade if the share of food imports in total imports was large.

Figure 4-3: Real Exchange Rate Index and Black Market Premium

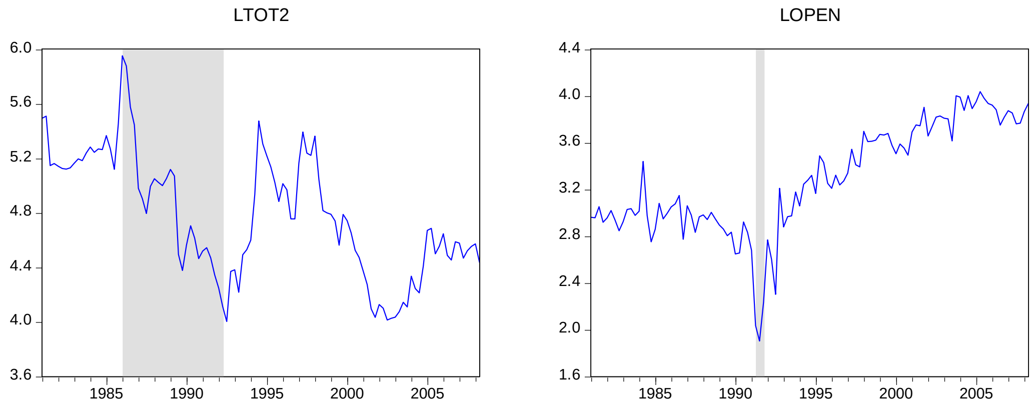
The black market premium (BMP) is the difference between the official exchange rate and the parallel market rate. The spike in the mid 1980s is mainly due to the overvaluation of the USD, the currency to which the birr was pegged. The increase in 1987 (to 1989) is likely to be associated with the macroeconomic policies of the Derg regime. The sharp increase of the premium in early 1992 is probably due to the anticipation of the devaluation of the birr. This is a usual feature of devaluations, as suggested in Edwards (1989:108). Since then, however, the premium has dropped considerably and from 1997 it is almost negligible.

The terms of trade (LTOT2) are significantly affected by the international price of coffee, which historically has revealed strong volatility.[127] Thus, the steady deterioration from 1986 to 1992 is mainly explained by the fall in coffee prices. The terms of trade observed a substantial improvement in 1993-94, before a temporary deterioration and then picking up around 1997. During the period 1998-01 there was another steady fall, followed by a strong improvement until 2005. Since then the index has been relatively steady. With regard to trade openness (LOPEN), we observe a mildly declining trend in the trade-GDP ratio up to 1990. In 1991, we observe a sharp drop that can be explained by the fall in trade volumes.[128] Since then, a steady increase ensued, as a natural consequence of the liberalisation of trade flows.

---

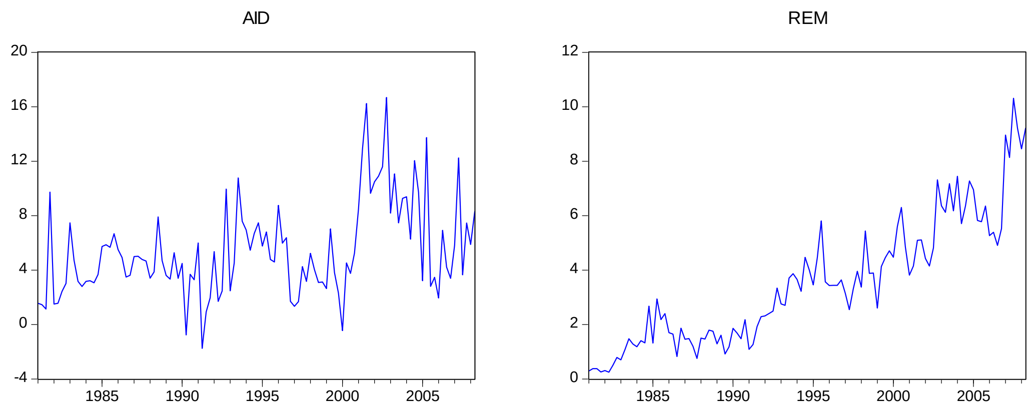
[127] The preferred specification (LTOT2) uses the international price of coffee as a proxy for export prices and export unit values from both advanced and developing economies (equal weights) as a proxy for import prices.
[128] This effect could have been compensated by a proportional fall in GDP, but in the absence of precise information the interpolation technique used does not capture it.

Figure 4-4: Terms of Trade and Openness



Foreign aid (as a share of GDP), which includes grants and concessional loans, shows considerable volatility. These have increased considerably in 2001, just after the end of the war with Eritrea, despite a drop during 2005 (election). With regard to remittances (as a share of GDP), these have been steadily increasing over the years, as a consequence of the growing Diaspora.

Figure 4-5: Foreign Aid Inflows (% GDP) and Remittances (% GDP)

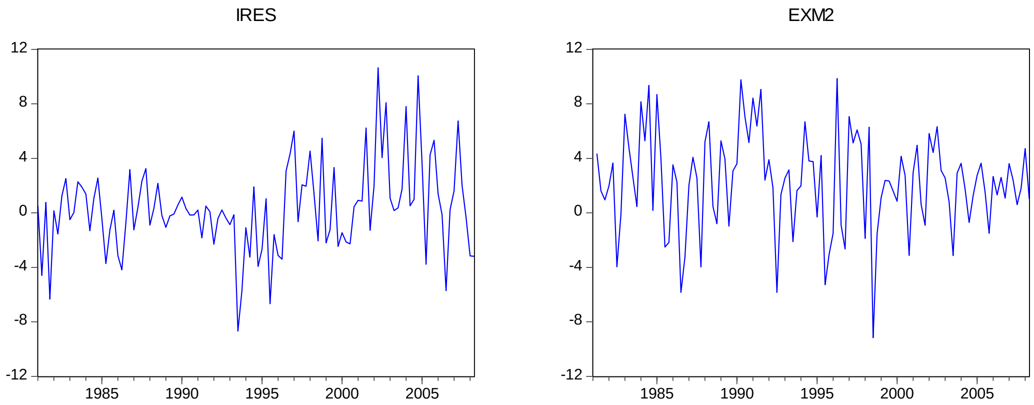


Changes in international reserves (as a share of GDP) looked fairly stable until 1993, after which they became significantly more volatile. However, we can observe than the 1981-1993 average is negative, whilst the 1994-2008 average is positive. Excess money growth shows high volatility (mainly due to M2) and appears to be stationary.[129]

---

[129] It should be noted that the volatility in these two variables (IRES and EXM2) does not come from the denominator (i.e. GDP). The interpolation technique provided a smooth estimate of quarterly GDP precisely to minimise spurious quarterly behaviour.

Figure 4-6: Change in International Reserves (% GDP) and Excess Credit Growth



4.5.3 Unit Root Tests and Seasonality

This section applies seasonal unit root tests to determine the order of integration of the variables and to assess whether there is evidence of stochastic seasonality. For that purpose we use the test proposed by Hylleberg et al (1990:216) for quarterly data. The test is based on the model:

$$\Delta_4 y_t = \pi_1 z_{1,t-1} + \pi_2 z_{2,t-1} + \pi_3 z_{3,t-1} + \pi_4 z_{3,t-2} + \sum_{j=1}^{p} \alpha_j^* \Delta_4 y_{t-j} + u_t$$

where $z_{1t} = (1 + L + L^2 + L^3)y_t$, $z_{2t} = -(1 - L + L^2 - L^3)y_t$, $z_{3t} = -(1 - L^2)y_t$ and L is the lag operator. The null hypotheses H_0: $\pi_1=0$, H_0: $\pi_2=0$ and H_0: $\pi_3=\pi_4=0$ correspond to tests for regular, semi-annual and annual unit roots, respectively. These hypotheses are tested by estimating the model above by OLS and using the relevant *t*-tests and *F*-tests. The critical values reported are from Franses and Hobijn (1997). It should be noted, however, that the asymptotic distributions of the test statistics under the respective null hypotheses depend on the deterministic terms in the model. This fact is taken into consideration since there is evidence that at least some of the series seem to be trended. The results of the unit root tests are reported below.

Table 4-6: Seasonal Unit Root Tests (Levels and First Differences)

Var.	Lag	H_0	Test	Stat	Var.	Lag	H_0	Test	Stat
LRER	0	$\pi_1=0$	$t_{\pi1}$	−2.19	DLRER	0	$\pi_1=0$	$t_{\pi1}$	−4.84***
		$\pi_2=0$	$t_{\pi2}$	−6.09***			$\pi_2=0$	$t_{\pi2}$	−5.15***
		$\pi_3=\pi_4=0$	F_{34}	49.61***			$\pi_3=\pi_4=0$	F_{34}	25.11***
LOPEN	0	$\pi_1=0$	$t_{\pi1}$	−2.06	DLOPEN	0	$\pi_1=0$	$t_{\pi1}$	−6.21***
		$\pi_2=0$	$t_{\pi2}$	−5.80***			$\pi_2=0$	$t_{\pi2}$	−4.87***
		$\pi_3=\pi_4=0$	F_{34}	23.44***			$\pi_3=\pi_4=0$	F_{34}	17.06***
LTOT2	0	$\pi_1=0$	$t_{\pi1}$	−2.76	DLTOT2	0	$\pi_1=0$	$t_{\pi1}$	−5.38***
		$\pi_2=0$	$t_{\pi2}$	−7.61***			$\pi_2=0$	$t_{\pi2}$	−6.44***
		$\pi_3=\pi_4=0$	F_{34}	32.83***			$\pi_3=\pi_4=0$	F_{34}	22.71***
AID	0	$\pi_1=0$	$t_{\pi1}$	−2.93	DAID	0	$\pi_1=0$	$t_{\pi1}$	−7.54***
		$\pi_2=0$	$t_{\pi2}$	−4.76***			$\pi_2=0$	$t_{\pi2}$	−4.72***
		$\pi_3=\pi_4=0$	F_{34}	25.73***			$\pi_3=\pi_4=0$	F_{34}	24.35***
REM	0	$\pi_1=0$	$t_{\pi1}$	−2.77	DREM	0	$\pi_1=0$	$t_{\pi1}$	−6.81***
		$\pi_2=0$	$t_{\pi2}$	−4.89***			$\pi_2=0$	$t_{\pi2}$	−4.74***
		$\pi_3=\pi_4=0$	F_{34}	36.37***			$\pi_3=\pi_4=0$	F_{34}	26.35***
EXM2	0	$\pi_1=0$	$t_{\pi1}$	−4.58***	DEXM2	0	$\pi_1=0$	$t_{\pi1}$	−7.25***
		$\pi_2=0$	$t_{\pi2}$	−4.74***			$\pi_2=0$	$t_{\pi2}$	−4.59***
		$\pi_3=\pi_4=0$	F_{34}	35.20***			$\pi_3=\pi_4=0$	F_{34}	29.13***
BMP	0	$\pi_1=0$	$t_{\pi1}$	−1.95	DBMP	0	$\pi_1=0$	$t_{\pi1}$	−4.04***
		$\pi_2=0$	$t_{\pi2}$	−6.64***			$\pi_2=0$	$t_{\pi2}$	−4.93***
		$\pi_3=\pi_4=0$	F_{34}	38.70***			$\pi_3=\pi_4=0$	F_{34}	31.68***
GEX	0	$\pi_1=0$	$t_{\pi1}$	−1.56	DGEX	0	$\pi_1=0$	$t_{\pi1}$	−5.60***
		$\pi_2=0$	$t_{\pi2}$	−3.98***			$\pi_2=0$	$t_{\pi2}$	−3.12**
		$\pi_3=\pi_4=0$	F_{34}	24.45***			$\pi_3=\pi_4=0$	F_{34}	17.48***
IRES	0	$\pi_1=0$	$t_{\pi1}$	−4.17***	DIRES	0	$\pi_1=0$	$t_{\pi1}$	8.58***
		$\pi_2=0$	$t_{\pi2}$	6.52***			$\pi_2=0$	$t_{\pi2}$	7.25***
		$\pi_3=\pi_4=0$	F_{34}	29.16***			$\pi_3=\pi_4=0$	F_{34}	30.12***

Obs.: The Schwarz Criterion was used (maximum set at 10 lags). The deterministic components included were: constant, trend, and seasonal dummies. For the differences, the trend was dropped. The results for the sub-sample 1995-2008 lead to similar conclusions, i.e. EXM2 and IRES potentially stationary. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

The null hypothesis of the HEGY test is that there is a unit root. We include a constant, a deterministic trend, and seasonal dummies in the test regression. The number of lagged seasonal differences (i.e. lag length) was selected according to the Schwarz Criterion. As expected, the results show that most variables have regular (zero frequency) unit roots (i.e. cannot reject π_1=0). The only exception is excess money growth (EXM2) and change in international reserves (IRES), which reject the unit root hypothesis. However, and perhaps more importantly, the presence of a semi-annual unit root (π_2=0) or annual unit root (π_3=π_4=0) is rejected for all variables. Finally, the HEGY test on the (first) differenced variables seems to reject the null hypotheses of unit roots.

Hence, the HEGY tests do not provide evidence of seasonal unit roots. The seasonal components do not seem to be time-dependent, suggesting that the patterns of the selected variables within the year remained relatively stable throughout the sample. The tests also suggest that most variables are integrated of order one, whilst two variables (EXM2 and IRES) appear to be stationary.

4.6 Empirical Results

4.6.1 Econometric Models

This section deals with the specification and estimation of our regression models. Taking into consideration previous theoretical and empirical studies, the initial long-run relation to be explored is:

$$LRER = \alpha_0 + \alpha_1 LOPEN + \alpha_2 LTOT2 + \alpha_3 AID + \alpha_4 REM + \alpha_5 t + \varepsilon_t$$

Where LRER is the log of the real effective exchange rate index, LOPEN the log of the ratio of total trade flows to GDP, LTOT2 is a proxy for the terms of trade, AID is the ratio of foreign aid flows to GDP, and REM the ratio of private transfers to GDP.[130] The deterministic time trend is aimed at capturing the Balassa-Samuelson effect. Moreover, seasonal dummy variables are also included in the specification to account for deterministic seasonal patterns. A number of other variables were also used: the black market premium (difference between the parallel exchange rate to the USD and the official exchange rate), the change in international reserves as percentage of GDP, the ratio of government consumption spending to GDP, and excess money growth. However, these variables are not found to be statistically significant.

The first set of results are computed for the unrestricted ECM specification – an ARDL(1,1,1,1,1). In terms of cointegration testing, the ECM-test proposed by Banerjee et al (1998) is a *t*-test on the lagged dependent variable, which in practice assesses the statistical significance of the adjustment coefficient of the error-correction term. The bounds test approach proposed by Pesaran et al (2001:307) also uses the conditional unrestricted ECM and performs an *F*-test (or Wald-test) on the long-run coefficients. The distribution of both tests is non-standard, and therefore we use the tabulated values from Pesaran et al (2001). We report results for the full sample (1981Q1-2008Q1) and for the later half (1995Q1-2008Q1). The advantage of the sub-sample is that it avoids the problematic period of the fall of the Derg regime (1991) as well as the devaluation in late 1992. It is also likely to provide better estimates of the determinants of the RER, since market forces are expected to exert stronger influence on the RER in the latter part of the sample.

[130] A similar approach is followed by Bourdet and Falck (2006).

Moreover, an alternative measure of the RER was also used as a robustness check. RERnf refers to the construction of the RER that uses the Ethiopian CPI for non-food items rather than the overall index. The difference between these indices is particularly significant in the last three years of the sample, when food prices rose sharply since 2005 as a consequence of the world food production shortfalls. Misspecification and cointegration tests are also reported.

Table 4-7: Unrestricted ECM Results

	Full Sample		Sub-Sample			
	DLRER	DLRER$_{nf}$	DLRER	DLRER	DLRER$_{nf}$	DLRER$_{nf}$
C	0.010	0.244	1.640***	1.608***	1.117*	0.990*
DLOPEN	0.004	−0.027	−0.062*	−0.064*	−0.068*	−0.055
DLTOT2	−0.026	−0.030	−0.001	0.004	0.012	0.017
DAID	−0.002*	−0.002*	−0.001	0.000	0.001	0.001
DREM	0.006	0.002	0.004	0.003	−0.004	−0.001
LOPEN(−1)	*0.043*	*0.030*	*−0.111*	*−0.113*	*−0.061*	*−0.041*
LTOT2(−1)	*−0.029*	*−0.034*	*0.069*	*0.075*	*0.026*	*0.023*
AID(−1)	*−0.004*	*−0.005*	*−0.001*		*0.000*	
REM(−1)	*0.004*	*0.002*	*0.000*		*−0.007*	
LRER$_{[nf]}$(−1)	*0.004*	*−0.024*	*−0.393*	*−0.393*	*−0.225*	*−0.205*
S1	0.021**	0.022**	0.012	0.011	0.011	0.012
S2	0.047***	0.021**	0.026**	0.025**	−0.001	0.002
S3	0.034***	0.024**	0.032**	0.031***	0.013	0.016
T	−0.001*	−0.001**	0.003***	0.003***	0.001	0.000
D924p	−0.758***	−0.723***				
D912p	0.191***	0.138***				
D911p	0.199***	0.163***				
D854p	−0.120***					
Obs	108	108	53	53	53	53
R–Squared	0.871	0.876	0.573	0.570	0.340	0.283
Serial Correlation	[0.112]	[0.102]	[0.621]	[0.622]	[0.838]	[0.363]
Functional Form	[0.416]	[0.233]	[0.468]	[0.347]	[0.001]	[0.015]
Normality	[0.201]	[0.789]	[0.778]	[0.704]	[0.974]	[0.951]
Heteroscedasticity	[0.465]	[0.458]	[0.383]	[0.273]	[0.003]	[0.003]
ECM test	0.165	−0.961	−3.771	−4.236**	−2.233	−2.276
F–test (deletion)	1.663	3.010	3.920*	6.759***	2.118	2.319

Obs.: Computed with MicroFit 4 (p-values in square brackets). The critical values for the cointegration tests are taken from Pesaran et al (2001). The significance of the long-run coefficients (*italic*) will be assessed at a later stage. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

Perhaps not surprisingly, cointegration is only found for the sub-sample (1995Q1-2008Q1). The results suggest that trade openness has a negative impact on the RER, therefore contributing to a depreciation. A positive shock to the terms of trade has a positive impact on the RER, suggesting that the income effect dominates the substitution effect. Moreover, neither foreign aid nor remittance flows seem to have a significant impact on the long-run path of the RER.

With regard to the full sample, higher values for serial correlation seem to arise from the unexplained appreciation (and then depreciation) of the birr in the mid-1980s. Since this particular behaviour of the RER was mainly a consequence of the USD movements against other hard currencies, there is little scope to explain these movements within Ethiopia. Hence, a model was also estimated for the period after 1986Q1, and although it passes all misspecification tests, there is still no cointegration amongst the variables. A number of other variables were also included in the main specification, but did not improve the results: black market premium (BMP), change in international reserves (IRES), government consumption spending (GEX), excess money growth (EXM2).

