CREDIT Research Paper

No. 10/05

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

by

Pedro M. G. Martins

Abstract

This paper uses the cointegrated vector autoregressive (CVAR) model to assess the dynamic relationship between foreign aid inflows, public expenditure, revenue and domestic borrowing in Ethiopia. It departs from the existing literature by using a unique quarterly fiscal dataset (1993-2008) and providing new insights into the formulation of testable fiscal hypotheses. The paper also derives and interprets structural shocks and places a strong focus on model specification. The results suggest the presence of three long-run relationships: the government budget constraint, a donor disbursement rule, and a financing trade-off. Foreign aid grants adjust to the level of development spending, which can be seen as an indication of (procyclical) aid conditionality. Moreover, domestic borrowing often compensates for lower levels of revenue and grants, highlighting the cost of aid unpredictability and revenue volatility. 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.

JEL Classification: C32, F35, O23, O55
Keywords: Fiscal Response, Foreign Aid, Time Series Models, Africa

---

Centre for Research in Economic Development and International Trade, University of Nottingham

CREDIT Research Paper

No. 10/05

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

by

Pedro M. G. Martins

Outline

1. Introduction
2. Literature Review
3. Methodology
4. Data
5. Estimation
6. Structural Analysis
7. Conclusion

The Authors

Pedro M. G. Martins was a DPhil student at the Institute of Development Studies (IDS), University of Sussex, UK. Email: pedromgmartins@gmail.com

Acknowledgements

This paper is based on the author's DPhil research, which was financially supported by the *Fundação para a Ciência e a Tecnologia*. The author would also like to thank the comments and suggestions from the participants of the 2008 Summer School in Econometrics (CVAR) in Copenhagen, the attendants of an IDS-Sussex seminar in 2009, and the participants of the 15th Annual Conference on Econometric Modelling for Africa in 2010.

Research Papers at www.nottingham.ac.uk/economics/credit/

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. This is the initial and probably the most important transmission channel through which aid impacts the recipient economy. Hence, 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.

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 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. 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:

a) 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 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.

b) 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).

c) 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.

d) 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.

e) 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.

f) Spending. This concept is defined by Hussain et al (2009) 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.

g) 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 seven 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 estimates the model, whereas section six addresses the identification issue and performs impulse response functions. Section seven concludes the paper.

2. Literature Review

There is a significant empirical literature on the fiscal impacts of aid. McGillivray and Morrissey (2001a) provide an excellent review of this literature, distinguishing between 'categorical fungibility' and 'fiscal response' studies. The distinctive feature of categorical fungibility studies is that these 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. A set of reduced form equations is then derived from the maximisation problem and estimated simultaneously, often by three-stage least squares (3SLS).

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.

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 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. 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 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 usually requires that we exclude one fiscal component from the system of equations (to avoid estimating an identity), with potential loss of information.[1]

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

[1] For example, Osei et al (2005) exclude non-tax revenue and external borrowing from the estimation.

to) 1.[2] 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).

Table 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).[3] 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

[2] 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.

[3] The VAR methodology does not provide numerical results that are directly comparable with traditional fiscal response models. 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.

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 moderately 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: 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 vary or simply because governments behave differently (in terms of policy-making).

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.[4]

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, …, T; Π_i are $p \times p$ matrices of parameters (i.e. coefficients to be estimated) with i = 1, 2, …, 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): $\varepsilon_t \sim \text{iid } N_p(0,\Omega)$. The residual covariance matrix (Ω) has dimensions $p \times p$, and contains

[4] 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.

information about possible contemporaneous effects. 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. 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:[5]

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

where,

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

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

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 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 $|A(1)| = |\Pi| = 0$.[6] 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 β.[7] 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

[6] This means that Π is a singular matrix (i.e. non-invertible).
[7] 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).

= 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. 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.

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,[8] 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

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

the common stochastic trends, which are responsible for the non-stationarity of the process. The MA representation of the VAR is given by:

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

where,

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

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 structural VAR.

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):

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

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: $\mathrm{a} = \alpha A_0^{-1}$, $A_j = \Gamma_j A_0^{-1}$, $\widetilde{\Phi} = \Phi A_0^{-1}$, 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).

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.

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.[9] 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 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,[10] 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

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

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 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, 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.

There can be substantial discrepancies between data sources. OECD-DAC grant figures for Ethiopia are, 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, some 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 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 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.

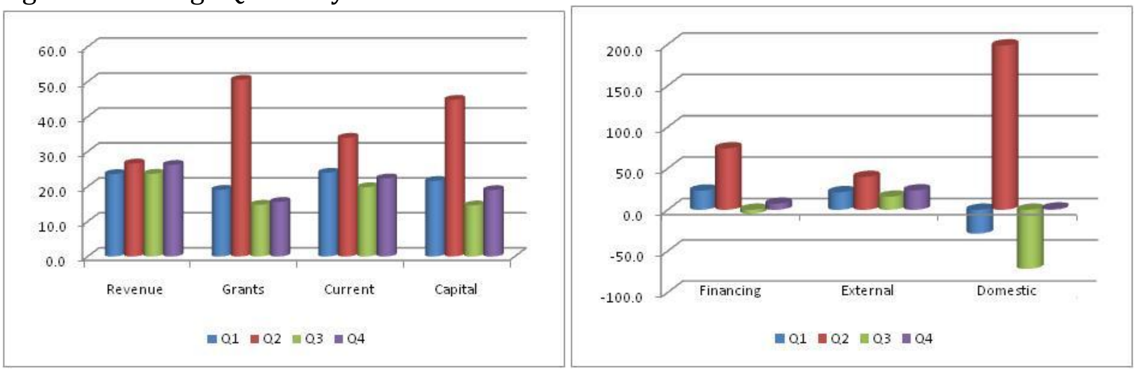
Table 5: Summary of Fiscal Variables (% GDP)

	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 previous table 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 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.[11]

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 recurrent 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.