The table below presents the results from the dynamic OLS (DOLS) approach. Only the long-run coefficients are reported, since the number of first-differenced regressors in this specification is rather large and they do not have a direct economic interpretation. The leads and lags were chosen to take into consideration the frequency of the data (k_1=4 and k_2=2). Moreover, we use adjusted variance matrices to correct for potential heteroscedasticity and serial correlation. Cointegration is evaluated through an ADF test on the residuals of each regression (without seasonal dummies, trend or constant). The lag length of the ADF regression is selected by the Schwarz Information Criterion.

Table 4-8: DOLS Long-Run Estimates

	Full Sample		Sub-Sample		
	LRER	$LRER_{nf}$	LRER	LRER	$LRER_{nf}$
C	6.103***	6.615***	4.398***	4.340***	5.724***
LOPEN	−0.295***	−0.478***	−0.296*	−0.289**	−0.635***
LTOT2	0.023	0.064	0.163**	0.169***	0.152***
AID	−0.021*	0.002	−0.001		0.016***
REM	0.101**	0.026	−0.003		−0.043***
DEVAL	−0.642***	−0.556***			
S1	0.016	0.022	0.002	0.002	0.009
S2	0.035	0.023	0.005	0.025	−0.012
S3	0.042**	0.030*	0.007	0.006	0.005
T	−0.003	−0.002	0.007**	0.008***	0.007***
Obs	102	102	51	51	51
R-Squared	0.969	0.988	0.923	0.923	0.988
Serial Correlation	[0.000]	[0.000]	[0.000]†	[0.001]†	[0.003]†
Functional Form	[0.000]	[0.000]	[0.217]	[0.290]	[0.107]
Normality	[0.323]	[0.507]	[0.646]	[0.666]	[0.933]
Heteroscedasticity	[0.078]	[0.119]	[0.403]	[0.278]	[0.564]
ADF test	−3.88***	−4.41***	−5.36***	−5.45***	−6.40***

Obs.: Computed with Microfit 4 (p-values in square brackets). The *t*-statistics use the adjusted covariance matrix. †*F*-version does not reject the null hypothesis of serially uncorrelated errors. ADF tests are carried out in EViews 6 and use MacKinnon (1996) one-sided p-values. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

The ADF test results suggest that these are meaningful economic (cointegrating) relations.[131] In terms of the estimated coefficients, the results suggest that increased trade openness contributes to the depreciation of the RER, while positive terms of trade shocks tend to appreciate the RER. These results are robust to both RER measures, although LTOT2 is not statistically significant in the full sample. These findings corroborate the results obtained from the unrestricted ECM specification. With regard to the other variables, foreign aid and remittance flows do not appear to be significant long-run determinants of the RER, except for the alternative RER measure in the post-1995. However, the estimated coefficients are rather small and not robust to the RER definition. All regressions pass the normality residual test, but do not reject other misspecification problems. Nevertheless, the standard errors used provide a correction for heteroscedastic residual and serial correlation. The CUSUM tests do not suggest any structural breaks in the sub-sample.

The table below reports results from the Phillips-Hansen approach to cointegration. The Fully-Modified OLS (FMOLS) methodology uses a semi-parametric correction to solve potential endogeneity and serial correlation, therefore providing standard errors that are more reliable than standard OLS. This procedure requires all regressors to be I(1) and not cointegrated amongst themselves. Robustness checks included different weighting schemes, truncation and variables.

Table 4-9: FMOLS Long-Run Estimates

	Full Sample		Sub-Sample	
	LRER	$LRER_{nf}$	LRER	$LRER_{nf}$
C	5.719***	6.576***	4.693***	5.821***
LOPEN	−0.291***	−0.443***	−0.305***	−0.379***
LTOT2	0.080	0.049	0.117***	0.045
AID	0.005	0.005	−0.006***	0.002
REM	0.068***	0.023*	0.008*	−0.008
DEVAL	−0.597***	−0.480***		
S1	0.031	0.032	0.000	0.021
S2	0.084**	0.071**	0.058***	0.025
S3	0.055	0.043	0.042***	0.006
T	−0.002	−0.004**	0.006***	0.000
Obs	108	108	53	53
ADF test	−3.49**	−5.23***	−5.22***	−3.94**

Obs.: Computed with MicroFit 4 (Bartlett weights with truncation lag 4). The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

131 It should be noted that residual-based cointegration tests are known to have low power (Type II error), especially in relatively small samples. This means that they may fail to reject the null hypothesis in the presence of cointegration.

The results reported above provide further evidence that trade openness depresses the RER, while the terms of trade are associated with a RER appreciation. The impact of the remaining variables is rather small or even statistically insignificant. Foreign aid is significant in the sub-sample, but not with the expected theoretical sign. Moreover, remittances appear to appreciate the RER, but the impact is not robust to the choice of the RER measure. The plot of the residuals does not suggest misspecification problems.

We now summarise the results obtained so far by focusing on the sub-sample, since these do not raise misspecification concerns and seem more reliable. The table below shows the long-run coefficient from the three cointegration approaches. It also reports their statistical significance through Wald tests, as proposed by Stock and Watson (1993) for the DOLS approach and Phillips and Hansen (1990) for the FMOLS method. The results corroborate the conclusions from the *t*-stats.

Table 4-10: Summary of RER Long-Run Determinants

	ECM	DOLS	FMOLS
LOPEN	-0.29^{***}	-0.29^{***}	-0.30^{***}
LTOT2	0.19^{***}	0.17^{***}	0.12^{***}
AID	0.00	0.00	-0.01^{***}
REM	0.00	0.00	0.01^{*}

Obs.: Computed with MicroFit 4. The results for DOLS are based on adjusted covariance matrix. Wald-test statistic is distributed as a Chi-Square(1). Disaggregating aid inflows into grants and loans does not change the conclusions. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

Overall, the three approaches to cointegration used in this paper provide strong evidence that trade openness and the external terms of trade have a significant impact on the long-run path of the RER. The former always entails a depreciation of the RER, as predicted by economic theory, whilst the latter tends to appreciate the RER, suggesting that the income effect is stronger than the substitution effect. With regard to capital inflows, there is little evidence of a significant impact of foreign aid or remittance inflows on the RER. Hence, the Dutch disease hypothesis does not seem to hold for Ethiopia. Other variables were also included in the specification, but little was gained in terms of explanatory power (i.e. relevance to either long-run or short-run) or improvement of misspecification tests. These included excess money growth (EXM2), black market premium (BMP), change in international reserves (IRES), government consumption expenditure (GEX), capital account balance (KA), and net errors and omissions from the BoP (NEO). The time trend was statistically significant in some specifications, possibly capturing the Balassa-Samuelson effect.

4.6.2 Unobserved Components

We now turn to a structural time series model to complement the empirical exercise undertaken above. The starting point is Harvey's basic structural model (BSM):

$$y_t = \mu_t + \gamma_t + \varepsilon_t$$
$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \beta_{t-1} + \zeta_t$$

where y_t is the observed variable (LRER), μ_t is the trend, γ_t the seasonal, and ε_t the irregular component. Note that the components are initially assumed to be stochastic, whilst the cycle is excluded from the specification. The BSM seems to be a good starting point for the empirical analysis since the economic theory on RER determination does not provide a strong argument for the presence of cycles. Moreover, the validity of this assumption can be analysed through spectral analysis. The initial estimation results suggest that the seasonal component is deterministic for both sample sizes, since the estimated variance of the component is not statistically significant. This corroborates the results from the HEGY test, and further validates the use of dummy variables to account for seasonality in the econometric (cointegration) models. Moreover, the variance of the irregular component is not statistically significant, indicating that the movements of the RER are totally explained by a stochastic trend and a deterministic seasonal component. The stochastic trend is then re-specified as a 'smooth trend' by setting the level variance to zero (σ_η = 0) whilst letting the slope variance unrestricted. Given the insights of the initial univariate model, we will now estimate and present the results of a mixed model, where we add potential explanatory variables to the structural model. The final specification is:

$$y_t = \mu_t + \gamma_t + \varepsilon_t + \sum_{i=1}^{4}\sum_{\tau=0}^{1} \alpha_{i\tau} x_{i,t-\tau} + \lambda DUM924s$$
$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \beta_{t-1} + \zeta_t$$

where x_i includes openness (LOPEN), terms of trade (LTOT2), foreign aid (AID) and remittances (REM). Both contemporaneous and one period lags are included for all explanatory variables. A level shift is also included to account for the devaluation in the last quarter of 1992. Since the results for the full sample do not suggest that any of the variables is statistically significant, we focus on the sub-sample. The sub-sample includes 53 observations and the order of trend smoothness value (p) is 3. The summary statistics suggest that the model passes the normality test, which is the Bowman-Shenton statistic based on the third and fourth moments of the residuals. The heteroscedasticity test (H) is the ratio of the squares of the last h residuals to the squares of the first h residuals (h is set at the closet integer of T/3) and it is centred around unity. Serial correlation is assessed through the Durbin-Watson test, serial correlation coefficients (r) at the first and last lags, and the (portmanteau) Box-Ljung statistic (Q) based on the first p autocorrelations. The results suggest only mild autocorrelation. Finally, the coefficient of determination based on the differences around seasonal means is 0.28, whilst the more common measure (R^2) is 0.89. In terms of the component's variances, we confirm that the level variance is set to zero (smooth trend), while the seasonal variance is estimated to be zero.

Table 4-11: Summary Statistics and Disturbances (Sub-Sample)

Summary Statistics		5% critical value [p-value]
T	53	
p	3	
std. error	0.027	
Normality	0.529	$\chi^2(2)$=5.99 [0.768]
H(13)	1.374	F(h,h)=2.48 [0.288]
DW	1.681	N(2,4/T)=1.548 [0.123]
r(1)	0.150	N(0,1/T)=−0.226 [0.863]
q	8	
r(q)	−0.305	N(0,1/T)=−0.226 [0.013]
Q(q,q−p)	11.578	$\chi^2(5)$=11.07 [0.041]
Rs^2	0.280	

Disturbances	Variance	q-ratio
Level	0	0.000
Slope	4.87E-05	0.113
Seasonal	0	0.000
Irregular	0.000431	1.000

Obs.: Allowing a stochastic level does not change results.

The next table shows the values of the state vector and regression effects. The coefficients of the explanatory values are interpreted as in classic econometric models, and support the conclusions from the cointegration analysis. The lag of trade openness is negatively correlated with the RER, while lagged terms of trade induce a RER appreciation. The fact that the one-period lags are statistically significant and not the contemporaneous values may suggest that the transmission mechanisms take a certain time to affect the RER. Neither foreign aid nor remittance flows are statistically significant, although the coefficient for workers' remittances is not far from significance. In fact, the suggestion that remittances may appreciate the RER supports the weak evidence from the FMOLS.

Table 4-12: State Vector Analysis and Regression Effects at period 2008(1)

State Vector	Value	Prob.	Regressors	Coefficient	RMSE	*t*-value	Prob.
Level	4.846	0.000	LOPEN	−0.048	0.038	−1.255	0.217
Slope	0.006	0.587	LOPEN(−1)	−0.078*	0.040	−1.973	0.055
Seasonal (χ^2 test)	8.930	0.030	LTOT2	−0.002	0.037	−0.061	0.952
S1	−0.011	0.073	LTOT2(−1)	0.074**	0.036	2.084	0.044
S2	0.000	0.995	AID	−0.001	0.001	−0.413	0.682
S3	0.018	0.006	AID(−1)	−0.001	0.001	−0.655	0.516
S4	−0.007	0.233	REM	0.007	0.004	1.682	0.100
			REM(−1)	0.000	0.004	−0.108	0.914

Obs.: Disaggregating aid inflows into grants and loans does not change the conclusions. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

We now undertake a graphical analysis of the residuals, in order to assess the robustness and reliability of the results. The first graph presents the standardised residuals, which do not suggest the presence of a significant outlier. Moreover, the histogram supports the normal distribution, while the CUSUM *t*-test confirms the stability of the model. The correlogram does not show significant autocorrelations, except perhaps for lag 8.

Figure 4-7: Standardised Residuals

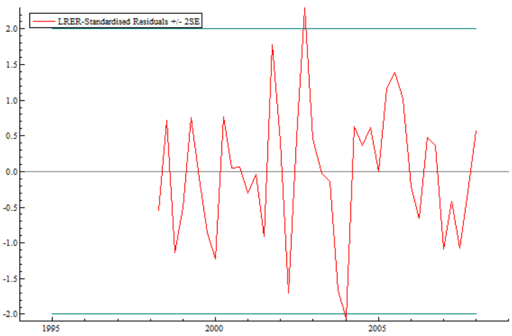


Figure 4-8: Histogram

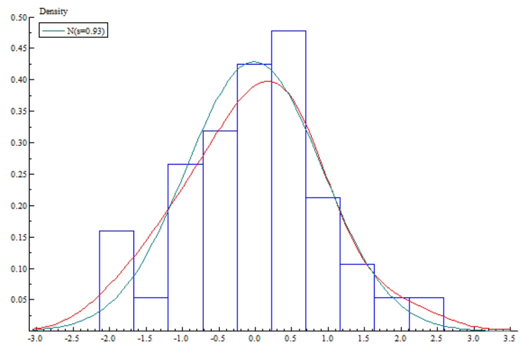


Figure 4-9: CUSUM t-Test

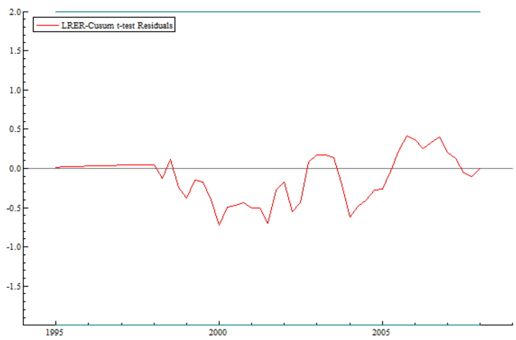
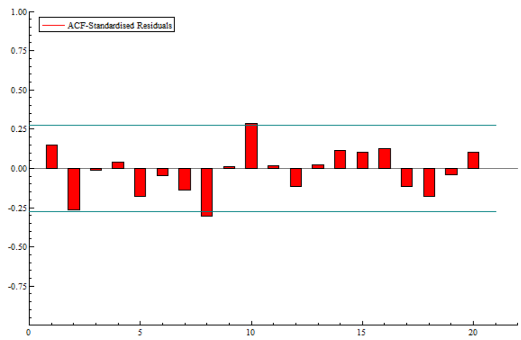
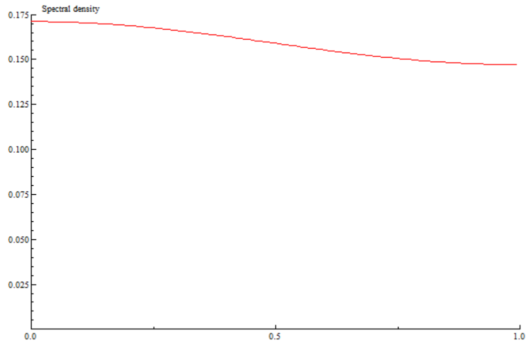


Figure 4-10: Correlogram



Finally, the spectral density is clearly flat, therefore corroborating the decision to exclude of the cycle component from the model. A marked peak in the graph would have suggested the presence of a cycle.

Figure 4-11: Spectral Density



In summary, the results from the unobserved components model suggest that the liberalisation of trade flows in Ethiopia (openness) have contributed to the depreciation of the RER, while positive terms of trade shocks contribute to appreciation pressures. Foreign aid inflows are not found to have a significant impact on the RER, while remittances may induce small appreciations.

4.7 Conclusions

This paper aimed to uncover the main determinants of RER fluctuations in Ethiopia. In particular, it tried to assess whether large capital inflows – mainly foreign aid and remittances – tend to cause the RER to appreciate as suggested by the theoretical literature. The classic Dutch disease model (Corden and Neary, 1982) was briefly reviewed, while the equilibrium RER approach (Edwards, 1989) underpinned the empirical exercise. The importance of real determinants ('fundamentals') as well as nominal (short-run) determinants was scrutinised. Despite the unequivocal theoretical prediction that large capital inflows will force the RER to appreciate, empirical studies have seldom found robust evidence of Dutch disease effects, especially with regard to foreign aid inflows. Hence, this paper reviewed some alternative arguments and proceeded to conduct a thorough empirical investigation of the RER dynamics in Ethiopia for the period 1995-2008.

Three cointegration frameworks were used: (i) the unrestricted error correction model proposed by Banerjee et al (1998); (ii) the dynamic OLS approach suggested by Saikkonen (1991) and Stock and Watson (1993); and (ii) the fully-modified OLS estimator of Phillips and Hansen (1990). Moreover, the unobserved components (UC) model (Harvey, 1992) provided an alternative empirical framework. Its main strength lies in its explicit modelling of the stochastic trend, and by extending the UC model to include explanatory variables we are able to evaluate whether capital inflows (as well as other 'fundamentals') have had a significant impact on RER determination.

The results emerging from both the econometric and time series models suggest the following conclusions. The openness measure has a negative impact on the long-run value of the RER, which means that reforms undertaken in the 1990s to liberalise trade flows and exchange markets have contributed to downward pressures on the RER (depreciation). Moreover, external terms of trade shocks have a positive impact on the RER (appreciation). The implication of this finding is that the income effect outweighs any potential substitution effects. With regard to capital inflows, neither foreign aid nor workers' remittances were found to be statistically significant, although the latter might be weakly associated with appreciation pressures. This may be due to the fact that a large share of private transfers is spent on domestic products. The lack of robust evidence that capital inflows appreciate the RER questions the Dutch disease thesis, but is not totally surprising.

> "Time-series models […] tend to find that it [RER] responds much less to variations in aid flows than it does to other exogenous foreign exchange flows, most notably commodity price or terms of trade variations." (Adam 2006:178)

Hence, this paper argues that the main fluctuations of the Ethiopian RER can be accounted by three main factors: (i) external commodity price shocks (measured by the terms of trade), mainly affecting coffee exports and oil imports; (ii) political events, such as the instability towards the end of the Derg regime that caused unusually high inflation; and (iii) economic policy, especially the liberalisation of trade flows and the exchange rate market. These effects are not likely to act through the capital account, since there are still several restrictions in Ethiopia. Moreover, the lack of significance of variables such as excess money supply and government consumption may suggest that Ethiopia has pursued sound macroeconomic policies since the fall of the Derg regime. To conclude, the results suggest that Ethiopia has been able to effectively manage large capital inflows, thus avoiding major episodes of macroeconomic instability. A prudent approach from the central bank and aid flows targeted at alleviating supply-side constraints (mostly through public investment) may have played an important role.

4.8 Bibliography

Adam, C. (2006), "Exogenous Inflows and Real Exchange Rates: Theoretical Quirk or Empirical Reality?" in P. Isard, L. Lipschitz, A. Mourmouras, and B. Yontcheva (eds) The Macroeconomic Management of Foreign Aid: Opportunities and Pitfalls (Washington: International Monetary Fund, 2006).

Adam C. and D. Bevan (2006), "Aid and the Supply Side: Public Investment, Export Performance, and Dutch Disease in Low-Income Countries." *The World Bank Economic Review*, 20(2), 261-290.

Adenauer I. and L. Vagassky (1998), "Aid and the Real Exchange Rate: Dutch Disease Effects in African Countries." *Intereconomics: Review of International Trade and Development*, 33(4), 177-185.

Asteriou, D. and S. Hall (2007), Applied Econometrics: A Modern Approach Using EViews and Microfit (Revised Edition), Palgrave MacMillan.

Bandara, J. (1991), "An Investigation of 'Dutch disease' Economics with a Miniature CGE Model." *Journal of Policy Modeling*, 13(1), 67–92.

Banerjee, A., J. Dolado and R. Mestre (1998), "Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework." *Journal of Time Series Analysis*, 19(3), 267-283.

Banerjee, A., J Dolado, D. Hendry and G. Smith (1986), "Exploring Equilibrium Relationships in Econometrics through Static Models: Some Monte Carlo Evidence." *Oxford Bulletin of Economics and Statistics*, 48(3), 253-277.

Benjamin, N. (1990), "Investment, the Real Exchange Rate, and Dutch Disease: A Two-Period General Equilibrium Model of Cameroon." *Journal of Policy Modeling*, 12(1), 77-92.

Bourdet, Y. and H. Falck (2006), "Emigrants' Remittances and Dutch Disease in Cape Verde." *International Economic Journal*, 20(3), 267-84.

Buffie, E., C. Adam, S. O'Connell and C. Pattillo (2008), "Riding the Wave: Monetary Responses to Aid Surges in Low-Income Countries." *European Economic Review*, 52(8), 1378–1395.

Buffie, E., C. Adam, S. O'Connell, and C. Pattillo (2004), "Exchange Rate Policy and the Management of Official and Private Capital Flows in Africa." International Monetary Fund, Staff Papers, 51(Special Issue), 126-60.

Cerra, V., S. Tekin and S. Turnovsky (2008), "Foreign Aid and Real Exchange Rate Adjustments in a Financially Constrained Dependent Economy." International Monetary Fund, IMF Working Paper 08/204.

Chow, G. and A. Lin (1971), "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series." *Review of Economics and Statistics*, 53(4), 372-375.

Collier, P. and J. Gunning (1992), "Aid and Exchange Rate Adjustment in African Trade Liberalisations." *The Economic Journal*, 102(413), 925-39

Corden, M. (1984). "Boom Sector and Dutch Disease Economics: Survey and Consolidation." *Oxford Economic Papers*, 36(3), 359-380.

Corden, M. and P. Neary (1982), "Booming Sector and De-Industrialisation in a Small Open Economy." *The Economic Journal*, 92 (368), 825-48.

Davidson, R. and J. MacKinnon (1993), Estimation and Inference in Econometrics, Oxford University Press, New York.

Degefa, D. (2001), "The Parallel Foreign Exchange Market and Macroeconomic Performance in Ethiopia." AERC Research Paper 107.

Devarajan, S., D. Go, J. Page, S. Robinson and K. Thierfelder (2008), "Aid, Growth, and Real Exchange Rate Dynamics." The World Bank, Policy Research Working Paper 4480

Ebrahim-zadeh, C. (2003), "Back to Basics. Dutch Disease: Too Much Wealth Managed Unwisely." International Monetary Fund, Finance & Development, 40(1).

Edwards, S. (1989), Real Exchange Rates, Devaluation and Adjustment: Exchange Rate Policy in Developing Countries, MIT Press, Cambridge, Massachusetts.

Elbadawi, I. (1999), "External Aid: Help or Hindrance to Export Orientation in Africa?" *Journal of African Economies*, 8(3), 578–616.

Engle, R. and C. Granger (1987), "Co-Integration and Error Correction: Representation, Estimation, and Testing." Econometrica, 55(2), 251-276.

Franses, P and B. Hobijn (1997), "Critical Values for Unit Root Tests in Seasonal Time Series." *Journal of Applied Statistics*, 24(1), 25-47.

Geda, A. (2006), "The Structure and Performance of Ethiopia's Financial Sector in the Pre- and Post-Reform Period with a Special Focus on Banking." WIDER Research Paper 2006/112.

Granger, C. and P. Newbold, (1974), "Spurious Regressions in Econometrics." *Journal of Econometrics*, 2(2), 111–120

Greene, W. (2003), Econometric Analysis (Fifth Edition), Prentice Hall, New Jersey.

Hallwood, P. and R. MacDonald (2000), International Money and Finance (Third Edition), Blackwell.

Harvey, A. (1993), Time Series Models (Second Edition), MIT Press, New York.

Hinkle, L. and P. Montiel, eds. (1999), Exchange Rate Misalignment: Concepts and Measurement for Developing Countries, World Bank, Oxford Press University.

Hylleberg, S., R. Engle, C. Granger and B. Yoo (1990), "Seasonal Integration and Cointegration." *Journal of Econometrics*, 44(1-2), 215-238.

Issa, H and B. Ouattara (2008), "Foreign Aid Flows and Real Exchange Rate: Evidence from Syria" *Journal of Economic Development*, 33(1), 133-46

Johnston, J. and J. DiNardo (1997), Econometric Methods (Fourth Edition), McGraw-Hill.

Kennedy, P (2003), A Guide to Econometrics (Fifth Edition), John Wiley & Sons

Koopman, S., A. Harvey, J. Doornik and N. Shephard (2007), STAMP 8: Structural Time Series Analyser and Modeller and Predictor, Timberlake Consultants Press, London.

Laplagne, P., M Treadgold and J. Baldry (2001), "A Model of Aid Impact on some South Pacific Microstates." *World Development*, 29(2), 365-383.

Lartey, E. (2007), "Capital Inflows and the Real Exchange Rate: An Empirical Study of Sub-Saharan Africa." *Journal of International Trade and Economic Development*, 16(3), 337-57.

Li, Y. and F. Rowe (2007), "Aid Inflows and the Real Effective Exchange Rate in Tanzania." World Bank Policy Research Working Paper 4456.

Lütkepohl, H. (2005), New Introduction to Multiple Time Series Analysis, Springer.

MacKinnon, J. (1996). "Numerical Distribution Functions for Unit Root and Cointegration Tests." *Journal of Applied Econometrics*, 11(6), 601-618.

Maddala, G. and I. Kim (1998), Unit roots, Cointegration, and Structural Change, Cambridge University Press.

Michaely, M. (1981), "Foreign Aid, Economic Structure, and Dependence." *Journal of Development Economics*, 9(3), 313-330.

Mongardini, J. and B. Rayner (2009), 'Grants, Remittances, and the Equilibrium Real Exchange Rate in Sub-Saharan Africa Countries.' International Monetary Fund, IMF Working Paper 09/75.

Newey, W. and K. West (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55(3), 703-708.

Nkusu, M. (2004), "Aid and the Dutch Disease in Low-Income Countries: Informed Diagnoses for Prudent Prognoses." International Monetary Fund, Working Paper 04/49.

Nyoni, T. (1998), "Foreign Aid and Economic Performance in Tanzania." *World Development*, 26(7), 1235-40.

Ogun, O. (1998), "Real Exchange Rate Movements and Export Growth: Nigeria, 1960-1990." AERC Research Paper 82.

Opoku-Afari, M. (2004), "Measuring the Real Effective Exchange Rate in Ghana." CREDIT Research Paper 04/11.

Opoku-Afari, M, O. Morrissey and T. Lloyd (2004), "Real Exchange Rate Response to Capital Inflows: A Dynamic Analysis for Ghana" CREDIT Research Paper 04/12.

Ouattara, B. and E. Strobl (2004), "Do Aid Inflows Cause Dutch Disease? A Case Study of the CFA Franc Countries." University of Manchester, School of Economic Studies, Discussion Paper 0330.

Pesaran, H., Y. Shin and R. Smith (2001), "Bounds Testing Approaches to the Analysis of Level Relationships." *Journal of Applied Econometrics*, 16(3), 289-326.

Pesaran, H. and B. Pesaran (1997), Working with Microfit 4.0: Interactive Econometric Analysis. Oxford University Press, London.

Pesaran, H. and Y. Shin (1999), "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis." Econometrics and Economic Theory in 20th Century: the Ragnar Frisch Centennial Symposium, eds. S. Strom, Cambridge University Press, Cambridge.

Phillips, P. and B. Hansen (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes." *The Review of Economic Studies*, 57(1), 99-125.

Prati, A. and T. Tressel (2006), "Aid Volatility and Dutch Disease: Is There a Role for Macroeconomic Policies?" International Monetary Fund, Working Paper 06/145.

Prati, A., R. Sahay and T. Tressel (2003), "Is There a Case for Sterilizing Foreign Aid Inflows?" Paper presented at the Research Workshop "Macroeconomic Challenges in Low Income Countries", October 23-24, Washington D.C.: IMF.

QMS (2007), EViews 6 User's Guide I, Quantitative Micro Software, LLC.

Rajan, R. and A. Subramanian (2005), "What Undermines Aid's Impact on Growth?" International Monetary Fund, Working Paper 05/126.

Rogoff, K. (1996), "The Purchasing Power Parity Puzzle." *Journal of Economic Literature*, 34(2), 647-668.

Sackey, H. (2001), "External Aid Inflows and the Real Exchange Rate in Ghana." AERC Research Paper 110.

Saikkonen, P. (1991), "Asymptotically Efficient Estimation of Cointegration Regressions." *Econometric Theory*, 7(1), 1-21.

Stock, J. and M. Watson (1993), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems." *Econometrica*, 61(4), 783-820.

Torvik, R. (2001), "Learning by Doing and the Dutch Disease." *European Economic Review*, 45(2), 285-306.

Van Wijnbergen, S. (1986), "Macroeconomic Aspects of the Effectiveness of Foreign Aid: On the Two-Gap Model, Home Goods Disequilibrium and Real Exchange Rate Misalignment." *Journal of International Economics*, 21(1-2), 123-36.

Van Wijnbergen, S. (1984), "The 'Dutch Disease': A Disease After All?" *The Economic Journal*, 94(373), 41-55.

Vos, R. (1998), "Aid Flows and 'Dutch Disease' in a General Equilibrium Framework for Pakistan." *Journal of Policy Modeling*, 20(1), 77-109.

White, H. and G. Wignaraja (1992), "Exchange Rates, Trade Liberalization and Aid: The Sri Lankan Experience." *World Development*, 20(10), 1471-80.

Wooldridge, J. (2003), Introductory Econometrics: A Modern Approach (Second Edition), South-Western.

Yano, M. and J. Nugent (1999), "Aid, Nontraded Goods, and the Transfer Paradox in Small Countries." *American Economic Review*, 89(3), 431-449.

Younger, S. (1992), "Aid and Dutch Disease: Macroeconomic Management When. Everybody Loves You." *World Development*, 20(11), 1587-97.

Zivot, E. and J. Wang (2006), Modeling Financial Time Series with S-PLUS (Second Edition), Springer-Verlag.

5. Aid Absorption and Spending in Africa: A Panel Cointegration Approach

Abstract: This paper focuses on the macroeconomic management of large inflows of foreign aid. It investigates the extent to which African countries have coordinated fiscal and macroeconomic responses to aid surges. In practice, we construct a panel dataset to investigate the level of aid 'absorption' and 'spending'. This paper departs from the recent empirical literature by utilising better measures for aid inflows and by employing cointegration analysis. The empirical short-run results suggest that, on average, Africa's low-income countries have absorbed two-thirds of (grant) aid receipts. This suggests that most of the foreign exchange provided by the aid inflows has been used to finance imports. The other third has been used to build up international reserves, perhaps to protect economies from future external shocks. In the long-run, absorption increases but remains below its maximum ('full absorption'). Moreover, we also show that aid resources have been fully spent, especially in support of public investment. There is only weak evidence that a share of aid flows have been 'saved', i.e. substituted domestic borrowing. Overall, these findings suggest that the macroeconomic management of aid inflows in Africa has been significantly better than often portrayed in comparable exercises. The implication is that African countries will be able to efficiently manage a gradual scaling up in aid resources.

5.1 Introduction

Foreign aid is often provided with the twin objectives of financing domestic expenditures and increasing the availability of foreign exchange. In Africa's low-income countries, external grants and concessional loans provide crucial resources to support the expansion of public investment programmes – e.g. building important socio-economic infrastructure that contributes to fostering economic growth and alleviating poverty. Moreover, these flows provide foreign exchange resources that allow countries to increase imports of capital goods, which stimulate economic output and are often associated with productivity gains.

This paper is mainly concerned with the fiscal and macroeconomic management challenges arising from large foreign aid inflows. For that purpose, we use the analytical framework proposed by the IMF (2005) and Hussain et al (2009) to investigate whether

African countries have pursued a coordinated strategy in terms of their fiscal and macroeconomic responses to large aid inflows. The lack of coordination between the government and the central bank may undermine the effective use of foreign aid resources, often contributing to inflationary pressures, the appreciation of the nominal exchange rate, high interest rates and accumulation of public debt (Buffie et al, 2004).

We construct a new panel dataset for African countries, covering the period 1980-2005. An important emphasis is placed on the definition, source and construction of the main variables. Although the vast literature on the macroeconomic impacts of foreign aid inflows predominantly uses OECD-DAC data on aid, we argue that this is not appropriate. One reason is that donor-reported statistics often overestimate the 'true' amount of aid. For example, costs relating to technical assistance are included in foreign aid statistics (e.g. OECD-DAC's) even though many of these payments never actually leave the donor country's banking system. Since these activities have no clear impact on the balance of payments or the fiscal budget, they should not be included in the analysis. Moreover, off-budgets are not likely to have significant fiscal effects. Therefore, we favour the use of official data from recipient countries to assess the questions at hand. In this study we use balance of payments (BOP) data for the macroeconomic variables (including external grants) and government data for the fiscal variables. The former is reported in the IMF's Balance of Payments Statistics (BOPS) by the respective central banks, whilst the latter is reported in the World Bank's Africa Database by World Bank country economists. This actually entails the construction of two different measures of foreign aid.

This paper also strives to use appropriate panel data methodologies. Despite the popularity of dynamic panel data (DPD) methods in applied research, these seem to be more suitable for panels with large N (e.g. countries) and small T (observations through time). For panels that incorporate both a significant number of cross-sections and annual observations – like this one – non-stationarity becomes a major concern for inference. Therefore, we use recently developed methods that have strong foundations in the analysis of time series data, namely, panel unit root tests, cointegration tests, and efficient estimators for assessing long-run relationships.

The next section provides a brief overview of the literature on the macroeconomic effects of aid. Moreover, it introduces the analytical framework that provides the background for this study and presents the few existing empirical results. Section 3 introduces the empirical methodologies to be utilised in this study. Section 4 explains the construction of the variables, whereas section 5 presents the empirical findings. Section 6 concludes the paper.

5.2 Literature Review

The Macroeconomic Management of Aid

There is a growing literature on the macroeconomic challenges associated with large foreign aid inflows. White (1992) is an important and often cited early contribution. The author critically surveys the debates relating to the impact of aid on domestic savings, the fiscal response, the real exchange rate and ultimately economic growth, thus providing an excellent synthesis of the theoretical and empirical contributions to the topic. However, academic interest in these lines of investigation may have suffered from the marked reduction in aid flows to developing countries during the 1990s. This declining trend was partly due to: (i) the collapse of the Soviet Union, eliminating the geo-political justification for providing aid inflows; (ii) rising concerns about the effectiveness of aid in achieving desired outcomes, namely policy reform, economic growth and poverty reduction ('aid fatigue'); and (iii) the economic recession that affected several donors in the early 1990s. Nonetheless, the early 2000s witnessed a renewed interest from the international donor community. The United Nations Millennium Declaration (and the subsequent Millennium Development Goals) provided the impetus that was quickly followed by promises to increase the availability of external finance to developing countries – in particular to Africa.[132] Naturally, this led to the revival of many debates concerning the impact of 'scaling up' aid inflows. The International Monetary Fund took a decisive lead, with publications such as Isard et al (2006), Heller (2005) and Gupta et al (2006). These works revisit the main foreign aid debates and provide an overview of current knowledge.

We can subdivide the main issues concerning the macroeconomics of aid into two main areas: (i) the fiscal sphere, which is influenced by recipient governments; and (ii) the monetary and exchange rate sphere, which is usually under the responsibility of central banks. The first incorporates questions about the impact of aid on the size and composition of public spending, domestic revenues, fiscal deficit, debt sustainability and aid dependency. This leads to policy decisions such as how much aid the government should spend and whether it should save some of the aid resources (e.g. to smooth the expendi-

[132] These were embedded in the 2002 Monterrey Consensus – an outcome of the United Nations International Conference on Financing for Development – and the 2005 Gleneagles G8 summit.

ture pattern when resources are scarce). The second area focuses on concerns of exchange rate appreciation, rising price inflation and high interest rates. This often leads to debates about the optimal level of sterilisation (e.g. Prati et al, 2003) and effective exchange rate regimes (e.g. Buffie et al, 2004). Nonetheless, these two areas of interest are interdependent and should be considered in tandem. Fiscal decisions crucially depend on macroeconomic circumstances (e.g. the interest rate on domestic public debt), while central bank objectives (e.g. low inflation) are partly influenced by the government's policy stance. This interdependence has led to the development of the analytical framework that we will now discuss.

Analytical Framework

The starting point of this empirical investigation is the analytical framework proposed by Hussain et al (2009).[133] The framework is used to investigate the macroeconomic management challenges and optimal policy responses to increases (surges) in foreign aid inflows. This is a crucial policy issue for low-income countries, which are often aid-dependent and may suffer from the volatility and unpredictability of aid flows. Hence, the framework emphasises the need to coordinate fiscal policy with monetary and exchange rate policy in order to minimise potential adverse effects and improve its efficiency. Hussain et al (2009) suggest the use of the following two interrelated concepts: (i) 'absorption', which is defined as the widening of the current account deficit (excluding aid) due to the aid surge; and (ii) 'spending', which is defined as the widening of the fiscal deficit (excluding aid) following an aid surge. Absorption can be seen as a measure of the degree of 'real resource transfer,'[134] whilst spending assesses "the extent to which the government uses aid to finance an increase in expenditures or a reduction in taxation" (Gupta et al, 2006:10). In the special cases of aid-in-kind and 'tied aid' (i.e. imports directly financed by aid), spending and absorption are equivalent.

In order to understand the implications of these concepts, we can make use of the relevant macroeconomic and fiscal identities. In terms of aid absorption, we start with the following balance of payments identity:

$$\Delta R = CAB + KAB$$

[133] Earlier versions of this paper are Berg et al (2007) and IMF (2005).
[134] "It measures the extent to which aid engenders a real resource transfer through higher imports or through a reduction in the domestic resources devoted to producing exports" (IMF, 2005:3).

where ΔR stands for changes in international reserves, *CAB* is the current account balance, and *KAB* the capital account balance. If we pull out aid inflows from both accounts,[135] we obtain:

$$\Delta R = NACAB + NAKAB + Aid$$

where *NACAB* is the non-aid current account balance, *NAKAB* the non-aid capital account balance, and *Aid* is the net aid inflow. Taking differences and rearranging, we obtain the following expression:

$$\Delta Aid = -\Delta NACAB - \Delta NAKAB - \Delta\Delta R$$

This identity provides some insights into the possible uses of additional aid inflows: (i) to widen the non-aid current account deficit (usually through higher imports); (ii) to widen the non-aid capital account deficit (potentially through capital outflows); and (iii) to increase the accumulation of international reserves. We can now express aid absorption as the deterioration of the non-aid current account balance that is attributed to aid (Aiyar and Ruthbah, 2008):

$$Absorption = -\Delta NACAB/\Delta Aid$$

Assuming that $\Delta Aid>0$, 'full absorption' is achieved when the non-aid current account deficit increases by the same amount of the extra aid inflow (the measure equals 1). A value close to 0 indicates a low level of absorption, and suggests that the additional foreign exchange provided by the aid inflow is used to increase international reserves and/or widen the non-aid capital account deficit.