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.

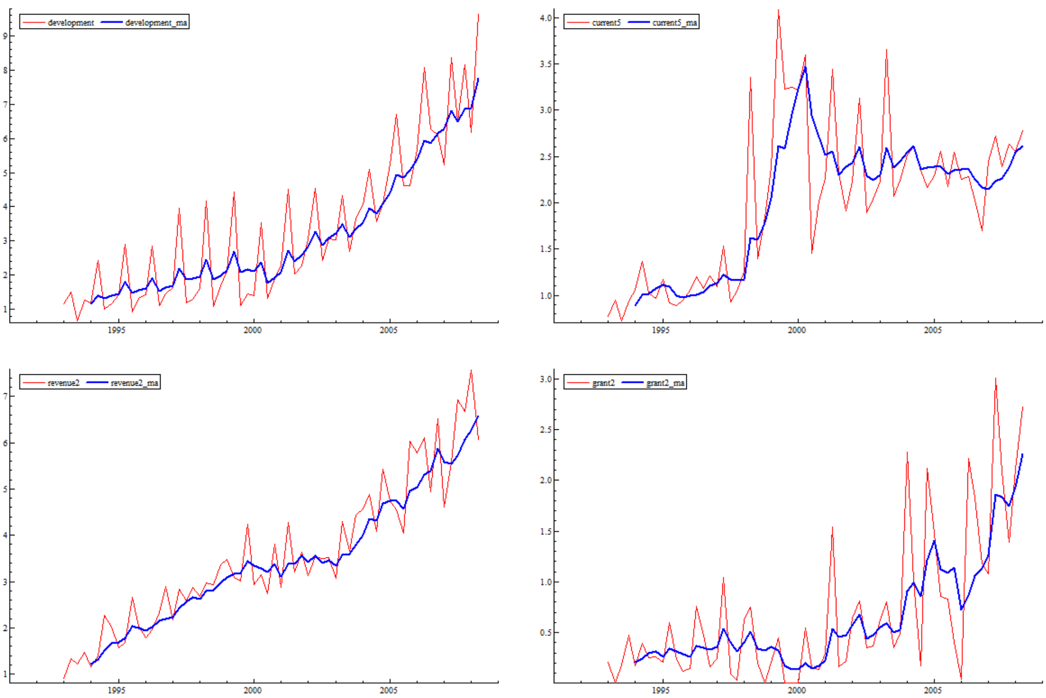
[11] A similar approach can be found in Osei et al (2005). Taking logarithms of the variables is not appropriate, since it would violate the budget identity.

> "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)

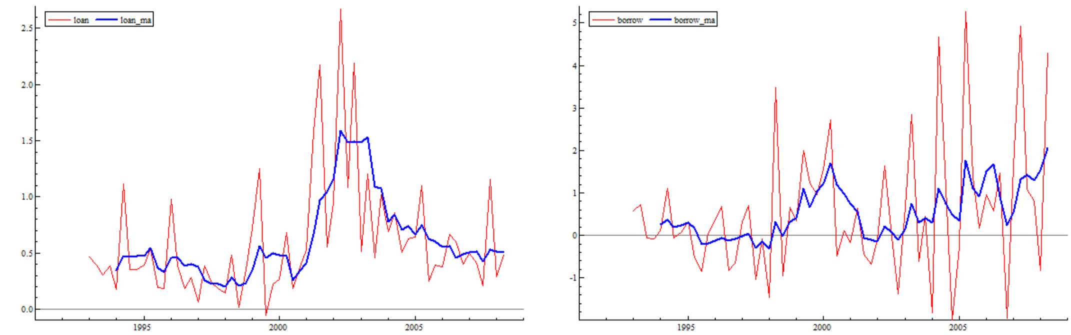
Therefore, Current will include defence related expenditures, interest payments, and 'other' recurrent spending.

The following plots show the fiscal variables in real terms[12] 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.

Figure 2: Variable Plots and 4-Period Moving Average



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



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.[13] 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.

[13] "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)

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.

5. Estimation

5.1 Unit Root Tests and Seasonality

We start by undertaking 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.[14] 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.

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

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

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).[15] 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 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). Moreover, F-tests may also be used to test the joint null hypothesis H_0: $\pi_2=\pi_3=\pi_4=0$, 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).

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

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 6: Seasonal Unit Root Tests (Levels and Regular Differences)

Var.	Lags	H_0	Test	Stat	Var.	Lags	H_0	Test	Stat
DEV	5	$\pi_1=0$	$t_{\pi1}$	0.32	ΔDEV	7	$\pi_1=0$	$t_{\pi1}$	−0.87
		$\pi_2=0$	$t_{\pi2}$	−1.29			$\pi_2=0$	$t_{\pi2}$	−1.83
		$\pi_3=\pi_4=0$	F_{34}	5.07			$\pi_3=\pi_4=0$	F_{34}	3.44
	0	$\pi_1=0$	$t_{\pi1}$	0.25		0	$\pi_1=0$	$t_{\pi1}$	-4.60^{***}
		$\pi_2=0$	$t_{\pi2}$	−0.27			$\pi_2=0$	$t_{\pi2}$	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	$\pi_1=0$	$t_{\pi1}$	−1.67	ΔCUR5	0	$\pi_1=0$	$t_{\pi1}$	-4.17^{***}
		$\pi_2=0$	$t_{\pi2}$	-2.90^{**}			$\pi_2=0$	$t_{\pi2}$	−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	$\pi_1=0$	$t_{\pi1}$	−0.09	ΔREV2	2	$\pi_1=0$	$t_{\pi1}$	-5.33^{***}
		$\pi_2=0$	$t_{\pi2}$	−2.39			$\pi_2=0$	$t_{\pi2}$	-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	$\pi_1=0$	$t_{\pi1}$	−1.60					
		$\pi_2=0$	$t_{\pi2}$	−2.24					
		$\pi_3=\pi_4=0$	F_{34}	9.88^{***}					
		$\pi_2=\pi_3=\pi_4=0$	F_{234}	8.18^{***}					
GR2	9	$\pi_1=0$	$t_{\pi1}$	0.09	ΔGR2	7	$\pi_1=0$	$t_{\pi1}$	−1.83
		$\pi_2=0$	$t_{\pi2}$	-3.19^{**}			$\pi_2=0$	$t_{\pi2}$	-2.88^{**}
		$\pi_3=\pi_4=0$	F_{34}	6.19^{*}			$\pi_3=\pi_4=0$	F_{34}	12.48^{***}
	0	$\pi_1=0$	$t_{\pi1}$	−1.65		0	$\pi_1=0$	$t_{\pi1}$	-4.94^{***}
		$\pi_2=0$	$t_{\pi2}$	-5.10^{***}			$\pi_2=0$	$t_{\pi2}$	-4.26^{***}
		$\pi_3=\pi_4=0$	F_{34}	11.55^{***}			$\pi_3=\pi_4=0$	F_{34}	10.12^{***}
LOA	0	$\pi_1=0$	$t_{\pi1}$	−1.90	ΔLOA	0	$\pi_1=0$	$t_{\pi1}$	-5.73^{***}
		$\pi_2=0$	$t_{\pi2}$	-3.44^{***}			$\pi_2=0$	$t_{\pi2}$	-3.70^{***}
		$\pi_3=\pi_4=0$	F_{34}	11.14^{***}			$\pi_3=\pi_4=0$	F_{34}	11.66^{***}
BOR	0	$\pi_1=0$	$t_{\pi1}$	-3.59^{**}	ΔBOR	0	$\pi_1=0$	$t_{\pi1}$	-7.55^{***}
		$\pi_2=0$	$t_{\pi2}$	-3.02^{**}			$\pi_2=0$	$t_{\pi2}$	-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.

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.[16] 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).

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 7: 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.

[16] "[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)

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,[17] 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.[18] 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.[19] 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

[17] "Including unmodeled stochastic variables may be problematic for inference and analysis purposes unless the variables satisfy exogeneity requirements" (Lütkepohl and Krätzig, 2004:67).
[18] 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).
[19] There is some evidence that the variable is weakly exogenous.

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 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.[20] The estimates of the new model are presented in the Appendix.

5.3 Misspecification Tests

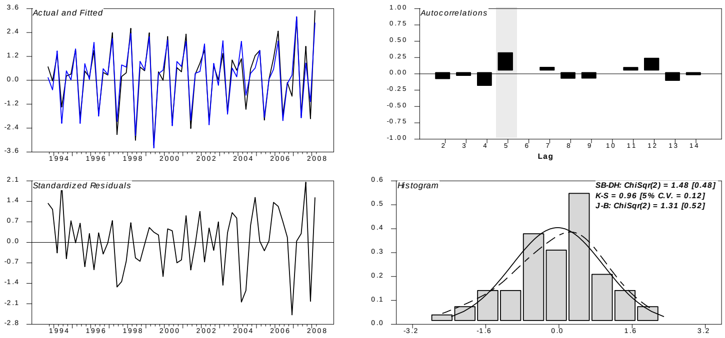
In this section we report several formal misspecifications tests, most of which focus on the system residuals. These tests are important to assess the validity of the assumptions underlying statistical model (VAR). It is useful to start with a graphical inspection of the residuals, since it can help 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

[20] 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.

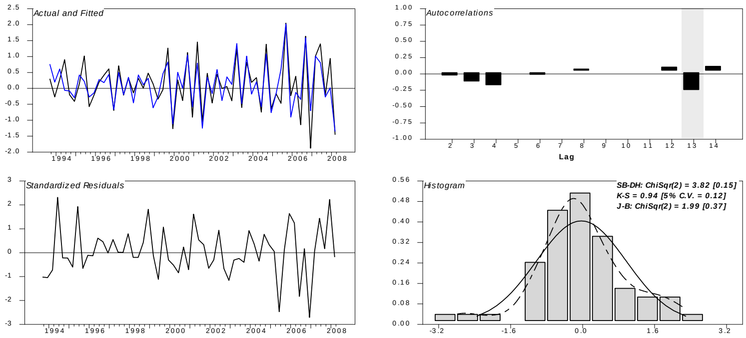
(bottom right panel); and (d) the standardised residuals (bottom left panel). The graphs do not suggest any particular problems, especially after the outliers were corrected.