In terms of aid spending, we start from the usual budget constraint facing the government:

$$I_G + C_G = T + Aid + B + L$$

where I_G stands for public investment, C_G public recurrent expenditures, T domestic revenue, B domestic borrowing and L external (non-concessional) loans. Re-arranging the budget constraint and differencing we obtain:

[135] Some aid inflows are included in the current account (e.g. current grants) while others are incorporated in the capital account (e.g. capital loans) – see Aiyar and Ruthbah (2008).

$$\Delta Aid = -\Delta NAGOB - (\Delta B + \Delta L)$$

where $NAGOB$ is the non-aid government overall balance, i.e. domestic revenues (T) minus total expenditures ($I_G + C_G$). Hence, the potential uses of the additional aid inflows are: (i) to widen the non-aid current account deficit (through higher public spending and/or lower domestic revenues); and (ii) reduce the need for deficit financing (either domestic or external). We can now express aid spending as:

$$Spending = -\Delta NAGOB/\Delta Aid$$

Similarly, 'full spending' is achieved when the additional aid inflows are utilised to expand the non-aid fiscal deficit (the measure equals 1), whereas a value close to 0 suggests that aid has not been significantly spent.

Table 5-1: Possible Combinations in Response to a Scaling Up of Aid

	Absorbed	Not Absorbed
Spent	• Government spends the aid • Central Bank sells the foreign exchange • Current account deficit widens	• Fiscal deficit widens (expenditures are increased) • Central Bank does not sell foreign exchange • International reserves are built up • Inflation increases
Not Spent	• Government expenditures are not increased • Central Bank sells the foreign exchange • Monetary growth is slowed; nominal exchange rates appreciate; inflation is lowered;	• Government expenditures are not increased • Taxes are not lowered • International reserves are built up

Source: Gupta et al (2006:12)

When we take these two concepts together, there are four potential scenarios to be considered:

(i) *Absorb and spend aid*. The government spends the extra aid inflow – either through higher public spending, lower domestic revenue (e.g. cutting taxes), or a mixture of both – while the central bank sells the foreign exchange in the currency market. The fiscal expansion stimulates aggregate demand, which in turn contributes to a higher (public and private) demand for imports. This effect does not create balance of payments problems since the aid inflow finances the increase in net imports – as more foreign currency becomes available to importers. Hence, the foreign exchange is absorbed by the economy through the widening of the non-aid current account deficit (Gupta et al, 2006). This policy combination leads to aid-financed widening deficits, while the central bank's balance sheet remains unaltered (see table below). However, some real exchange rate appreciation may take place to enable this reallocation of resources. The choice of exchange rate regime will affect the mechanism through which the (potential) real exchange rate appreciation may occur – nominal appreciation in a 'pure float' versus higher domestic inflation in a 'fixed peg' (Hussain et al, 2009). This absorb-and-spend combination is often considered to be the ideal policy response to a surge in aid inflows.

(ii) *Absorb but not spend aid*. The government decides not to spend the aid inflow,[136] while the central bank sells the foreign exchange. Foreign aid is thus used to reduce the government's seigniorage requirement since it substitutes domestic borrowing in financing the government deficit (Buffie et al, 2004). Moreover, the central bank sterilises the monetary impact of domestically financed fiscal deficits (Gupta et al, 2006). This policy scenario usually leads to slower monetary growth and alleviates inflationary pressures. Hussain et al (2009) suggest that this could be an appropriate policy response in countries that have not achieved stabilisation – hence facing high domestic deficits and high inflation – or have a large stock of domestic public debt. A reduction in the level of outstanding public debt could 'crowd in' the private sector (both investment and consumption) through its effect on interest rates (Hussain et al, 2009).[137] This increase in aggregate demand would then feed into higher net imports, which would then be financed by the additional foreign exchange available in the currency market.

(iii) *Spend and not absorb aid*. The government spends the additional aid inflow (non-aid fiscal deficit widens), while the central bank allows its foreign exchange reserves to increase. In this case, the extra foreign exchange is not made available to importers but instead is used to build up international reserves. This policy response is similar to a fiscal stimulus in the absence of foreign aid (Hussain et al, 2009). The increase in government spending must be financed by either: (i) monetising the fiscal expansion (i.e. printing domestic currency), which increases money supply and therefore inflation; or (ii) sterilising the monetary expansion (by issuing securities, usually treasury bills), which could lead

[136] It is assumed that neither public spending is increased nor revenues lowered (through tax cuts), which means that aggregate demand remains unchanged. However, a 'balanced budget' approach (i.e. a combination of higher/lower spending and taxes that leaves the non-aid fiscal deficit unchanged) is compatible with this result and can have significant impact on aggregate demand via the fiscal multiplier.

[137] "When debt reaches low levels, however, there are typically limits to the extent to which the financial system can effectively channel additional resources to the private sector. Further attempts to absorb without spending may amount to 'pushing on a string', increasing excess liquidity or even causing capital outflows rather than increased domestic activity" (IMF, 2005:4).

to higher interest rates and potentially crowd out the private sector (Hussain et al, 2009). There is no real resource transfer due to the absence of an increase in net imports. The IMF (2005) argues that this is a "common but problematic response, often reflecting inadequate coordination of monetary and fiscal policies." The net effect on the real exchange rate is uncertain: higher (unmet) demand for net imports contributes to depreciation (via the nominal exchange rate), whilst higher inflation works in the opposite way.

(iv) *Neither absorb nor spend aid*. The government does not use the additional aid inflow to widen the non-aid fiscal deficit, while the central bank increases its foreign exchange reserves. In this scenario, the government 'saves' the incremental aid and the availability of foreign exchange in the currency market is not increased. Once again, this could be a viable (short-run) strategy if the government needs to retire onerous debts (or smooth volatile aid inflows) and foreign reserves are at a precariously low level (Gupta et al, 2006). In the absence of a fiscal expansion, aggregate demand is not affected and there are no pressures on the exchange rate or domestic prices (Hussain et al, 2009). In the long-run, however, this may not be a politically viable strategy due to external and domestic pressures.

The following table provides an example of an aid inflow of 100 monetary units, which formalises the discussion presented above.

Table 5-2: Example with Aid Inflow of 100 Monetary Units

	Absorbed				Not Absorbed			
Spent	Central Bank Balance Sheet				Central Bank Balance Sheet			
	NIR	0	RM	0	NIR	+100	RM	+100
	NDA	0			NDA	0		
	Fiscal Accounts				Fiscal Accounts			
	EF	+100	DEF	+100	EF	+100	DEF	+100
	DF	0			DF	0		
Not Spent	Central Bank Balance Sheet				Central Bank Balance Sheet			
	NIR	0	RM	−100	NIR	+100	RM	0
	NDA	−100			NDA	−100		
	Fiscal Accounts				Fiscal Accounts			
	EF	+100	DEF	0	EF	+100	DEF	0
	DF	−100			DF	−100		

Obs.: 'NIR' net international reserves, 'NDA' net domestic assets, 'RM' reserve money, 'EF' external financing, 'DF' domestic financing, and 'DEF' deficit (excluding aid).
Source: Adapted from Berg et al (2007:8)

A final observation is warranted with regard to aid absorption. The central bank can choose to build up reserves or sell foreign exchange under any exchange rate regime (IMF, 2005:9). Consider the 'corner solutions'. In a pure float, the central bank sells the full amount of foreign exchange (ΔR = 0), so that aid is absorbed and/or exits through the capital account. However, this is an extreme and unusual case. Many African countries operate a de facto managed float, where the central bank intervenes in the foreign exchange market (e.g. accumulation of reserves) due to a 'fear of floating' (Calvo and Reinhart, 2002:388). For this reason, any combination of the three uses of aid would be possible. In a fixed regime (e.g. CFA Franc Zone), the central bank accumulates the foreign exchange to defend the fixed peg ($\Delta R = \Delta Aid$) and none of the aid is absorbed. However, as the fiscal stimulus contributes to higher demand for net imports the central bank may decide to sell the foreign exchange to defend the peg – nominal depreciation pressures in 'spend and not absorb' – which may lead to full absorption (IMF, 2005:9). Hence, one could argue that while the exchange rate regime may condition the short-term response to aid, in the long-run, countries with different exchange rate and monetary frameworks may adopt similar policy responses. This assumption supports the main empirical framework proposed by this paper (pooled mean group estimator), which constrains the long-run impacts to be identical across countries but allows for short-run heterogeneous effects.

The aim of this section was to briefly introduce the analytical framework and concepts that are going to be used in the empirical assessment. The next section presents the empirical findings of the relevant literature.

Empirical Results

Hussain et al (2009) apply the framework described above to five African countries with the objective of examining their policy responses.[138] The countries that have recently experienced a surge in aid inflows and are included in the sample are Ethiopia, Ghana, Mozambique, Tanzania, and Uganda. The table below presents the results.

[138] The same results can be found in IMF (2005).

Table 5-3: Aid Absorption and Spending (% GDP)

Country	Period		ΔNACAB	ΔAid	Aid Absorbed?	ΔGE	ΔT	ΔAid	Aid Spent?
	Pre-Surge	Aid Surge							
Ethiopia	1999-00	2001-03	–1.6	8.0	20%	0.7	1.5	5.9	0%
Ghana	1999-00	2001-03	10.0	5.5	0%	2.3	1.9	6.0	7%
Mozambique	1998-99	2000-02	–3.9	5.9	66%	6.7	1.3	5.0	100%
Tanzania	1998-99	2000-04	2.3	2.2	0%	4.0	0.4	3.9	91%
Uganda	1999-00	2001-03	–1.3	4.7	27%	2.5	0.1	3.2	74%

Source: Adapted from Hussain et al (2009:499-501). All variables are defined as a percentage of GDP. The results are truncated at 0 and 100. The actual 'absorption' value for Ghana and Tanzania is negative. Moreover, the actual 'spending' value for Ethiopia is –14 percent, whereas for Mozambique is 108 percent.

The results suggest that, with the exception of Mozambique, foreign aid inflows were not significantly absorbed. In Ethiopia and Uganda, only 20 and 27 percent of the additional aid was absorbed, respectively. Moreover, Ghana and Tanzania have not used the extra aid inflow to widen the current account deficit. In fact, these countries experienced an improvement in the non-aid current account balance. With regard to aid spending, the estimate for Mozambique suggests that all aid was 'spent', meaning that the non-aid government balance deteriorated by the full amount of aid. In Ethiopia, however, none of the additional aid was 'spent'. In fact, there was an improvement of the non-aid government deficit, since revenue collection increased by more than the increase in government expenditures. The conclusion would be that Ethiopia is saving aid resources, possibly to substitute for domestic borrowing or even to retire public debt. Most of the additional aid inflows were spent in Tanzania and Uganda, whereas in Ghana most of the aid resources were saved. As a result of these findings, the usual policy prescription is that African countries need to significantly improve the management of aid inflows – through better coordination between the government (e.g. treasury) and the central bank.[139]

Foster and Killick (2006) also follow this approach to explore the consequences of scaling up aid flows in four African countries: Mauritania, Sierra Leone, Mozambique and Tanzania (the latter two overlapped with the IMF study). They conclude that aid has been fully absorbed and spent in Mauritania, whereas in Sierra Leone it has been mostly absorbed and partly spent (54 percent, if debt relief excluded).

However, this methodology has a few limitations. For example, the estimates from the absorption and spending equations will be very sensitive to the point in time in which they are evaluated. Defining the pre-aid surge and the surge period will be critical for the results and perhaps the policy conclusions. Moreover, this simple methodology ignores potential dynamic effects. Absorption and spending may well increase after the surge

[139] McKinley (2005) suggests that countries may not fully spend aid inflows due to a 'fear of inflation', while a 'fear of appreciation' hampers full absorption by the central bank.

period (time lag). Finally, one needs to use these concepts with caution, since full 'spending' can be achieved through a total displacement of domestic revenues, in which case aid flows cause a proportional decrease in domestic revenues with no increase in government expenditures. The concept does not distinguish between what would be a desired outcome (e.g. increased developmental expenditures), and a potentially perverse effect that increases aid dependency and threatens long-term sustainability. For that purpose, we will also investigate the impact of aid on public investment.

More recently, Aiyar and Ruthbah (2008) have employed a panel econometric approach to investigate these issues in more detail. Their sample consists of 95 countries, distributed over three categories: sub-Saharan Africa (45), Latin America and Caribbean (19), and Other (31). Annual data for the period 1970-2004 is collected from the World Development Indicators (WDI), Global Development Finance (GDF) and the World Economic Outlook (WEO). The aid variable is taken from the WDI, which is in fact compiled by the OECD-DAC. Most variables in their analysis are measured as a share of GDP. The authors use a system Generalised Method of Moments (system GMM) estimator to evaluate the level of aid absorption and spending. The logarithm of average income per capita (LYPC) and the volatility of aid flows are used as country-specific time-invariant control variables. The logarithm of the terms of trade (LTOT) and inflation are country- and time-specific determinants.

Table 5-4: Aid Absorption and Spending Regressions

	NACABY			NAGOBY		
	Full	Africa	A/Y>0.1	Full	Africa	A/Y>0.1
Lagged Dep. Var.	0.64***	0.63***	0.60***	0.65***	0.62***	0.63***
AIDY	−0.31***	−0.41***	−0.45***	−0.56***	−0.79***	−0.68***
LTOT	−0.20	1.09	2.28			
LYPC	1.34**	1.17*	2.21	−1.63***	−1.72	−0.75
Autocracy Index	−0.01	−0.04	0.21	0.04	0.41***	0.02
Aid Volatility	−0.35***	−0.36***	−0.40***	−0.02	0.13	−0.22*
Inflation				−0.29	−0.91*	−1.35**
Observations	2218	1087	535	1290	539	241
Countries	72	37	18	62	31	15

Source: Compiled from Aiyar and Ruthbah (2008). Aid-dependent countries are those where foreign aid is higher than 10 percent of GDP. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

Their results suggest that aid absorption is statistically significant in the short-run, with a coefficient ranging for −0.31 to −0.45. Technically, this means that a 1 percentage point increase in the aid-GDP ratio will lead to about a third of a percentage point increase in the non-aid current account deficit (as a share of GDP).[140] The results are stronger for the African and aid-dependent samples. Moreover, income per capita seems to have a positive impact on the non-aid current account balance, whilst aid volatility contributes to its deterioration. The results for aid spending appear to be stronger than for absorption. This is not an unexpected result, since countries are often blamed for spending more aid than they absorb (no real resource transfer). In this case, the impact for Africa appears to be stronger than for aid-dependent countries. Finally, the autocracy index appears to improve the government balance while inflation has the opposite effect. For aid-dependent countries, only inflation and aid volatility seem to be statistically significant.

To complement their analysis, Aiyar and Ruthbah (2008) also estimate the impact of foreign aid inflows on the accumulation of international reserves and total domestic investment. Their results suggest that aid has no impact on the accumulation of foreign reserves. Moreover, income per capita and the terms of trade may have a positive impact. Finally, foreign aid contributes to a modest increase in total domestic investment, when controlled for public investment.

Table 5-5: International Reserves and Investment Regressions

	dRY			INVY		
	Full	Africa	A/Y>0.1	Full	Africa	A/Y>0.1
Lagged Dep. Var.	0.02	0.04	−0.05	0.48***	0.43***	0.42***
AIDY	0.05	0.01	0.06	0.14***	0.15***	0.19***
L(ToT)	0.54	0.61*	0.52	−1.79**	−1.70*	−0.59
L(YPC avg)	0.37	0.52*	0.28**	0.60	0.39	0.16
Autocracy Index	−0.02	0.04	0.04	−0.01	0.08	−0.04
Aid Volatility	−0.05	0.01	−0.02	0.20***	0.23***	0.25*
INVPY				0.54***	0.60***	0.68***
Observations	2073	1007	485	1813	888	389
Countries	72	37	18	72	37	18

Source: Compiled from Aiyar and Ruthbah (2008). The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

The table below summarises the short- and long-run effects of an increase in aid inflows. The long-run coefficients are obtained by dividing the estimated (short-run) coefficient by 1 minus the coefficient of the lagged dependent variable. The long-run results indicate that aid has a more than proportional effect on the non-aid current account balance and the non-aid government balance. Although there is no impact on the accumulation of international reserves, aid seems to contribute to a modest increase in total investment. Finally, the authors argue that the unabsorbed aid is leaving the countries through the capital account.

[140] The authors interpret these results differently. For the first column they suggest that (in the short-run) "about 30 cents out of every dollar is absorbed" (Aiyar and Ruthbah, 2008:10).

Table 5-6: Impact of a 1 percentage point increase in the Aid-GDP Ratio

	Full Sample		Africa		A/Y > 0.1	
	Short-Run	Long-Run	Short-Run	Long-Run	Short-Run	Long-Run
Absorption	0.30***	0.83***	0.41***	1.11***	0.45***	1.13***
Spending	0.56***	1.60***	0.79***	2.14***	0.68***	1.48***
Reserves	0.05	0.05	0.01	0.00	0.06	0.00
Investment	0.14***	0.26***	0.15***	0.26***	0.19***	0.33***

Source: Aiyar and Ruthbah (2008:13) for full sample and author's calculations for remaining samples. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

This study will revisit the empirical evidence on aid absorption and spending, with a special focus on low-income countries in Africa. For that purpose we compiled data from several international sources and constructed a new (balanced) panel dataset. Contrary to what is common practice in this field of research, we do not use the OECD-DAC's dataset on aid flows, but instead collect consistent aid data as reported by the recipients. Furthermore, we use alternative panel data methodologies, which we argue are more appropriate to deal with this type of macroeconomic dataset.

5.3 Methodology

This paper uses (linear) panel data regression methods to evaluate how African countries have managed foreign aid inflows. Panel data is a special case of pooled time-series cross-section, in which the same cross-section (e.g. individual) is surveyed over time. In this paper, the cross-section includes African countries, for which annual observations of a number of variables were collected. Baltagi (2008:6-11) provides a good summary of the advantages and disadvantages of using panel data. Here we focus on the aspects that contrast macro panels to time series regressions. Some of the advantages include: (i) controlling for individual heterogeneity;[141] (ii) more informative data, variability, degrees of freedom and efficiency, as well as less collinearity among the variables; (iii) allowing the construction and testing of more complicated behavioural models; and (iv) panel unit root tests that have more power and have standard asymptotic distributions.

In terms of its limitations, the most serious are: (i) the 'poolability' (homogeneity) assumption, although there are formal tests to evaluate its validity; (ii) potential cross-sectional dependence, which complicates the analysis; (iii) some tests and methods require balanced panels; and (iii) cross-country data consistency. With these features in

[141] Unobserved heterogeneity or time-invariant variables that are correlated with explanatory variables (such as history, institutions and political regimes) may cause omitted variable bias in time series regressions.

mind, we now proceed to the presentation of two important methodological approaches – dynamic panel data (DPD) methods and cointegration analysis.

5.3.1 Dynamic Panel Data

Economic relationships often incorporate some degree of dynamic behaviour. To capture this feature, dynamic panel data (DPD) models – which include a lagged dependent variable – are usually considered (Baltagi, 2008:147):

$$y_{it} = \delta y_{i,t-1} + \beta x_{it} + u_{it}$$

where δ is a scalar, x_{it} is a 1 x k vector of explanatory variables and β is a k x 1 vector of coefficients. For the purpose of illustration, assume that u_{it} is a one-way error component model:

$$u_{it} = \mu_i + \nu_{it}$$

where $\mu_i \sim$ IID$(0,\sigma^2_\mu)$ and $\nu_{it} \sim$ IID$(0,\sigma^2_\nu)$ independent of each other and among themselves.

This DPD model is characterised by two sources of persistence over time: (i) autocorrelation due to the lagged dependent variable; and (ii) individual effects capturing country heterogeneity (Baltagi, 2008:147). Estimation of DPD models raises several problems in both fixed- and random-effects. For example, the lagged dependent variable is correlated with the disturbance term (since $y_{i,t-1}$ is a function of μ_i), even if the ν_{it} is not serially correlated (Greene, 2003:307). The OLS estimator is biased and inconsistent in finite samples, especially if T is small. In fact, the coefficients of the explanatory variables will be subject to a downward bias in absolute terms (i.e. biased towards zero). Even for T=30 the fixed-effects (FE) estimator can present a significant bias (Baltagi, 2008:148). The solution is thus to use instrumental variables (IV) regressions or generalised method of moments (GMM) estimators (Greene, 2003:308-14).

Arellano and Bond (1991) developed one- and two-step GMM estimators for dynamic panels ('difference GMM'). They obtain additional instruments by using orthogonality conditions between the lagged dependent variables and the disturbance terms. The difference GMM does not require any prior knowledge of the initial conditions or even the distribution of v_i and μ_i. However, if the dependent variable is very persistent (close to a random walk), then the lagged levels are poor instruments for first-differences and difference GMM performs poorly. Blundell and Bond (1998) develop a 'system GMM' estimator for DPD models to solve the problem of 'weak instruments'. The Blundell-Bond estimator combines moment conditions for the model in first-differences with those for the model in levels. The procedure uses lagged differences of y_{it} as instruments for the equation in levels and lagged levels of y_{it} as instruments for the equation in first-differences. Moreover, it requires a stationary restriction on the initial conditions process (Baltagi, 2008:161). The validity of the moment conditions imposed are usually assessed by a test of over-identifying restrictions (either Hansen's or Sargan's).

The main advantages of these GMM estimators relate to their perceived robustness to heteroscedasticity and non-normality of the disturbances. Moreover, the use of instrumental variables helps address biases arising from reverse causality. Nonetheless, there are some remaining concerns about the efficiency of such methods. The violation of moment conditions (e.g. presence of non-stationarity), will yield inconsistent estimates. Moreover, Roodman (2009) argues that the number (and quality) of instruments generated by difference and system GMM methods can affect the asymptotic properties of the estimators and specification tests. In samples with large T, instrument proliferation can be particularly serious, inducing two main types of problems: (i) overfitting endogenous variables; and (ii) imprecise estimates of the optimal weighting matrix. Greene (2003:307) provides another strong criticism. He argues that introducing a lagged dependent variable to an otherwise long-run (static) equation will significantly change its interpretation, especially for the independent variables. In the case of a DPD model, the coefficient on x_{it} merely represents the effect of new information, rather than the full set of information that influences y_{it}. Finally, it is often argued that while DPD methods are appropriate for panels with a small T, but when T is sufficiently large other methods should be preferred. Hence, we now turn to panel data methods that were specifically developed for 'long' panels.

5.3.2 Panel Cointegration

Traditional panel data econometrics rests on micro panels that usually include thousands of households or hundreds of firms (large N), which are tracked over a few survey rounds (small T). This study, however, uses macroeconomic variables that are collected for several African countries over a significant number of years. The use of panel datasets with these characteristics – large N and large T – presents new challenges to researchers. Panels with a significant temporal dimension are subject to spurious relationships, especially since macroeconomic variables are often characterised by non-stationarity. According to Baltagi (2008:273), the accumulation of observations through time generated two strands of ideas: (i) the use of heterogeneous regressions (one for each country) instead of accepting coefficient homogeneity (implicit in pooled regressions), e.g. Pesaran et al (1999); and (ii) the extension of time series methods (estimators and tests) to panels in order to deal with non-stationarity and cointegration, e.g. Kao and Chiang (2000) and Pedroni (2000).[142] We will pursue both strategies in this paper.

Cointegration analysis in a panel data setting entails similar steps to those usually employed in time series analysis: (i) unit root testing; (ii) cointegration testing; and (iii) estimation of long-run relationships. We take these in turn.

Unit Root Tests

The first step requires an analysis of the stationarity properties of the variables. Panel unit root tests have become a fast-growing area of research in econometrics with a view to improving the perceived low power of individual unit root tests – particularly in small samples. These tests are often grouped into two main categories: (i) first-generation tests, which assume cross-sectional independence – e.g. Levin et al (2002), Im et al (2003), Maddala and Wu (1999), and Choi (2001); and (ii) second-generation tests, which explicitly allow for some form of cross-sectional dependence – e.g. Pesaran (2007). As a starting point, consider the following autoregressive (AR) process for panel data:

$$y_{it} = \rho_i y_{i,t-1} + \delta_i Z_{it} + u_{it}$$

where ρ_i is the AR coefficient and the error term u_{it} is assumed to be independent and identically distributed (i.i.d.). Moreover, Z_{it} includes individual deterministic effects, such as constants ('fixed effects') and linear time trends, which capture cross-sectional heterogeneity.

[142] Moreover, the estimators for panel cointegrated models and related statistical tests are often found to have different asymptotic properties from their time series counterparts (Baltagi, 2008:298). An important contribution is Phillips and Moon (1999, 2000), who analyse the limiting distribution of double indexed integrated processes.

Levin et al (2002)[143] propose a test (LLC) that can be seen as a panel extension of the augmented Dickey-Fuller (ADF) test:

$$\Delta y_{it} = \alpha y_{i,t-1} + \sum_{j=1}^{p_i} \beta_{ij} \Delta y_{i,t-j} + \delta_i Z_{it} + u_{it}$$

Since the lag length of the differenced terms (p_i) is unknown, Levin et al (2002:5-7) suggest the following three-step procedure: (i) carry out separate ADF regressions for each individual and generate two orthogonalised residuals;[144] (ii) estimate the ratio of long-run to short-run innovation standard deviation for each individual; (iii) compute the pooled *t*-statistics, with the average number of observations per individual and average lag length. In this test, the associated AR coefficient is constrained to be homogenous across individuals (i.e. $\alpha_i = \alpha$ for all i). Hence, the null hypothesis assumes a common unit root (H_0: $\alpha = \rho - 1 = 0$) against the alternative hypothesis that each time series is stationary (H_1: $\alpha < 0$). The authors show that the pooled *t*-statistic has a limiting normal distribution under the null hypothesis. This test is often recommended for moderate sized panels, especially for N>10 and T>25.

Im et al (2003)[145] extend the LLC test by allowing heterogeneity on the AR coefficient. In practice, the test entails the estimation of individual ADF regressions, and then combining this information to perform a panel unit root test. This approach allows for different specifications of the coefficients (α_i for each cross-section), the residual variance and lag-length (Asteriou and Hall, 2007:368). The authors propose a *t*-bar statistic, based on the average of the individual unit root (ADF) test statistics. This statistic evaluates whether the coefficient α is non-stationary across all individuals (H_0: $\alpha_i = 0$ for all i), against the alternative hypothesis that at least a fraction of the series is stationary (H_1: $\alpha_i < 0$ for at least one i). Both LLC and IPS tests require N to be small enough relative to T, whilst the LLC test also requires a strongly balanced panel (Baltagi, 2008:280).

Breitung (2000) uses Monte Carlo experiments to show that the power of the LLC and IPS tests statistics is sensitive to the specification of the deterministic components, such as the inclusion of individual specific trends (Baltagi, 2008:280). He proposes a test statistic based on modifications to the LLC steps to overcome these difficulties. Breitung's test

---

143 Originally published in 1992 and one of the first panel unit root tests in the literature.

144 Here, the lag order of the differenced terms (pi) is allowed to vary across individuals and is usually determined by a lag selection criterion (to correct for serial correlation).

145 The IPS test was originally published in 1997.

statistic assumes a common unit root process and is also shown to be asymptotically distributed as a standard normal. The test is often suggested for samples of around N=20 and T=30.

Maddala and Wu (1999:637) and Choi (2001) suggest the use of nonparametric Fisher tests. The main feature of these tests is that they combine the probability limit values (*p*-values) of unit root tests from each cross-section rather than average test statistics. Fisher tests are usually implemented using individual ADF or Phillips-Perron unit root tests, and their asymptotic distribution follows a chi-square (*P*-test).[146] Choi (2001) also proposes an alternative Fischer-type statistic that follows a standard normal distribution (*Z*-test). Both IPS and Fischer-type tests combine information of individual unit root tests, but simulation studies suggest that Fischer tests have better power properties than the IPS test. The disadvantage of Fischer-type tests relates to the need to derive *p*-values through Monte Carlo simulations.

Hadri (2000) proposes a residual-based Lagrange multiplier (LM) test, which is in fact a panel generalisation of the KPSS test (Baltagi, 2008:282). The test uses the residuals from individual OLS regressions of y_{it} on deterministic components (constant and trend) to compute the LM statistic. This test also differs from the previous in the sense that it is a stationarity test. The null hypothesis assumes no unit root in any of the time series (all panels stationary), against the alternative of non-stationarity for, at least, some cross-sections.

The main drawback of the first-generation tests described above relates to the assumption that the data is independent and identically distributed (i.i.d.) across individuals (cross-section independence). In practice, this means that the movements of a given variable through time are independent across countries. This restrictive assumption has often been challenged by empirical studies, and it should be evaluated on a case-by-case basis.[147] Some cross-sectional dependence tests include Pesaran (2004) and a Breusch-Pagan LM statistic (for $T>N$). Banerjee et al (2005) show that in the presence of cross-section dependence, first-generation tests tend to have serious size distortions and therefore

[146] Maddala and Wu (1999:645) also suggest a bootstrap procedure to account for cross-sectional dependence, but size distortions are only decreased rather than eliminated.

[147] Levin et al (2002) suggest 'demeaning' the data in order to attenuate the biases caused by the presence of cross-sectional dependence, which involves subtracting cross-sectional averages (for each time period) from the series before the use of unit root tests. Nonetheless, this procedure cannot ensure the successful elimination of the bias.

perform poorly. This often leads to the over-rejection of the null hypothesis (unit root) when the sources of non-stationarity are common across individuals.

These findings led to the development of unit root tests for panels with cross-sectional dependence (second-generation tests). Pesaran (2007) suggests a simple method to remove the influence of cross-sectional dependence, which involves augmenting standard ADF regressions with the cross-section averages of lagged levels and first-differences of the individual series. These individual cross-sectionally augmented Dickey-Fuller (CADF) statistics (or the corresponding p-values) can then be used to develop modified versions of standard panel unit root tests – such as IPS's t-bar, Maddala and Wu's P, or Choi's Z. The tests are applicable for both when $N>T$ and $T>N$, and are shown to have good size and power properties, even when N and T are relatively small (e.g. 10). However, the t-bar statistic (CIPS) can only be computed for balanced panels. For unbalanced panels, the modified Z test can be reported.

Table 5-7: Characteristics of Unit Root Tests

Test	Null	Alternative Hypothesis	Deterministic Components	Autocorrelation Correction	Cross-Section Dependence	Unbalanced Panel (Gaps)
LLC	UR	No UR	None, F, T	Lags	demean	No (–)
Breitung	UR	No UR	F, T	Lags	robust[1]	No (–)
IPS	UR	Some CS without UR	None, F, T	Lags	demean	Yes (No)
Fisher	UR	Some CS without UR	None, F, T	Lags/Kernel	demean	Yes (Yes)
Hadri	No UR	Some CS with UR	F, T	Kernel	robust[1]	No (–)
Pesaran	UR	Some CS without UR	F, T	Lags	robust	Yes (No)

Obs.: 'UR' unit root, 'CS' cross-sections, 'None' no exogenous variables, 'F' fixed effect, 'T' individual effect and individual trend. [1] Stata's 'xtunitroot' command computes robust versions that account for cross-sectional dependence.

Source: Compiled from QMS (2007:110, corrected) and Stata's 'xtunitroot' command help.

Cointegration Tests

The panel unit root tests proposed above aim to assess the order of integration of the variables. If the main variables are found to be integrated of order one, then we should use panel cointegration tests to address the non-stationarity of the series. As before, some of these tests were developed as extensions of earlier tests for time series data.

Pedroni (1999, 2004:604) provides cointegration tests for heterogeneous panels based on the two-step cointegration approach of Engle and Granger (1987). Pedroni uses the residuals from the static (long-run) regression and constructs seven panel cointegration test statistics: four of them are based on pooling (within-dimension or 'panel statistics test'), which assumes homogeneity of the AR term, whilst the remaining are less restrictive (between-dimension or 'group statistics test') as they allow for heterogeneity of the

AR term. The assumption has implications on the computation of the second step and the specification of the alternative hypothesis. The *v*-statistic is analogous to the long-run variance ratio statistic for time series, while the *rho*-statistic is equivalent to the semi-parametric 'rho' statistic of Phillips and Perron (1988). The other two are panel extensions of the (non-parametric) Phillips-Perron and (parametric) ADF *t*-statistics, respectively. These tests allow for heterogeneous slope coefficients, fixed effects and individual specific deterministic trends, but are only valid if the variables are I(1). Pedroni (1999) derived critical values for the null hypothesis of no cointegration.

Kao (1999) proposes residual-based DF and ADF tests similar to Pedroni's, but specifies the initial regression with individual intercepts ('fixed effects'), no deterministic trend and homogeneous regression coefficients. Kao's tests converge to a standard normal distribution by sequential limit theory (Baltagi, 2008:293). Both Kao and Pedroni tests assume the presence of a single cointegrating vector, although Pedroni's test allows it to be heterogeneous across individuals.

Maddala and Wu (1999) propose a Fisher cointegration test based on the multivariate framework of Johansen (1988). They suggest combining the *p*-values of individual (system-based) cointegration tests in order to obtain a panel test statistic. Moreover, Larsson et al (2001) suggest a likelihood ratio statistic (LR-bar) that averages individual rank trace statistics. However, the authors find that the test requires a large number of temporal observations. Both of these tests allow for multiple cointegrating vectors in each cross-section.

Westerlund (2007) suggests four cointegration tests that are an extension of Banerjee et al (1998). These tests are based on structural rather than residual dynamics and allow for a large degree of heterogeneity (e.g. individual specific short-run dynamics, intercepts, linear trends and slope parameters).[148] All variables are assumed to be I(1). Moreover, bootstrapping provides robust critical values in cases of cross-section dependence. The tests assess the null hypothesis that the error correction term in a conditional ECM is zero – i.e. no cointegration (Baltagi, 2008:306).

Banerjee et al (2004) argue that although these tests allow for cross-sectional dependence (via the effects of short-run dynamics), they do not consider long-run dependence, in-

[148] Westerlund (2007:710) argues that residual-based cointegration tests require the long-run cointegration vector in levels to equal the short-run adjustment process in differences – also known as 'common factor restrictions'. The trade-off is the assumption of weak exogeneity that ECM-based tests depend upon.

duced by cross-sectional cointegration. The authors demonstrate that in that case, panel cointegration tests may be significantly oversized (Baltagi, 2008:302-3). Moreover, most cointegration tests may be misleading in the presence of stationary data, as they require all data to be I(1).

Estimation of the Long-Run

A complementary issue relates to the efficient estimation of long-run economic relationships. In the presence of cointegrating non-stationary variables, one would like to be able to efficiently estimate and test the relevant cointegrating vectors. For that purpose, a number of panel estimators have been suggested in the literature. Once again, most of them are developed as extensions of well-known time series methods. An important difference is that the panel OLS estimator of the (long-run) static regression model, contrary to its time series counterpart, is inconsistent (Baltagi, 2008:299).

Kao and Chiang (2000) propose a panel dynamic OLS estimator (DOLS) which is a generalisation of the method originally proposed by Saikkonen (1991) and Stock and Watson (1993) for time series regressions. The regression equation is:

$$y_{it} = \alpha_i + \beta' X_{it} + \sum_{j=-q}^{q} c_{ij}\, \Delta X_{i,t+j} + \varepsilon_{it}$$

where X_{it} is a vector of explanatory variables, β the estimated long-run impact, q the number of leads and lags of the first-differenced data, and c_{ij} the associated parameters. The estimator assumes cross-sectional independence and is asymptotically normally distributed. The authors provide Monte Carlo results suggesting that the finite-sample properties of the DOLS estimator are superior to both fully-modified OLS (FMOLS)[149] and OLS estimators.

Pesaran et al (1999) suggest a (maximum-likelihood) pooled mean group (PMG) estimator for dynamic heterogeneous panels. The procedure fits an autoregressive distributed lag (ARDL) model to the data, which can be re-specified as an error correction equation to facilitate economic interpretation. Consider the following error correction representation of an ARDL(p, q, q,…, q) model:

[149] The panel FMOLS estimator developed by Phillips and Moon (1999) and Pedroni (2000) is a generalisation of the estimator proposed by Phillips and Hansen (1990).

$$\Delta y_{it} = \phi_i y_{i,t-1} + \beta_i' X_{it-1} + \sum_{j=1}^{p-1} \lambda_{ij} \Delta y_{i,t-j} + \sum_{j=0}^{q-1} \delta_{ij}' \Delta X_{i,t-j} + \mu_i + \varepsilon_{it}$$

where X is a vector of explanatory variables, β_i contains information about the long-run impacts, ϕ_i is the error correction term (due to normalisation), and δ_{ij} incorporates short-run information. The PMG can be seen as an intermediate procedure, somewhere between the mean group (MG) estimator and the dynamic fixed-effects (DFE) approach.[150] The MG estimator is obtained by estimating N independent regressions and then averaging the (unweighted) coefficients, whilst the DFE requires pooling the data and assuming that the slope coefficients and error variances are identical. The PMG, however, restricts the long-run coefficients to be same (β=β_i for all i), but allows the short-run coefficients and error variances to vary across countries (Pesaran et al, 1999:621). This approach can be used whether the regressors are I(0) or I(1) (Pesaran et al, 1999:625).