Figure 3: Fitted Values, Autocorrelograms, Distributions and Residuals

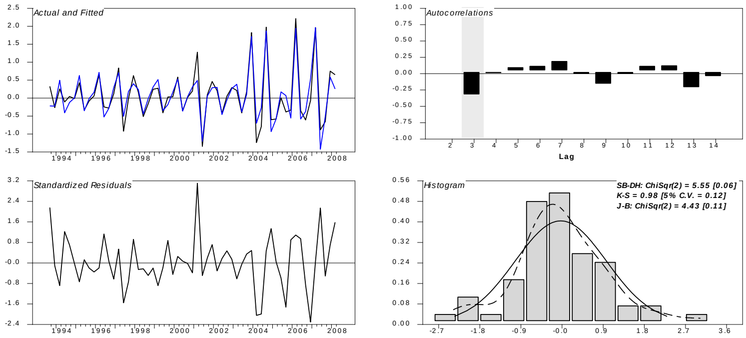
DDEVELOPMENT

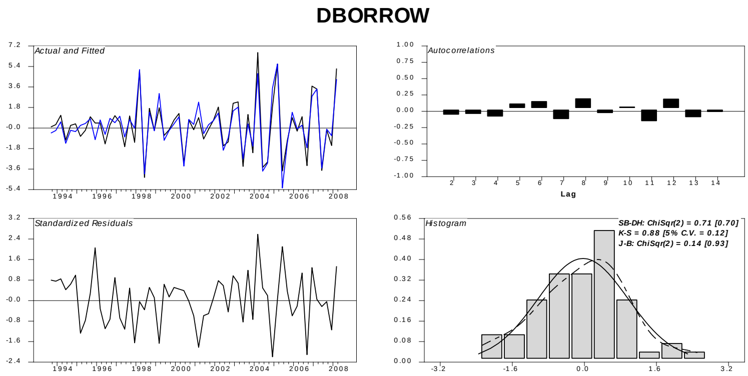


DREVENUE2



DGRANT2





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,[21] 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.[22] 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 serve this purpose, as they apply a penalising factor related to the number of estimated parameters (as a result of increasing the lag length).

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

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

---

[21] 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.

[22] The null hypothesis (H_k) suggests that the VAR model does not have significant coefficients at lag k+1.

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.[23]

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.[24]

Table 8: 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 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).

[23] 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.

[24] 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.

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 9: 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 10: 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). The test does not reject the hypothesis of multivariate normality ($\chi^2(8)=10.289$). 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 11: 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 12: 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.[25] 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.

5.4 Determination of the Cointegration Rank

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.[26] 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.[27] 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. 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-

[25] This value is computed as $X = \Phi^{-1}(1 - 0.025)^{1/T}$, where Φ is the cumulative normal distribution function.
[26] It is also known as the Johansen test or the LR test for cointegrating rank.
[27] Deterministic components in the cointegrating vector will influence the distribution of the test.

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.

Table 13: Trace Test

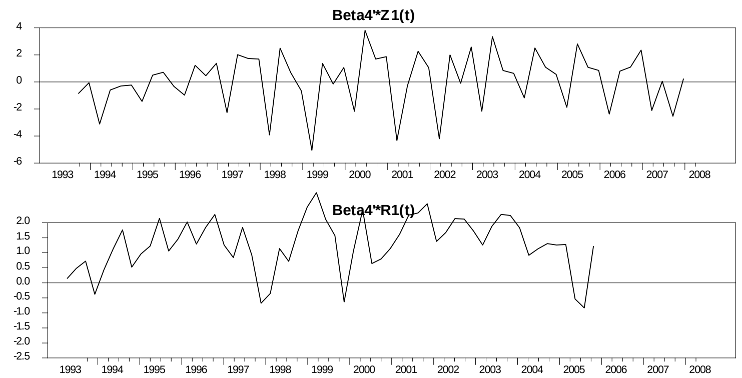
p–r	r	Eigenvalue	Trace	Trace*	Frac95	P-Value	P-Value*
4	0	0.825	256.104	247.230	82.501	0.000	0.000
3	1	0.709	153.342	149.598	57.316	0.000	0.000
2	2	0.672	80.591	79.418	35.956	0.000	0.000
1	3	0.221	14.743	14.669	18.155	0.143	0.147

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 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 5: Roots of the Companion Matrix

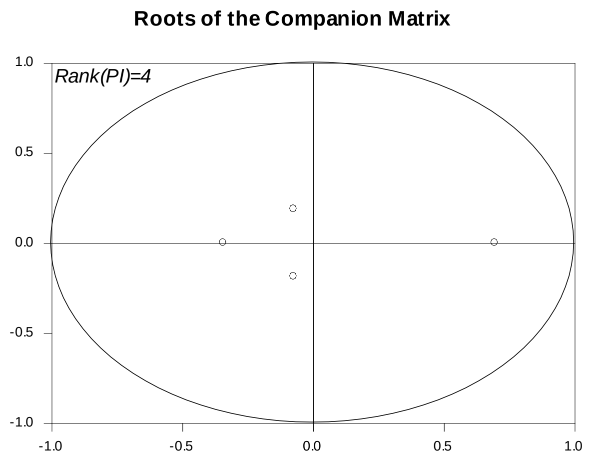


Table 14: Roots of the Companion Matrix

	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 Appendix). 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 in the Appendix.