5.4 Data

The data used in this paper was collected from the International Monetary Fund's (IMF) Balance of Payments Statistics (BoPS) and the World Bank's Africa Database. Complementary sources included the United Nations' National Accounts Main Aggregates Database, the IMF's World Economic Outlook (WEO), and the World Bank's World Development Indicators (WDI). There was a significant effort to construct a balanced panel for all 53 African countries covering the period 1970-2007. However, data for 1970-1979 is scarce for many countries, whereas data for 2006-2007 is usually based on estimates or projections. Moreover, data on aid flows for 2006 often contains outliers due to very large debt relief grants, which cannot be satisfactorily expunged. Hence, we have built a balanced panel for 1980-2005 for the macroeconomic variables, while for the fiscal variables we have a balanced panel for 1990-2005. It should be noted that our aid variables only include grants, due to the lack of data on concessional foreign loans. Nonetheless, there is a strong argument to separate these since aid grants and aid loans often have significantly different economic impacts.[151] Finally, seven countries had to be excluded from the initial

[150] A further estimation alternative would be Zellner's seemingly unrelated regression (SUR), but this procedure requires that N is significantly smaller than T, which unfortunately is not our case. Moreover, Pesaran et al (1999:626) argue that Swamy's static random coefficient model is asymptotically equivalent to the MG estimator.
[151] In practice, we allow 'aid loans' to remain lumped with foreign non-concessional loans.

sample. These countries either reached independence only in the 1990s (Eritrea and Namibia) or lack reliable data (Congo DR, Djibouti, Liberia, Somalia and Zimbabwe).

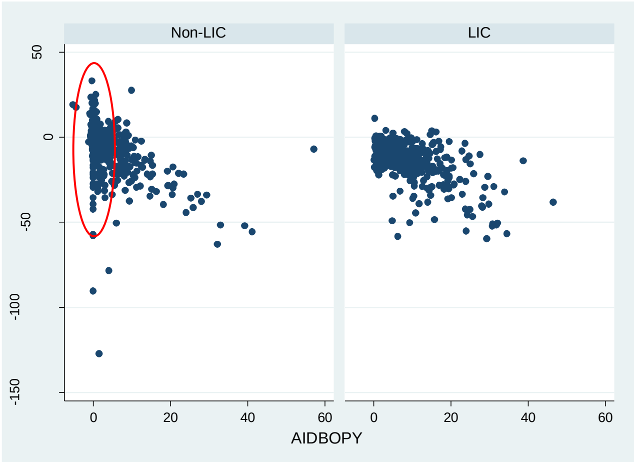
The list of variables includes:

NACABY	Non-Aid Current Account Balance (% GDP)
AIDBOPY	Aid Grants (% GDP), as reported by the Balance of Payments Statistics
LTOT	Logarithm of the Terms of Trade
DRY	Change in International Reserves (% GDP)
NAGOBY	Non-Aid Government Overall Balance (% GDP)
AIDGOVY	Aid Grants (% GDP), as reported by the World Bank's Africa Database
INF	Inflation Rate (CPI, percentage change)
INVGY	Gross Public Fixed Capital Formation (% GDP)
BORY	Domestic Financing (% GDP)

The following graphs provide pair-wise plots of the main variables of interest. The full sample of African countries is utilised, as well as a sub-sample incorporating low-income countries (LICs) only.[152] The plots confirm the strong negative correlation between foreign aid and the macroeconomic and fiscal balances. This suggests that aid inflows are used, at least to a certain extent, to increase the (non-aid) current account and budget deficits. However, there is an important concern arising from the observation of these graphs. It appears that richer countries may potentially distort the analysis. This is because middle-income countries tend to be less aid-dependent, and therefore the relationship between aid inflows and other economic variables can be significantly weaker. The inclusion of these countries may thus affect the magnitude and significance of the estimated coefficients, leading us to believe that aid absorption and spending is lower than desired. Further plots lead to similar conclusions for reserve accumulation and public investment. Finally, some of the richer countries are (at times) net 'donors', which further complicates the analysis.

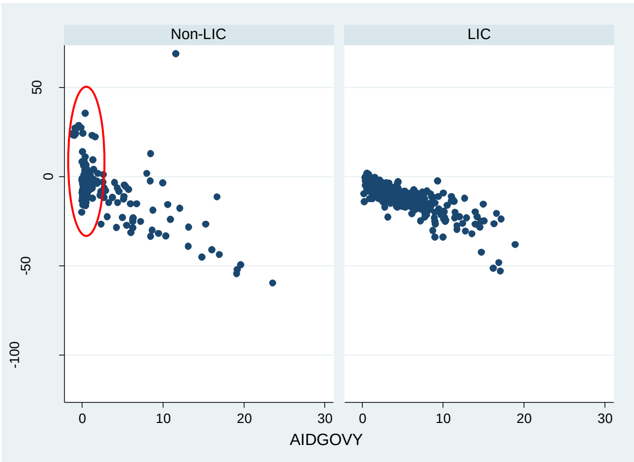
[152] As defined by the World Bank (July 2009).

Figure 5-1: Non-Aid Current Account Balances and Foreign Aid



Obs.: Excludes Lesotho (LSO)

Figure 5-2: Non-Aid Government Balances and Foreign Aid



Obs.: Excludes the Republic of Congo (COG)

The table below presents pair-wise correlations between the main variables of interest. The results corroborate the decision to exclude middle-income countries from the analysis, as for both macroeconomic and fiscal dimensions the (negative) correlations of the non-aid balances with foreign aid inflows are significantly stronger for low-income countries.

Table 5-8: Correlations

	ALL		LIC	
1980-2005	NACABY	AIDBOPY	NACABY	AIDBOPY
NACABY	1.00		1.00	
AIDBOPY	−0.56	1.00	−0.63	1.00
1990-2005	NAGOBY	AIDGOVY	NAGOBY	AIDGOVY
NAGOBY	1.00		1.00	
AIDGOVY	−0.72	1.00	−0.83	1.00

For the reasons presented above, this study will continue the analysis for the 25 African low-income countries in the sample. The following tables present basic statistics on the main variables.[153] As expected, both NACABY and NAGOBY have negative means, with fairly low maximums (surpluses). This highlights the importance of aid inflows in balancing these accounts. The average for AIDBOPY is higher than that for AIDGOVY, probably reflecting the presence of 'off-budgets' – i.e. aid flows not recorded in the budget, perhaps because they are implemented by the donor. BORY has a positive (but low) mean value.

Table 5-9: Basic Statistics for Macroeconomic Variables (1980-2005)

	Obs.	Mean	SD	Min	Max
NACABY	650	–12.2	9.6	–59.7	11.1
AIDBOPY	650	7.7	6.5	0.2	46.5
LTOT	650	4.7	0.4	2.7	5.8
DRY	650	–0.5	3.1	–16.0	34.9

Table 5-10: Basic Statistics for Fiscal Variables (1990-2005)

	Obs.	Mean	SD	Min	Max
NAGOBY	400	–11.6	7.4	–53.0	1.9
AIDGOVY	400	4.9	3.5	0.2	18.9
INF	400	13.4	20.5	–10.9	183.3
INVGY	400	7.6	3.8	1.4	32.2
BORY	400	0.8	2.5	–6.7	13.8

5.5 Empirical Results

This section undertakes a comprehensive econometric exercise to evaluate the uses of foreign aid in Africa's low-income countries. The sample for the macroeconomic specification (absorption) runs from 1980 to 2005, while the fiscal regressions (spending) cover the period 1990-2005. The previous section has demonstrated that the inclusion of middle-income countries – for whom foreign aid plays a much lesser role – can significantly distort the analysis. Hence, the sample in this section is restricted to the 25 African low-income countries in our sample.

Since we are dealing with macroeconomic and fiscal variables that are often found to be non-stationary, we first undertake panel unit root tests to evaluate their order of integration. The results provide evidence that at least some of the variables are non-stationary. We then apply panel cointegration tests to assess whether there are long-run relationships

[153] Not surprisingly, the full sample shows lower absolute averages and higher standard deviations for the aid variables.

amongst the variables of interest. Finally, long-run relations are estimated through appropriate and efficient methods.

5.5.1 Panel Unit Roots

We start with the application of panel unit root tests. A detailed description of the specific characteristics of each test was provided in a previous section. All test specifications include a deterministic time trend. In the LLC, IPS and Fisher-type tests, cross-sectional means are subtracted to minimise problems arising from cross-sectional dependence. The Pesaran test and the versions of the Breitung and Hadri tests used here allow for cross-sectional dependence.[154] However, this version of the Breitung test requires $T>N$. In the LLC and IPS tests, the Bayesian (Schwarz) information criterion (BIC) is used to determine the country-specific lag length for the ADF regressions, with a maximum lag of 3. Moreover, the Bartlett kernel was used to estimate the long-run variance in the LLC test, with maximum lags determined by the Newey and West bandwidth selection algorithm. Finally, the Fisher-ADF and Pesaran's CADF tests include 2 lags.

Table 5-11: Panel Unit Root Tests

	LLC	IPS	Fisher		Breitung	Hadri	Pesaran	
	t*	W-t-bar	ADF-Pm	PP-Pm	λ	z	t-bar	z
NACABY	−1.20*	−2.56***	0.99**	7.33***	−0.09	16.77***	−1.80	2.76
AIDBOPY	−3.85***	−3.82***	1.49*	7.12***	−1.73**	17.40***	−1.86	2.43
LTOT	−1.09	−2.51***	1.72**	4.87***	−2.10**	19.45***	−1.75	3.03
DRY	−13.25***	−13.87***	4.89***	29.26***	−6.48***	1.68**	−2.98***	−3.58***
NAGOBY	−4.04***	−4.14***	0.88	7.77***	n/a	8.06***	−1.96	1.66
AIDGOVY	−5.44***	−4.19***	3.89***	7.93***	n/a	7.80***	−1.84	2.23
INF	−51.40***	−18.44***	1.65**	6.26***	n/a	6.29***	−0.91	6.85
INVGY	−5.21***	−5.40***	2.47***	6.85***	n/a	6.83***	−1.75	2.68
BORY	−8.53***	−6.90***	1.57*	15.16***	n/a	1.74**	−2.41	−0.60

Obs.: Test results generated by Stata. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

The test results provide mixed evidence on the order of integration of the variables. The LLC test strongly rejects the null hypothesis of unit roots, except for LTOT and NACABY. The IPS test rejects the presence of unit roots for all variables. The results for the Fisher-type tests seem to depend on the underlying unit root test chosen. The Phillips-Perron option rejects the null hypothesis for all variables, whilst the ADF alternative presents significantly weaker evidence for some. For example, it cannot reject the presence of unit roots in NAGOBY, and has higher *p*-values for most variables. The last four columns show

[154] The robust versions of the Breitung and Hadri tests are implemented by the new STATA command 'xtunitroot.'

the tests that are robust to the presence of cross-sectional dependence. This variant of the Breitung test is only valid for the longer panel (*T*>*N*). The evidence it provides is mixed, with NACABY appearing to be non-stationary, while the other macroeconomic variables reject unit roots at 5 percent. The Hadri test has a different null hypothesis (stationarity) and provides strong evidence that (at least) some panels have unit roots. This test is an interesting alternative since it challenges the usually strong null hypothesis that all panels have unit roots. Finally, the Pesaran CADF test suggests that all variables have unit roots, except for DRY. The results from the CADF test are robust to the lag structure and specification of determinist components – with the exception of BORY, where a lower lag order (1) suggests that the variable is stationary.

Hence, while the IPS and Fisher-PP test results lead to the conclusion that all variables are stationary, the Hadri and Pesaran tests suggest the opposite (with the exception of DRY for the CADF test). The remaining tests (LLC, Fisher-ADF and Breitung) provide mixed evidence.[155] This observation may lead us to believe that there is some level of cross-sectional dependence affecting the results. Although the cross-sectional averages were subtracted from each series (de-meaning) prior to applying the LLC, IPS and Fisher-type tests,[156] there may still be some residual dependence left, which leads to the over-rejection of the null hypothesis of unit roots. However, we have also applied the original versions of the Hadri and Breitung tests, which are not robust to cross-sectional dependence, and we achieved fairly similar conclusions (no de-meaning applied either). Overall, it is fair to conclude that there is (at least) some non-stationarity that needs to be properly addressed.

5.5.2 Cointegration Tests

Despite the fact that (some of) the data is non-stationary, we may still be able to make valid inference if there is a meaningful relationship amongst the variables of interest. This will be the case if we find a linear combination that produces stationary error terms. The table below reports the results from several cointegration tests. The top row describes the variables included in the tentative cointegrating vectors. The Pedroni and Kao tests use the Bayesian information criterion (BIC) to automatically select the appropriate lag length (maximum set to 3). Moreover, spectral estimation is undertaken by the Bartlett kernel

[155] As noted before, the LLC and IPS tests require *N* to be relatively smaller than *T*, which is not the case here.
[156] For each time period, the mean of the series (across panels) is calculated and then subtracted from the observations.

with the bandwidth selected by the Newey-West algorithm. Whilst the Pedroni and Kao tests are based on the residuals of the long-run static regression, the Westerlund test assesses the significance of the adjustment coefficient in the ECM specification. For the latter we specify the error correction equations with one lag and use a Bartlett kernel window of 3. Deterministic time trends are not included in the specifications since these are generally found to weaken cointegration results. This is later supported by their lack of statistical significance in the error correction models. All tests are derived under the null hypothesis of no cointegration.

Table 5-12: Cointegration Tests

	Statistic	NACABY AIDBOPY LTOT	DRY AIDBOPY LTOT	NAGOBY AIDGOVY INF	INVGY AIDGOVY INF	BORY AIDGOVY INF
Pedroni	Panel-*v*	0.05	−1.26	−1.04	−0.68	2.22
	Panel-*rho*	−3.35***	−9.18***	−2.56***	−0.11	−2.60***
	Panel-PP	−6.58***	−13.69***	−6.86***	−3.06***	−8.60***
	Panel-ADF	−7.12***	−14.02***	−6.89***	−3.55***	−8.05***
	Group-rho	−2.48***	−7.20***	−0.80	1.34	−1.13
	Group-PP	−8.55***	−15.29***	−8.41***	−3.80***	−9.86***
	Group-ADF	−7.86***	−15.42***	−8.42***	−4.37***	−7.90***
Kao	*t*	−2.71***	−2.16**	−2.23**	−1.18	−1.93**
Westerlund	G*t*	−3.55***	−5.91***	−0.65	0.78	0.96
	G*a*	−1.73**	−3.81**	3.97	4.69	3.86
	P*a*	−5.20***	−5.58***	0.64	2.13	3.40
	P*a*	−5.02***	−6.41***	2.18	2.87	2.86

Obs.: Test results generated by EViews and 'xtwest' Stata module. Panel tests tend to have higher power than Group tests, since pooling increases efficiency. Pedroni's Panel statistics are weighted, as well as (all of) Westerlund's. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

The first column examines a vector of variables that includes the non-aid current account balance (NACABY), foreign aid inflows (AIDBOPY) and the logarithm of the terms of trade (LTOT). With the exception of Pedroni's Panel-*v* statistic, all tests reject the null hypothesis of 'no cointegration' among the variables. Hence, while unit root tests provided support for the presence of stochastic trends in the data, cointegration tests suggest that these trends have cancelled each other out – leading to stationary residuals. In practice, this means that these variables have a significant long-run relationship. The second column evaluates whether changes in international reserves (DRY), foreign aid inflows (AIDBOPY) and the terms of trade (LTOT) share a common stochastic trend. Once again, the results strongly suggest the presence of cointegration, but this can be a result of the fact that DRY is a stationary variable – as suggested by most unit root tests.

With regard to the third column, we test whether there is a relationship between the non-aid government balance (NAGOBY), foreign aid inflows (AIDGOVY), and inflation (INF).

Here, the four Westerlund statistics and two Pedroni tests do not reject the null. Moreover, the fourth column – which investigates whether public investment (INVGY), aid inflows (AIDGOVY) and inflation (INF) are a cointegrating relation – provides similar results, and so does the last one. Since the fiscal sample is significantly shorter (and in fact $N>T$) it may be that some cointegration tests (especially Westerlund's) have poor power properties.[157] If we set the lag to zero and exclude INF from the fiscal vectors, the majority of Westerlund's tests reject the null, which may highlight the lack of power of the test.

We did not perform the Maddala and Wu Fisher tests because they can be quite onerous to implement and may provide unreliable results. Since these tests require fitting vector autoregression (VAR) models to each cross-sectional unit, the usual caveats of the time series literature apply. This means that the (individual) tests are only appropriate if the VAR model is correctly specified, which requires a significant amount of individual testing – for example, checking (residual) serial correlation. Moreover, because T is relatively short for a time series study, these tests may have poor size properties.

Overall, the results appear to suggest that the variables of interest are cointegrated, which means that we have uncovered meaningful long-run relationships. However, these tests have some limitations. In the presence of cross-sectional dependence/cointegration, the test results may be biased. Moreover, these tests are developed under the assumption that all variables are I(1). If some of the variables are truly stationary (e.g. DRY), inference might be invalid. Nonetheless, the next section may provide further evidence of cointegration if, as expected, the error correction terms are statistically significant.

5.5.3 Specification and Estimation (Long-Run)

We now use panel data estimation methods to investigate, amongst other things, the impact of foreign aid inflows on the non-aid current account balance and the non-aid government overall balance. Our empirical specifications are similar to Aiyar and Ruthbah (2008), but do not include the time-invariant country-specific control variables.[158] Hence we have:

---

[157] Low power means that the test is not able to reject the null hypothesis when the alternative is correct (Type II error).

[158] These variables are average income per capita and a measure of aid volatility. The methods used in this paper (except for system GMM) do not allow the inclusion of time-invariant variables, since these induce perfect multicollinearity. Moreover, the autocracy index is also excluded since it exhibits little variation through time. An alternative income per capita variable (with annual observations) would create significant distortions due to its interaction with the denominator of our variables (GDP).

$$y_{it} = \alpha_i + \beta_1 AIDY_{it} + \beta_2 x_{it} + \varepsilon_{it}$$

where y_{it} includes the non-aid current account balance (NACABY), accumulation of international reserves (DRY), non-aid government overall balance (NAGOBY), public investment (INVGY) and domestic financing (BORY) – all expressed as a share of GDP. $AIDY_{it}$ is the relevant foreign aid variable, whilst x_{it} is a control variable: the logarithm of the terms of trade (LTOT) in the macroeconomic specifications (firs two) and inflation (INF) in the fiscal specification (last three). Potential reverse causality between the fiscal variables and inflation is addressed in some of the empirical methodologies utilised. The estimates for β_1 contain information about the impact of aid on y_{it}.[159]

The panel data analysis is conducted for the 25 African low-income countries in our sample. The tables below report the results from a number of alternative estimation methods. The aim is to analyse the robustness of the results to different empirical strategies. We start by applying the popular system GMM (SYS-GMM) estimator in the context of a (fixed-effects) lagged dependent variable model. In comparison with the OLS (OLS-FE) and difference GMM (DIF-GMM) alternatives, this estimator is likely to minimise the bias and inconsistency associated with the presence of a lagged dependent variable. However, given the relatively large T in this study, we argue that a methodology based on the time series properties of the data may provide more efficient estimates of the coefficients of interest. Therefore, we complement the SYS-GMM results with the dynamic OLS (DOLS) approach and the maximum-likelihood estimates (MLE) for the error correction model. The DOLS methodology entails the estimation of the static long-run relation augmented by leads and lags of the first-differenced explanatory variables. We chose to include two leads and two lags in the specification, and report robust standard errors. This strategy improves the efficiency of the long-run estimates, but does not provide much guidance on short-run behaviour. Therefore, we also use Pesaran's pooled mean group (PMG) estimator, which uses the panel extension of the single-equation autoregressive distributed lag (ARDL) model. It can be shown that the ARDL has an error correction representation, which is a particularly convenient feature for aiding economic interpretation. We are then able to efficiently estimate the long-run relationships whilst providing information about short-run behaviour (e.g. contemporaneous impacts and speed of adjustment to equilibrium). Another advantage is that while the long-run coefficients are assumed to be

[159] The PMG methodology, for example, requires that foreign aid is exogenous. This may not constitute a major concern since our aid variable only includes aid grants. Aid loans (e.g. IMF lending) tend to be more responsive to domestic conditions (e.g. balance of payments crisis and fiscal imbalances).

homogeneous (i.e. identical across panels), the short-run coefficients are allowed to be country-specific (heterogeneity). This methodology is appropriate for non-stationary panels where N and T are relatively large. For example, Pesaran et al (1999) apply their approach to two empirical examples with the following dimensions: (i) T=32 and N=24; and (ii) T=17 and N=10. We also estimate a mean group (MG) alternative, which allows the long-run parameters to vary, and then test the PMG's poolability assumption through a Hausman test. Finally, we also report the dynamic fixed-effects (DFE) estimator, which assumes short- and long-run parameter homogeneity.[160]

Table 5-13: Estimation Results for NACABY

	SYS-GMM NACABY	DOLS NACABY	DFE D.NACABY	PMG D.NACABY	MG D.NACABY
C	16.01	12.57	5.86	−3.11***	3.00
AIDBOPY	−0.43***	−0.86**			
LTOT	−3.66*	−3.80			
NACABY(−1)	0.62***		−0.31***	−0.40***	−0.64***
D.AIDBOPY			−0.63***	−0.63***	−0.76***
D.LTOT			0.32	1.14	−0.51
Cross	25	25	25	25	25
Time	25	21	25	25	25
Hausman					0.93
Long-run					
AIDBOPY	*−1.13****	−0.86**	−0.76***	−0.62***	−0.96**
LTOT	*−3.66**	−3.80	−5.28**	0.31	−2.88

Obs.: Robust standard errors. Coefficients in *italic* are calculated from the estimation output. SYS-GMM generates 327 instruments for 625 observations. The speed of adjustment for SYS-GMM equals one minus the coefficient on the lagged dependent variable (0.38). The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

We start by investigating the relationship between foreign aid inflows (AIDBOPY), the logarithm of the terms of trade (LTOT) and the non-aid current account balance (NACABY). In the SYS-GMM, the estimated coefficients of the explanatory variables are usually taken to represent short-term impacts, whilst long-run impacts are approximated by the short-term coefficient divided by 1 minus the coefficient of the lagged dependent variable. The results suggest that an increase by 1 percentage point in the aid-GDP ratio leads to an immediate deterioration of the non-aid current account balance by about 0.4 percentage points. Alternatively, we could say that about 40 percent of the aid inflow is being absorbed in the short-run (see Berg et al, 2007). In the long-run, its impact increases to around −1.1 percentage points (full absorption). The second column reports the dynamic OLS (DOLS) specification, which only provides information on the long-run. The coefficient is also significantly high (−0.86), suggesting almost full absorption. The last three columns provide the dynamic fixed-effects (DFE), pooled mean group (PMG) and

[160] This will yield different results from the OLS-FE approach since it is based on the error correction specification rather than the lagged dependent variable model. The latter is, in fact, an ARDL(1,0,…,0).

mean group (MG) estimates. Whilst the PMG constrains the long-run coefficient to be identical across countries (homogeneity), the MG allows the long-run effects to be country-specific (and reports the averaged responses). The fact that the error correction term (coefficient on the lagged dependent variable) is statistically significant provides further evidence of the existence of a long-run relationship.[161] Moreover, its magnitude for the MG (−0.6) suggests that more than half of the equilibrium error is corrected in one year, whilst for the other methods adjustment towards equilibrium appears to be slower. The short-run aid impact estimate is −0.6 for the PMG and the DFE, and −0.8 for the MG, whilst the long-run impacts vary between −0.6 and −1.0. On average, these results suggest that around two-thirds of foreign aid is absorbed in the short-run, with a modest increase in the long-run. The SYS-GMM seems to underestimate the short-term impact of aid and overestimate its long-run effect.

To test the validity of the pooling assumption and decide on the preferred specification (PMG versus MG) we undertake a Hausman test. The test assesses whether the differences in long-run coefficients are not systematic (null hypothesis), and follows a chi-square distribution with two degrees of freedom. Given that the test does not reject the null (supporting long-run homogeneity), preference should be given to the PMG since it is more efficient (less parameters to estimate). Overall, the terms of trade do not appear to be statistically significant.

The table below reports results on a potential association between foreign aid inflows and the accumulation of international reserves (DRY). Overall, there is little support for a long-run relationship between the variables. In fact, the significance of the error correction term alone (in MG) suggests that DRY is self-correcting, hence stationary, corroborating the conclusions from unit root tests. However, there is some evidence of significant short-run effects. According to the PMG and MG estimates, an increase by 1 percentage point in the aid-GDP ratio will lead to an increase in the accumulation of international reserves of around 0.3 percentage points (a minus sign indicates increase) – i.e. 30 percent of aid is used to build up international reserves. Central banks may adopt this strategy to protect their economies from future external shocks or even to smooth the availability of foreign exchange in an environment of volatile and unpredictable aid inflows. The lack of significance in the short-run coefficient in the DFE equation is probably due to the invalidity of the pooling assumption.

[161] This coincides with the coefficient on the lagged dependent variable due to normalisation.

Table 5-14: Estimation Results for DRY

	SYS-GMM DRY	DOLS DRY	DFE D.DRY	PMG D.DRY	MG D.DRY
C	−5.48	−2.91	−1.34	0.37***	−1.28
AIDBOPY	0.00	0.13			
LTOT	1.05	0.28			
DRY(−1)	0.08		−0.89***	−0.86***	−0.97***
D.AIDBOPY			−0.06	−0.27***	−0. 31***
D.LTOT			0.26	−0.64	−0.92
Cross	25	25	25	25	25
Time	25	21	25	25	25
Hausman					0.17
Long-Run					
AIDBOPY	*0.00*	0.13	0.01	−0.11***	−0.11
LTOT	*1.14*	0.28	0.18	−0.04	−0.39

Obs.: Robust standard errors. Coefficients in *italic* are calculated from the estimation output. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

We now turn to the second main empirical question. In this case, we are trying to uncover the relationship between the foreign aid inflows (AIDGOVY), inflation (INF) and the non-aid government overall balance (NAGOBY). In the short-run, the SYS-GMM, DFE and PMG estimators show that an increase by 1 percentage point in the aid-GDP ratio leads to a proportional deterioration of the non-aid government balance (full spending). Moreover, inflation has a small negative impact on the government balance.[162] In terms of its long-run impact, aid inflows contribute to a relatively large (more than proportional) widening of the public deficit (around −1.5 percent). This may be due to a possible positive correlation between aid grants (AIDGOVY) and external loans to the government (either concessional or commercial). With regard to inflation, the long-run coefficients are around −0.1. Once again, the Hausman test favours the utilisation of the PMG approach over its less restrictive alternative. The error correction term is statistically significant and relatively large, suggesting a fast adjustment to long-run equilibrium.

[162] There might be concerns of reverse causality (e.g. higher fiscal deficits causing higher inflation), since only the first two methodologies provide corrections for endogeneity. However, an increase in the non-aid fiscal deficit does not necessarily translate into an increase in money supply. It can be covered by the additional aid inflow, which appears to be the case. Moreover, the coefficients are almost identical to those from SYS-GMM.

Table 5-15: Estimation Results for NAGOBY

	SYS-GMM NAGOBY	DOLS NAGOBY	DFE D.NAGOBY	PMG D.NAGOBY	MG D.NAGOBY
C	2. 92***	1.60	−2.06***	−2.83***	−2.18***
AIDGOVY	−1.03***	−1.74***			
INF	−0.04***	−0.13***			
NAGOBY(−1)	0.35***		−0.63***	−0.61***	−0.81***
D.AIDGOVY			−1.04***	−1.03***	−1.31***
D.INF			−0.03**	−0.06**	−0.06**
Cross	25	25	25	25	25
Time	15	11	15	15	15
Hausman					0.92
Long-Run					
AIDGOVY	*−1.58****	−1.74***	−1.50***	−1.32***	−1.74***
INF	*−0.06****	−0.13***	−0.06***	−0.02*	−0.03

Obs.: Robust standard errors. Coefficients in *italic* are calculated from the estimation output. Dropping INF altogether makes the MG short-run aid coefficient fall to −1.2, and both PMG and MG long-run aid coefficients to converge around −1.55. The DFE, DOLS and SYS-GMM results are not significantly affected. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

We now assess the impact of foreign aid inflows (AIDGOVY) on public investment (INVGY). In terms of the short-run impacts, the estimation methods indicate that an increase by 1 percentage point in the aid-GDP ratio leads to an increase of about 0.4 percentage points in the public investment ratio. This impact rises up to around 0.7 percentage points in the long-run. The MG procedure fails to find a robust association between the two variables, potentially due to the fact that when T is small the lagged dependent variable bias leads to the underestimation of their true values (Pesaran et al, 1999:627). The PMG performs better because this bias is reduced by the pooling assumption, which causes an upward bias (Pesaran et al, 1999:628). Since it requires the estimation of fewer parameter coefficients, it is less onerous on the degrees of freedom – MG requires the estimation of 48 extra parameters.[163] Pesaran et al (1999:629) also note that the MG can be quite sensitive to outliers. Their impact is more severe than on the PMG, probably due to the use of unweighted averages. The extra column reports the results for the MG excluding the inflation variable, which is insignificant. The results are now in line with the other methodologies.

[163] The two long-run coefficients of the explanatory variables are now allowed to vary, i.e. (25−1)*2 = 48.

Table 5-16: Estimation Results for INVGY

	SYS-GMM INVGY	DOLS INVGY	DFE D.INVGY	PMG D.INVGY	MG D.INVGY	MG D.INVGY
C	2.12***	4.16***	1.75***	1.35***	3.10***	2.31***
AIDGOVY	0.35***	0.65***				
INF	0.01	0.03				
INVGY(−1)	0.47***		−0.39***	−0.39***	−0.57***	−0.51***
D.AIDGOVY			0.37***	0.35***	0.27***	0.36***
D.INF			0.01	0.01	0.02	
Cross	25	25	25	25	25	25
Time	15	11	15	15	15	15
Hausman					2.27	
Long-Run						
AIDGOVY	*0.66****	0.65***	0.58***	0.73***	0.17	0.60**
INF	*0.02*	0.03	0.01	0.04***	0.04	

Obs.: Robust standard errors. Coefficients in *italic* are calculated from the estimation output. Dropping INF altogether makes the PMG long-run aid coefficient increase to about 1. The results for DFE, DOLS and SYS-GMM are not significantly affected. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

Finally, we look at the potential impact of foreign aid inflows on domestic financing (BORY). The results do not seem to support a long-run relationship between the variables, since the coefficient in the DFE is barely significant and not statistically significant for the remaining regressions. However, SYS-GMM and DFE estimates suggest a short-run impact of –0.15. This indicates that a small share of aid inflows may be used to reduce domestic public debt. However, this relation may be obfuscated (in the PMG and MG) by the time aggregation of the variable. If quarterly data were available, this relationship may have been stronger, as many countries use this strategy to mitigate the impact of unpredictable aid inflows. For example, when aid flows fall below the average, governments borrow to finance planned expenditures; when aid flows are above the average, government repay the loans. Nonetheless, the yearly data does not reveal a considerable impact.

Table 5-17: Estimation Results for BORY

	SYS-GMM BORY	DOLS BORY	DFE D.BORY	PMG D.BORY	MG D.BORY
C	1.56***	0.96	1.36***	0.90***	0.32
AIDGOVY	−0.15***	0.02			
INF	−0.01	−0.01			
BORY(−1)	0.13		−0.82***	−0.84***	−0.95***
D.AIDGOVY			−0.15**	−0.23	−0.09
D.INF			−0.01	0.03	0.04
Cross	25	25	25	25	25
Time	15	11	15	15	15
Hausman					0.70
Long-Run					
AIDGOVY	*−0.17*	0.02	−0.13*	−0.02	−0.17
INF	*0.01*	−0.01	−0.01	−0.01*	0.04

Obs.: Robust standard errors. Coefficients in *italic* are calculated from the estimation output. Dropping INF altogether does not significantly affect the results (even for DFE). The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

The table below provides a summary of the impacts of foreign aid inflows on the macroeconomic and fiscal sphere. Starting with absorption, the results suggest that foreign aid inflows have had significant short- and long-run impact on the non-aid current account balance (NACABY). The short-run results for the SYS-GMM are significantly lower than the other empirical methods, possibly due to a downward bias induced by the presence of the lagged dependent variable. Overall, it seems that around two-thirds of the aid flows are used to increase the (non-aid) current account deficit, most likely through making foreign exchange available to domestic importers of goods and services. In the long-run, the impact of aid is significantly higher for the SYS-GMM, which even suggests full absorption. The PMG coefficient is lower than that of the DFE and MG, but on the whole the evidence points to a high level of absorption. With regard to the accumulation of international reserves (DRY), the fixed-effects models do not reveal a significant short-run impact, while the heterogeneous alternatives suggest that about one-third of the foreign exchange provided by aid transfers is kept as central banks' foreign reserves. In the long-run, only the PMG appears to indicate a statistically significant effect, albeit lower than the short-run impact. In light of the PMG and MG results, and bearing in mind the macroeconomic identity below, there is only weak evidence that aid flows are 'exiting' through the capital account (capital outflows), as the sum of the (short-run) impacts on DRY and NACABY is approximately 1.[164]

$$AIDBOPY = \Delta RY - (NACABY + NAKABY)$$

The implication is that short-run aid absorption in African countries is higher than previously suggested by Berg et al (2007) and Aiyar and Ruthbah (2008). Moreover, aid resources are also found to be used (in the short-run) to build up international reserves, perhaps to strengthen the capacity to weather external shocks.

[164] Note that an increase in DRY means a fall in international reserves.

Table 5-18: The Impact of Aid Inflows

	SYS-GMM	DOLS	DFE	PMG	MG[1]
Short-run					
NACABY	−0.43***	n/a	−0.63***	−0.63***	−0.76***
DRY	0.08	n/a	−0.06	−0.27***	−0. 31***
NAGOBY	−1.03***	n/a	−1.04***	−1.03***	−1.31***
INVGY	0.47***	n/a	0.37***	0.35***	0.36***
BORY	−0.15***	n/a	−0.15**	−0.23	−0.09
Long-Run					
NACABY	−1.13***	−0.86**	−0.76***	−0.62***	−0.96**
DRY	0.00	0.13	0.01	−0.11***	−0.11
NAGOBY	−1.58***	−1.74***	−1.50***	−1.32***	−1.74***
INVGY	0.66***	0.65***	0.58***	0.73***	0.60***
BORY	−0.17	0.02	−0.13*	−0.02	−0.17

Obs.: [1] The MG results for INVGY exclude inflation. The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

Turning to spending, the empirical results imply that aid inflows have had large short- and long-run impacts on the non-aid government balance (NAGOBY). In fact, aid is fully spent in the short-run. This means that the full amount of aid is used to either (i) boost public expenditures; (ii) reduce taxes; or (iii) a mixture of both. Full spending is not compatible with the hypothesis that governments 'save' aid resources to pay government debt (either domestic or foreign). In the long-run, the impact on NAGOBY grows to about 1.5, a more than proportional impact. This finding is similar to that of Aiyar and Ruthbah (2008). This may either be a symptom of aid illusion, or the consequence of a positive correlation between aid grants (AIDGOVY) and foreign loans. With regard to government investment (INVGY), about a third of aid resources are used to finance public investment programmes (in the short-run), rising to two-thirds in the long-run. However, the SYS-GMM short-run coefficient is notably higher. Finally, only the fixed-effects models uncover significant short-run impacts on domestic borrowing (BORY), even though these are comparatively smaller than those for other variables. There is little evidence supporting a long-run relationship between aid and domestic financing. These results can be analysed with the support of the following fiscal identity (budget constraint),

$$AIDGOVY = (INVGY + CY - TY) - (BORY + LY)$$

where CY stands for public recurrent spending, TY for domestic revenue, and LY for external lending (including concessional loans). Bearing in mind the caveat of potential endogeneity (although addressed by GMM and DOLS estimators), we may argue that the short-run impact on public investment is somewhere between one-third and one-half of the aid inflow, leaving about two-thirds or one-half for either increasing recurrent expenditures or lowering domestic revenues (e.g. taxes). In fact, recurrent spending is the most obvious candidate, since it often includes several development-related activities (e.g. wages of nurses, textbooks, etc.). In the long-run, the impact of aid on public investment increases to two-thirds.

Further to these economic observations, the empirical results may also provide some information about the 'small sample' behaviour of the estimators. As the temporal dimension (T) increases, the downward bias induced by the lagged dependent variable tends to decline and even OLS-FE may become consistent. However, these performance gains are likely to be higher/faster for the ECM models. This may explain why, in the macroeconomic sample, the SYS-GMM short-run coefficients are significantly lower than in the other approaches.[165] However, the SYS-GMM is likely to outperform the ECM approach in shorter panels (small T), such as the fiscal sample. In fact, the short-run coefficients are now higher than those for the PMG and MG – with the exception of NAGOBY, which are similar. Hence, it appears that the downward bias is stronger in the SYS-GMM estimator in the macroeconomic sample (T=26), whilst the bias is larger for the ECM-type models in the fiscal sample (T=16). This implies a trade-off between these different methodologies.

On the whole, our preferred model is the PMG estimator for two main reasons. Firstly, it appears that its estimates remain robust in the shorter panel, as opposed to those from the MG. This robustness may be explained by the fact that imposing parameter homogeneity often causes an upward bias (in absolute terms) in the lagged dependent variable (Pesaran et al, 1999:628). Hence, the potential downward bias induced by small T may actually be reduced or even cancelled out. It can be seen that, in general, the absolute magnitude of the estimated error correction coefficients follow the sequence MG>PMG>DFE. Moreover, the MG is also more sensitive to outliers. Secondly, the PMG assumptions are more appealing in economic terms. We allow heterogeneity in the short-run responses and the speed of adjustment to equilibrium, while constraining the long-run relationships to be the same. This is an appealing middle ground between the strong pooling assumptions of the DFE (and indeed GMM) estimator and the flexibility of the MG estimator.

However, we should bear in mind the weaknesses of the empirical analysis presented here. The results from our estimation strategies (including the system GMM) may be sensitive to the presence of cross-sectional error dependence. If T was significantly larger

---

165 We have also estimated OLS-FE and DIF-GMM for the lagged dependent variable model, and the results indicate that these estimators tend to underestimate both short- and long-run impacts in relation to SYS-GMM.

than N (which unfortunately is not our case) we could model and test the cross-correlation of the error terms through seemingly unrelated regressions (SUR). Nonetheless, we can test the assumption of cross-sectional independence with the Breusch-Pagan LM test statistic and Pesaran's (2004) CD statistic. The LM test follows a chi-square distribution with $N(N-1)/2$ degrees of freedom but requires $T>N$, whilst the Pesaran test is asymptotically normal. For both tests we use the DFE specification presented above. Although the results from the LM statistic strongly reject the hypothesis of cross-sectional independence, the CD statistic provides much weaker evidence of violations. We recall that the LM statistic relies on large T and small N and thus may not perform well when both dimensions are of similar magnitudes. In fact, the computations were just about possible for the macroeconomic sample, where T=26 and N=25. Thus, we may argue that there is only weak evidence that the assumption of cross-sectional dependence of the error structures is violated by the data, and thus the empirical estimates are not likely to be significantly biased. Furthermore, panel estimators that are robust to cross-sectional dependence are only at an embryonic stage – e.g. see recent literature on 'common correlated effects' estimator (Kapetanios et al, 2009).