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.[28] 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.

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.

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

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 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 15: 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). These tests are based on the fiscal concepts presented earlier. 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.[29] 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.

[29] 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 16: 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 (non-development) 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,[30] 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. 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.

Table 17: 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

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

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. 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 18: (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 19: 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.[31]

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'.

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.[32] Hence, there is a strong possibility that the residuals will be correlated, even though this study uses a quarterly dataset.[33] Since a straightforward identification of the

[31] 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.
[32] The short-run is generically identified by the zero restrictions on the contemporaneous matrix A_0. The challenge is to recover the structural parameters.
[33] 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.

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.

Table 20: 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 21: 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.[34] Finally, the fourth strategy is redundant for our case, since we already have a partial model but still have significant residual correlations.

[34] 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.

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.

Table 22: MA Representation and Decomposition of the Trend

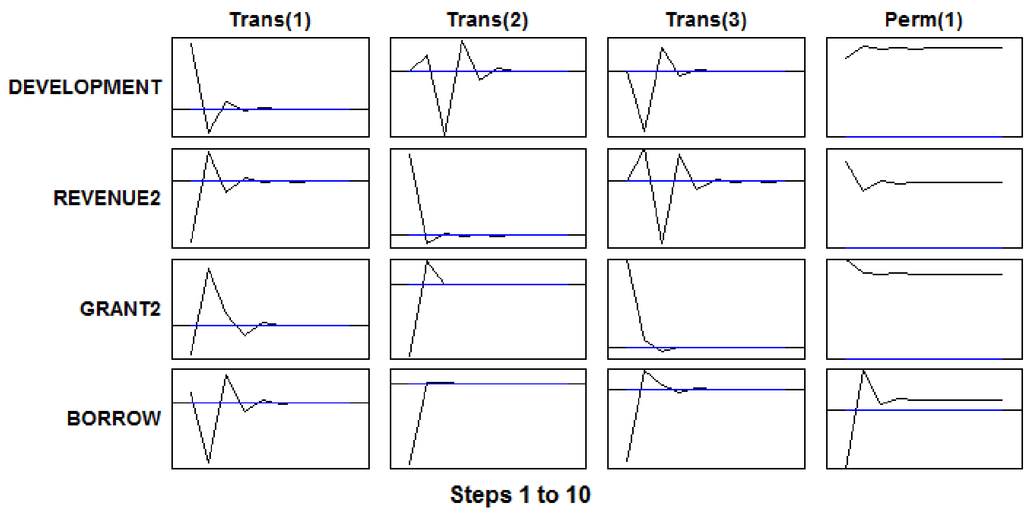
```
The Coefficients of the Common Trends:

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.[35] 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.[36] 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 6: Impulse Response Functions



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

[35] 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.

[36] The lack of significance intervals does not allow us to infer whether this is significantly different from zero.

Table 23: 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

7. 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, 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.

These results are not totally surprising. External aid inflows have financed a considerable share of capital expenditure in Ethiopia, either through earmarked investment projects or (indirectly) through budget support (Geda, 2001:171). This certainly explains the strong correlation between foreign aid grants and development expenditures. However, earmarked aid inflows usually lead to two simultaneous entries in the budget accounts, one under (capital) spending and another under (total) revenue. This suggests that the dynamic behaviour is mainly capturing the causality between budget support grants and development spending, which in turn may reveal the specific characteristics of donor-recipient relations in Ethiopia. According to Furtado and Smith (2007), the Ethiopian government is "more assured of its own directions, of its entitlement to set the development agenda, and of its stature vis-à-vis donors than are the governments of many other low-income countries." This can be partly explained by wellestablished domestic governance structures and long-standing sovereignty. Hence, while we would usually expect the causality to run from foreign aid to development expenditures – i.e. an increase in aid would lead to higher spending – the case of Ethiopia appears to be different. The government sets its spending targets according to its own development objectives, and then tries to find resources to finance those ambitions – usually in the following order: domestic revenue, aid inflows, and domestic borrowing. So far, donors have been keen to finance these expenditures, albeit with some level of unpredictability.

This leads us to the next point. When donors withhold aid disbursements due to political or bureaucratic issues, or domestic revenue falls short of expectations, the government is faced with a difficult choice: either postpone spending plans, or resort to domestic borrowing to cover its financing needs. The empirical results appear to suggest that the government tends to opt for the latter. Conversely, a positive aid shock (e.g. a large disbursement of budget support grants) is partly used to retire domestic loans, most likely the ones incurred by the previous aid shortfall. Again, this resonates with the hypothesis of a strong 'developmental state', since the priority is to keep expenditures at the desired (planned) level.

Finally, we do not find evidence that foreign aid is either used to boost the (non-development) recurrent budget or reduce the tax effort. This can be explained by the fact that recurrent expenditures have been mostly financed by domestic revenues (UNECA, 2002:89). In terms of the tax effort, we could even expect a positive effect. Since Ethiopia's economy is significantly constrained by the lack of international reserves, aid inflows provide vital foreign exchange to finance imports that in other circumstances would not take place. In turn, these imports will contribute to increase trade-related revenues. In fact, detailed fiscal data suggests that tax revenues accruing from international trade have been increasing despite the trade reforms implemented since the early 1990s (e.g. import levels have more than compensated the reduction in import tariffs and the abolishment of export taxes).

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.

References

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.

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.

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

Furtado, X. and J. Smith (2007), "Ethiopia: Aid, Ownership, and Sovereignty." Global Economic Governance Working Paper 2007/28, University of Oxford.

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.

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.

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 (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 (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), 188-136.

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

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.

UNECA (2002). Economic Report on Africa 2003: Tracking Performance and Progress, (Chapter 3: Ethiopia – Good Policies, Decent Outcomes). United Nations Economic Commission for Africa, Addis Ababa.

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.

Appendix

CVAR Estimates

Table 24: CVAR Model with full rank

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