Table 5-19: Cross-Sectional Independence Tests

DFE model	NACABY AIDBOPY LTOT	DRY AIDBOPY LTOT	NAGOBY AIDGOVY INF	INVGY AIDGOVY INF	BORY AIDGOVY INF
Breusch-Pagan LM statistic	395.88***	366.76***	n/a	n/a	n/a
Pesaran CD statistic	1.83*	2.24**	1.69*	–0.89	–0.26

Obs.: Tests results generated by the Stata's 'xttest2' and 'xtcsd' modules. CD test results based on DOLS regressions do not reject independence, except for the INVGY equation at 10 percent, whilst the LM statistic cannot be computed (even for the macroeconomic sample). The asterisks represent significance at the 10 percent (*), 5 percent (**), and 1 percent (***) confidence levels.

Finally, the estimates only represent country averages. Policy responses may vary from country to country, and therefore this analysis does not preclude the use of other methodologies to unveil country-specific macroeconomic responses.

5.6 Conclusion

This paper revisits the issue of macroeconomic management of large aid disbursements. We have constructed a new panel dataset to investigate the level of aid 'absorption' and 'spending' in Africa's low-income countries. Our results suggest that, in the short-run, recipient countries have absorbed about two-thirds of the aid inflow, using them to increase the non-aid current account deficit. Moreover, around one-third of the foreign exchange provided by these inflows has been used to build up international reserves, perhaps to protect economies from future external shocks.[166] In the long-run, absorption of foreign exchange appears to increase further without reaching its maximum (full absorption). In terms of 'aid spending', recipient countries appear to have fully spent the amount of aid, using it to increase the non-aid government deficit. In particular, a substantial percentage of these inflows went to finance public investment expenditures. There is only weak evidence that some aid flows have been 'saved', i.e. used to substitute for domestic borrowing. Overall, these findings suggest that the macroeconomic management of aid inflows in Africa has been better than often suggested by comparable exercises.

These results challenge some of the conclusions from Aiyar and Ruthbah (2008), namely that short-run absorption is usually low, with aid exiting through the capital account. This may be due to the use of an inappropriate measure of aid flows (DAC's donor reported aid) or the application of a methodology that neglects the time series properties of the data. However, we corroborate the result that spending is higher than absorption, which represents an injection of domestic liquidity in the recipient country.

5.7 Bibliography

Agénor, P. and P. Montiel (1999), Development Macroeconomics (Second Edition), Princeton University Press, Chichester, West Sussex.

Aiyar, S. and U. Ruthbah (2008), 'Where Did All the Aid Go? An Empirical Analysis of Absorption and Spending.' IMF Working Paper 08/34, Washington, D.C.

Arellano, M. and S. Bond (1991), 'Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations.' *Review of Economic Studies*, 58:277-297.

Asteriou, D. and S. Hall (2007), Applied Econometrics: A Modern Approach (Revised Edition), Palgrave McMillan, Basingstoke.

Baltagi, B. (2008), Econometric Analysis of Panel Data (Fourth Edition), John Wiley & Sons, Chichester.

Banerjee, A., M. Marcellino and C. Osbat (2005), 'Testing for PPP: Should We Use Panel Methods?' *Empirical Economics*, 30:77–91.

Banerjee, A., M. Marcellino and C. Osbat (2004), 'Some Cautions on the Use of Panel Methods for Integrated Series of Macroeconomic Data,' *Econometrics Journal*, 7:322-340.

[166] Buffie et al (2004) suggest that a 'managed float' is the most attractive approach to manage shocks to aid inflows, therefore arguing that African central banks have been correct to intervene in the foreign exchange market.

Banerjee, A., J. Dolado and R. Mestre (1998), 'Error-Correction Mechanism Tests for Cointegration in a Single-Equation Framework,' *Journal of Time Series Analysis*, 19:267–83.

Berg, A., S. Aiyar, M. Hussain, S. Roache, T. Mirzoev and A. Mahone (2007), 'The Macroeconomics of Scaling Up Aid: Lessons from Recent Experience,' IMF Occasional Paper 253.

Breitung, J. (2000), 'The Local Power of Some Unit Root Tests for Panel Data.' in B. Baltagi (ed.), Advances in Econometrics, Vol. 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels, 161–178, Elsevier Science.

Buffie, E., C. Adam, S. O'Connell and C. Pattillo (2004), 'Exchange Rate Policy and the Management of Official and Private Capital Flows in Africa.' *IMF Staff Papers*, 51, Special Issue, 126-160.

Calvo, G. and C. Reinhart (2002), 'Fear of Floating.' *Quarterly Journal of Economics*, 117(2):379-408.

Choi, I. (2001), 'Unit Root Tests for Panel Data.' *Journal of International Money and Finance*, 20(2):249–272.

Engle, R. and C. Granger (1987), "Co-integration and Error-Correction: Representation, Estimation and Testing." *Econometrica*, 55, 251-276.

Foster, M. and T. Killick (2006), 'What Would Doubling Aid do for Macroeconomic Management in Africa,' Overseas Development Institute (ODI) Working Paper 264.

Greene, W. (2003), Econometric Analysis (Fifth Edition), Prentice Hall, New Jersey.

Gupta, S., R. Powell and Y. Yang (2006), Macroeconomic Challenges of Scaling Up Aid to Africa: A Checklist for Practitioners. International Monetary Fund, Washington, D.C.

Heller, P. (2005), 'Pity the Finance Minister: Issues in Managing a Substantial Scaling Up of Aid Flows,' IMF Working Paper 05/180.

Hadri, K. (2000), 'Testing for Stationarity in Heterogeneous Panel Data.' *Econometrics Journal*, 3:148–161.

Hussain, M., A. Berg and S. Aiyar (2009), 'The Macroeconomic Management of Increased Aid: Policy Lessons from Recent Experience.' *Review of Development Economics*, 13(3), 491-509.

Im, K., H. Pesaran and Y. Shin (2003), 'Testing for Unit Roots in Heterogeneous Panels.' *Journal of Econometrics*, 115:53–74.

IMF (2005), 'The Macroeconomics of Managing Increased Aid Inflows: Experiences of Low-Income Countries and Policy Implications,' SM/05/306 (August 8, 2005), unpublished, Policy Development and Review Department

Isard, P., L. Lipschitz, A. Mourmouras, and B. Yontcheva, eds. (2006), The Macroeconomic Management of Foreign Aid: Opportunities and Pitfalls, Washington: International Monetary Fund.

Johansen, S. (1988), "Statistical Analysis of Cointegration Vectors." *Journal of Economic Dynamics and Control*, 12(2-3), 231-254

Kao, C. and M. Chiang (2000), 'On the Estimation and Inference of a Cointegrated Regression in Panel Data.' In B. Baltagi (ed.), Advances in Econometrics, Vol. 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels, 179-222, Elsevier Science.

Kao, C. (1999), 'Spurious Regression and Residual-Based Tests for Cointegration in Panel Data.' *Journal of Econometrics*, 90:1–44.

Kapetanios, G., H. Pesaran and T. Yamagata (2009), 'Panels with Nonstationary Multifactor Error Structures.' Cambridge Working Papers in Economics, 0651 (Revised), University of Cambridge.

Larsson, R., J. Lyhagen and M. Lothgren (2001), 'Likelihood-Based Cointegration Tests in Heterogeneous Panels.' *Econometrics Journal*, 4:109–142.

Levin, A., C. Lin and C. Chu (2002), 'Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties.' *Journal of Econometrics*, 108:1–24.

Maddala, G. and S. Wu (1999). 'A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test.' *Oxford Bulletin of Economics and Statistics*, 61:631–52.

McKinley, T. (2005), "Why is the Dutch Disease Always a Disease? The Macroeconomic Consequences of Scaling Up ODA." UNDP, International Poverty Centre, Working Paper 10.

Pedroni, P. (2004), 'Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis.' *Econometric Theory*, 20, 597-625.

Pedroni, P. (2000), 'Fully Modified OLS for Heterogeneous Cointegrated Panels.' In B. Baltagi (ed.), Advances in Econometrics, Vol. 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels, 93-130, Elsevier Science.

Pedroni, P. (1999), 'Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors.' *Oxford Bulletin of Economics and Statistics*, 61:653–70.

Pesaran, H. (2007), 'A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence.' *Journal of Applied Econometrics,* 22(2), 265-312.

Pesaran, H. (2004), 'General Diagnostic Tests for Cross Section Dependence in Panels.' Cambridge Working Papers in Economics 0435, Faculty of Economics, University of Cambridge.

Pesaran, H., Y. Shin and R. Smith (1999), 'Pooled Mean Group Estimation of Dynamic Heterogeneous Panels.' *Journal of the American Statistical Association*, 94(446):621–634.

Phillips, P. and H. Moon (1999), 'Linear Regression Limit Theory for Nonstationary Panel Data.' *Econometrica*, 67(5):1057-1111.

Phillips, P. and B. Hansen (1990), 'Statistical Inference in Instrumental Variables Regression with I(1) Processes.' *The Review of Economic Studies*, 57(1), 99-125.

Prati, A., R. Sahay and T. Tressel (2003), 'Is There a Case for Sterilizing Foreign Aid Inflows?' Paper presented at the Research Workshop 'Macroeconomic Challenges in Low Income Countries', October 23-24, Washington D.C.: IMF.

QMS (2007), EViews 6 User's Guide II, Quantitative Micro Software, LLC.

Roodman, D. (2009), 'A Note on the Theme of Too Many Instruments.' *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158.

Saikkonen, P. (1991), 'Asymptotically Efficient Estimation of Cointegration Regressions.' *Econometric Theory*, 7(1), 1-21.

Stock, J. and M. Watson (1993), 'A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems.' *Econometrica*, 61(4), 783-820.

Westerlund, J. (2007), 'Testing for Error Correction in Panel Data.' *Oxford Bulletin of Economics and Statistics*, 69(6), 709-748.

White, H. (2005), 'The Case for Doubling Aid.' *IDS Bulletin*, 36(3), 8-13.

White, H. (1992), 'The Macroeconomic Impact of Development Aid: A Critical Survey.' *Journal of Development Studies*, 28(2), 163-240.

6. Conclusion

This thesis makes an empirical contribution to the literature on the macroeconomics of foreign aid flows. Through conducting a systematic investigation of how fiscal behaviour and the real exchange rate mediate the impact of aid on economic growth, this thesis provides the following contributions: (i) new insights into the fiscal dynamics of aid; (ii) an empirical investigation of Dutch disease symptoms in one of the largest aid-recipient countries in Africa; and (iii) a re-assessment of the macroeconomic responses to aid inflows for a panel of African countries.

The first two papers (chapters 2 and 3) focus on the fiscal sector, whereas the third paper (chapter 4) analyses the potential effects on the real exchange rate. These three empirical assessments use data for Ethiopia, one of the most aid-dependent countries in Africa. The last paper (chapter 5) evaluates the fiscal and macroeconomic management of aid flows for a panel of 25 African countries. This section synthesises the main conclusions and policy implications of the research.

The first two papers make contributions to the fiscal response literature. The first paper uses annual fiscal data covering the period 1964-2005 to investigate the impact of foreign aid on the government sector. It uses a traditional fiscal response model where aid inflows are disaggregated into grants and loans. This is, to our knowledge, the first fiscal response study focusing on Ethiopia, and considerable attention was devoted to the collection of a consistent fiscal dataset. Moreover, it undertakes single-equation cointegration analysis to provide a further robustness check on the empirical results. Overall, the results suggest that foreign aid is positively correlated with public spending, especially investment expenditures. This is a desirable effect, since Ethiopia requires a significant boost in investment in order to foster economic growth and alleviate poverty. Aid is found to be negatively correlated with domestic borrowing, perhaps due to the volatility and unpredictability of aid disbursements. In fact, domestic borrowing appears to compensate aid shortfalls in order to smooth the expenditure pattern. There is only weak evidence that aid flows cause tax displacement. Finally, there are signs of aid heterogeneity, since grants and concessional loans tend to induce different impacts.

The second paper strengthens the fiscal analysis conducted in the first paper in two main ways. Firstly, it follows a recent trend in the literature by investigating fiscal relationships in a cointegrated vector autoregressive (CVAR) model. However, it improves on existing studies through the formulation and testing of interesting hypotheses about fiscal behaviour, as well as placing a strong focus on model specification and robustness checks. Secondly, the paper provides, to our knowledge, the first fiscal response analysis of a quarterly fiscal dataset. The advantages of this approach are: (i) allowing for important intra-year dynamics; (ii) providing a larger sample size; and (iii) minimising the likelihood of structural breaks, which often arise from economic reforms undertaken in the late 1980s and early 1990s in many African countries. The econometric results suggest the presence of three (empirical) long-run relations: (i) the government budget constraint; (ii) a possible donor disbursement rule; and (iii) a financing (trade-off) rule. While domestic fiscal variables adjust to budget imbalances, foreign aid grants seem to adjust to the level of development spending, which can be seen as an indication of (procyclical) aid conditionality. Since 'earmarked' grants have a contemporaneous impact, this finding implies that budget support grants are provided after the government makes a strong commitment to poverty reduction. Moreover, domestic borrowing often compensates for lower levels of revenue and grants, highlighting the cost of aid unpredictability and revenue volatility. The moving average representation of the CVAR suggests that unanticipated shocks to foreign aid grants do not have permanent effects on the remaining fiscal variables. There is also evidence of aid heterogeneity, since aid grants and loans entail different dynamics.

The empirical results obtained in the first two papers are consistent and highlight some interesting policy implications. One of the main (common) findings is the strong negative correlation between aid and domestic borrowing, suggesting that aid finances a substantial share of public spending indirectly. This may have negative implications on the fiscal accounts if interest payments on domestic debt are considerably high. This could be seen as the cost of unpredictable and volatile aid inflows. Hence, efforts to make aid more predictable are likely to improve medium-term fiscal planning. This would also ensure an improvement in the 'additionality' of aid resources. Ideally, foreign aid should directly finance new capital spending (and associated recurrent developmental expenditures) that contribute to enhancing human development and expanding the economy's productive capacity. Finally, we argue that aid should also be targeted to building up national capacities to mobilise domestic revenue. This would ensure a gradual shift from aid-dependence to a stronger reliance on domestic sources of finance.

The third paper investigates the second main channel through which aid influences economic performance by providing an evaluation of the determinants of the Ethiopian real exchange rate (RER). The main focus of the empirical analysis is, therefore, to assess whether foreign aid inflows are associated with appreciations of the RER – Dutch disease hypothesis. For that purpose, we compile a quarterly macroeconomic dataset (1981-2008) and use several single-equation cointegration frameworks. Moreover, we also use an unobserved components (UC) model as an alternative empirical framework. The results emerging from both the econometric and UC models for the period 1995-2008 suggest that trade openness tends to depreciate the RER, whilst external terms of trade shocks cause an appreciation. With regard to capital inflows, neither foreign aid nor workers' remittances were found to be statistically significant, although the latter might be weakly associated with appreciation pressures. This may be due to the fact that a larger share of private transfers is spent on domestic products. Hence, the paper argues that the main fluctuations of the Ethiopian RER can be accounted by three main factors: (i) external commodity price shocks (measured by the terms of trade), mainly affecting coffee exports and oil imports; (ii) political events, such as the instability towards the end of the Derg regime that caused unusually high inflation; and (iii) economic policy, especially the liberalisation of trade flows and the exchange rate market. Moreover, the lack of significance of variables such as excess money supply and government consumption may suggest that Ethiopia has pursued sound macroeconomic policies since the fall of the Derg regime.

The results from the third paper demonstrate that, in the case of Ethiopia, foreign aid inflows are not associated with RER appreciations as standard theoretical models predict. This might be due to resources being channelled to improve the productive capacity of the economy. These supply-side spillovers may indeed provide a justification as to why it has been difficult to empirically verify these theoretical predictions. Moreover, while most studies in this literature either find that aid induces appreciation or depreciation of the RER, we fail to find a robust association between the two variables. This suggests that the positive productivity effects of aid cancel out any pressures arising from the additional demand for domestic goods.

Finally, the last paper (chapter 5) investigates the fiscal and macroeconomic management of aid flows for a panel of 25 African low-income countries. It complements the other three papers by analysing whether the government and central bank policy responses have played a role in the management of aid inflows. We compile a new (balanced) fiscal and macroeconomic panel dataset to investigate the extent to which aid inflows contribute to a widening of the non-aid current account deficit ('absorption') and the non-aid fiscal deficit ('spending'). The main purpose of using these concepts is to analyse the level of coordination between recipient governments and central banks. We place a strong focus on understanding the statistical properties of the data (e.g. non-stationarity) and therefore use several panel cointegration methods. The evidence suggests that foreign aid inflows are significantly 'absorbed' in the short-run (about two-thirds), which means that most aid resources are utilised to increase the (non-aid) current account deficit (e.g. boost imports). Absorption increases in the long-run, but remains below full absorption. Aid resources are also used to increase international reserves, but only in the short-run. Moreover, aid flows are fully 'spent', in the sense that the flows generate (at least) a proportional increase in the (non-aid) fiscal deficit. A significant share is found to be devoted to support public investment. Finally, we did not find a significant impact on domestic borrowing. This may be due to the quality of the variable (which often includes 'net errors and omissions'), the lack of a simultaneous fiscal decision framework, or simply because the time aggregation (annual data) does not allow us to uncover the effects found for Ethiopia.

The results from the final paper suggest that African low-income countries have, on average, pursued coordinated fiscal and macroeconomic responses to aid inflows. This finding contrasts with previous studies that have argued that African countries have not coordinated fiscal policies with monetary and exchange rate polices. The fact that 'spending' tends to be higher than (external) 'absorption' suggests that aid inflows provide an injection of domestic liquidity in the recipient country. Although this may raise fears of inflationary pressures, it may be a necessary approach in the face of unpredictable and volatile aid inflows.

Overall, this thesis provides significant contributions to the literature on the fiscal and macroeconomic impact of aid inflows. In the case of Ethiopia, we demonstrate that foreign aid inflows may create management problems, illustrated by the relationship between aid and domestic borrowing. We argue that while it is desirable that recipient governments smooth their expenditure patterns, significant efficiency gains could be achieved if donors were to improve the predictability of their aid disbursements. Nonetheless, we did not find evidence of tax displacement and Dutch disease symptoms in Ethiopia, despite being one of the largest aid recipients in the world. This will possibly ease concerns about the potential for foreign aid transfers to undermine economic growth. Finally, we found that policy responses by African government and central banks have been adequate in the face of the recent increase in aid inflows. This analysis could provide donors and recipient governments with the empirical evidence to design better aid interventions, and effectively manage large inflows of foreign finance while minimising the potential negative impacts.