Table 25: 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)

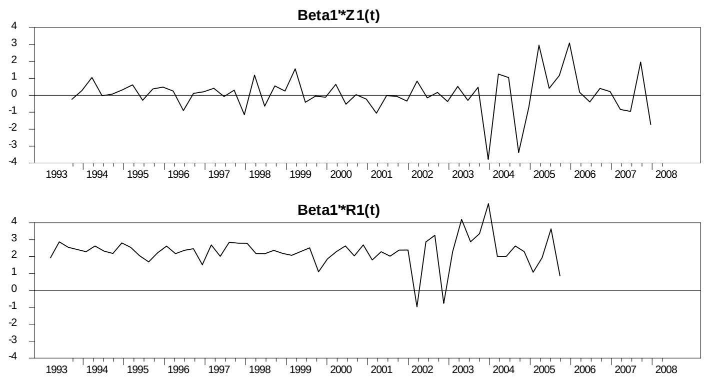
PI

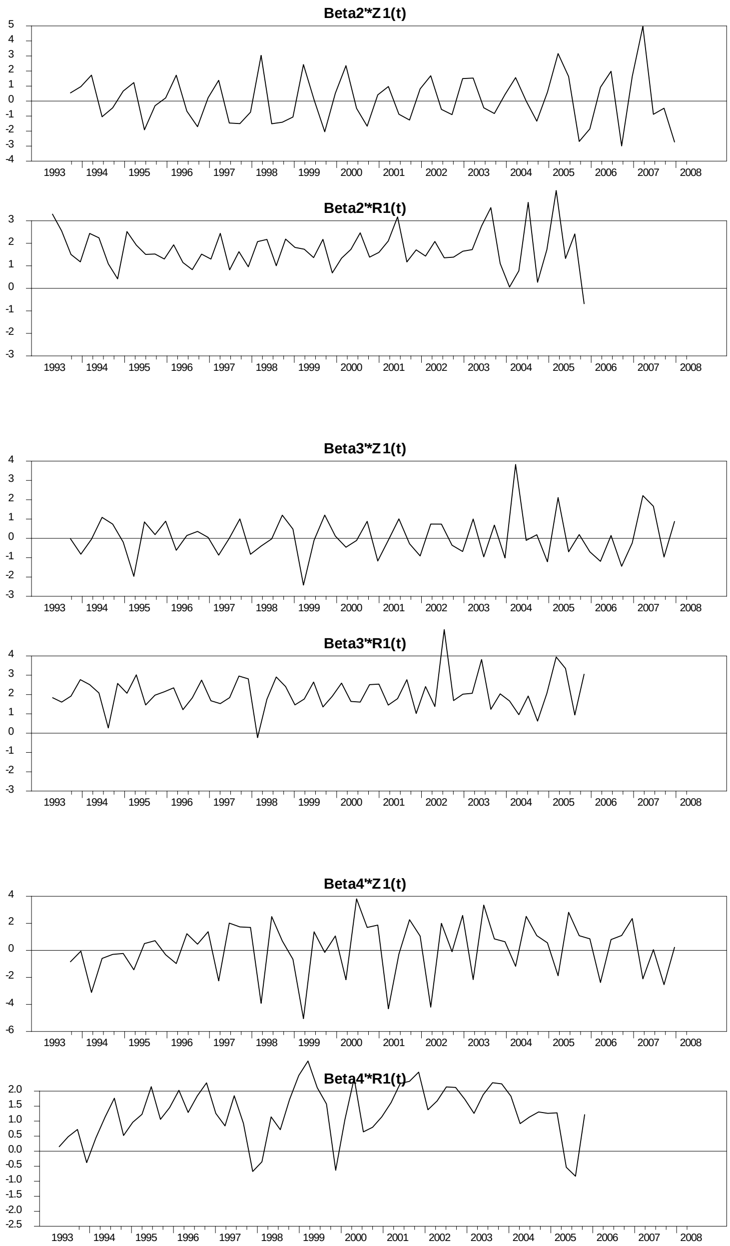
		REVENUE2	GRANT2	BORROW	LOAN	CURRENT5	TREND
DDEVEL	-0.514	0.754	0.231	0.178	0.352	-0.392	-0.010
	(-5.400)	(5.530)	(1.455)	(2.639)	(4.943)	(-5.009)	(-2.859)
DREVEN	0.542	-1.042	0.049	0.003	-0.312	0.395	0.021
	(5.784)	(-7.768)	(0.312)	(0.045)	(-4.456)	(5.130)	(5.794)
DGRANT	0.274	0.190	-0.926	0.000	0.043	-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

Cointegrating Relations

Figure 7: Cointegrating Relations





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.

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.

We mainly report forward recursive tests, since the backward recursive tests did not reveal any problems. 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.

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. 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.

Figure 8: Eigenvalues

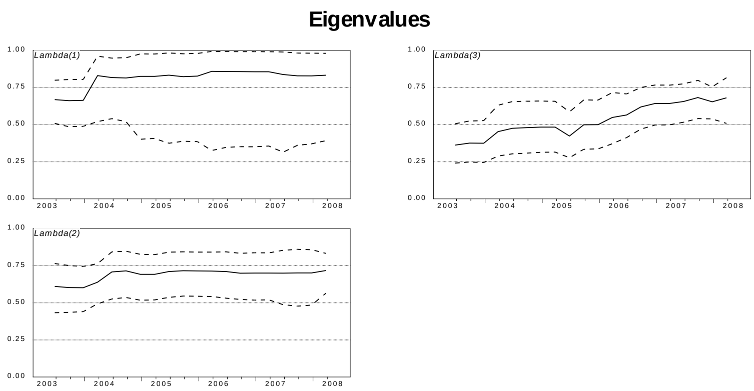
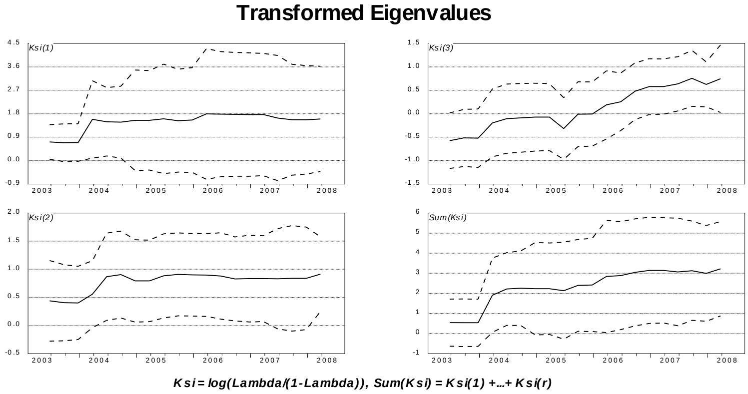
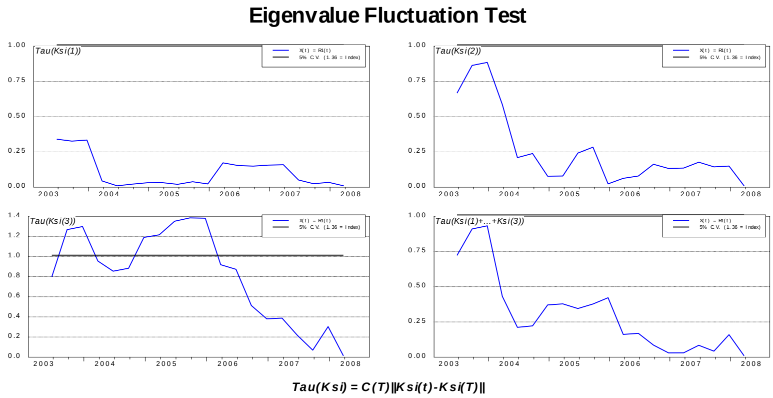


Figure 9: Transformed Eigenvalues



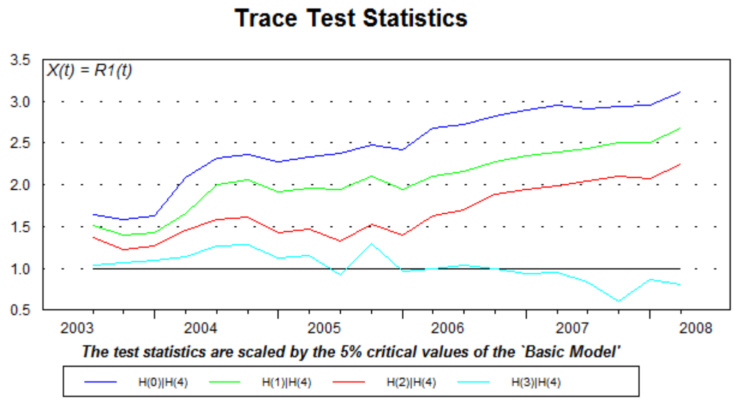
The following graph is the 'fluctuation test of the (transformed) eigenvalues'. The test is a *supremum* test and often regarded as rather conservative (Dennis, 2006:97-8). This means that if the test is rejected, there is strong evidence of non-constancy of the eigenvalues. The test statistic is scaled 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 10: 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 11: 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 12: Test Beta Constancy

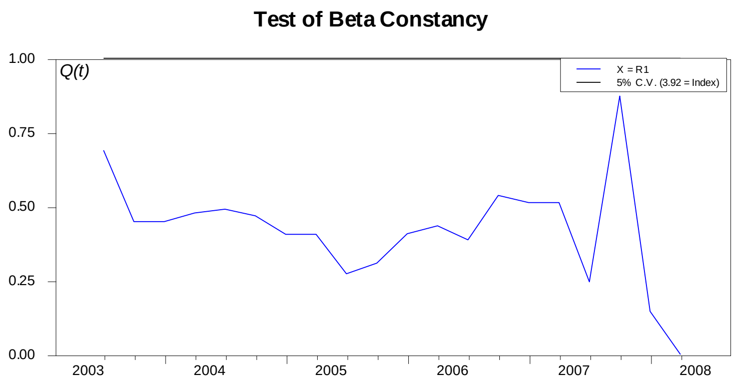
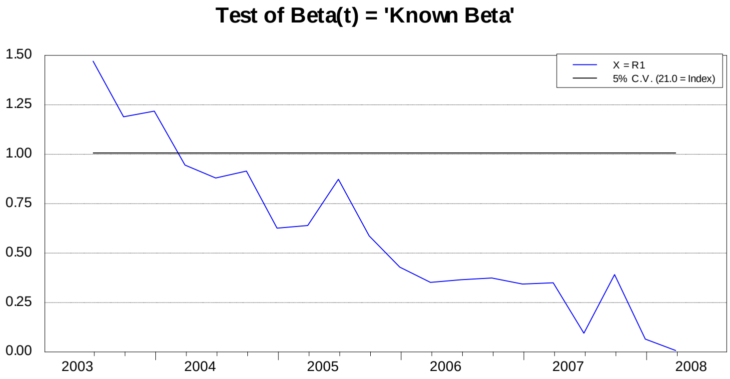


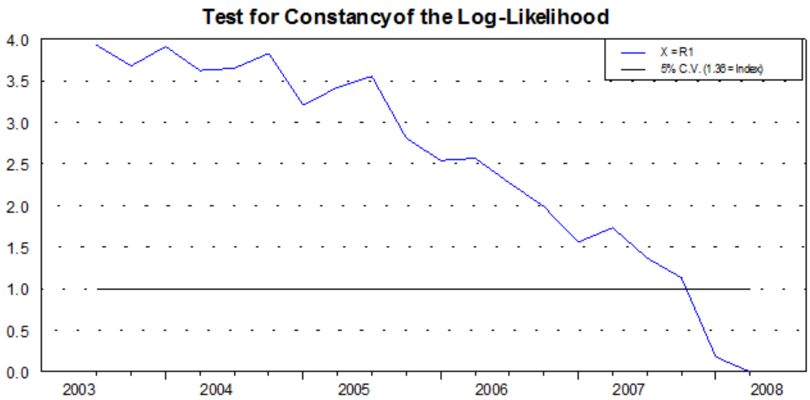
Figure 13: Test of Beta(t) = 'Known Beta'



Constancy of the (Maximised) Log-Likelihood Function

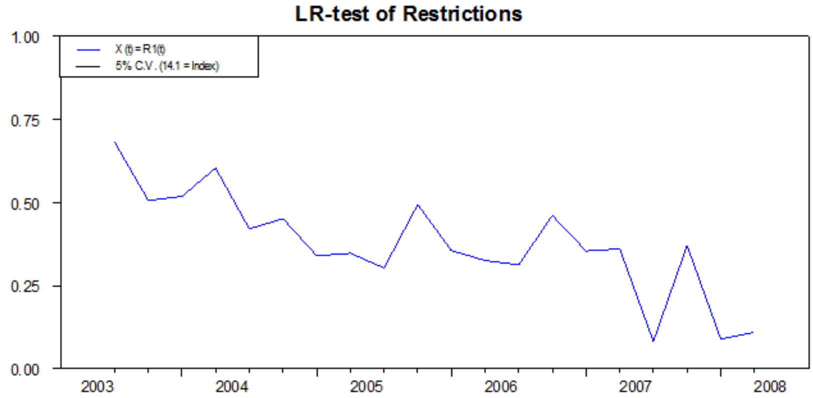
This is "essentially a test on the variances that measures the distance of the subsample and full sample estimates of the covariance matrix Ω" (Dennis, 2006:101). Again, the test statistic is scaled by the critical value. The result seems to suggest that there may be problems in our model. While this 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 14: 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.

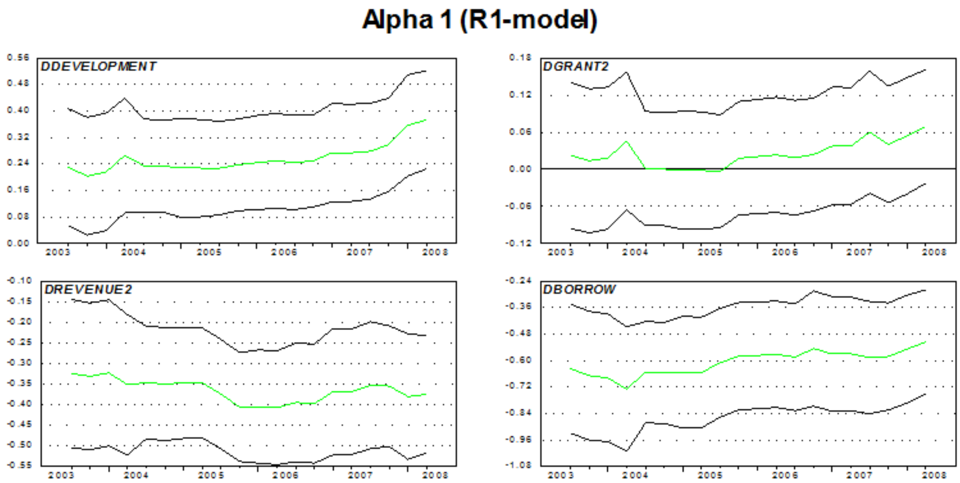
Figure 15: LR-test of Restrictions



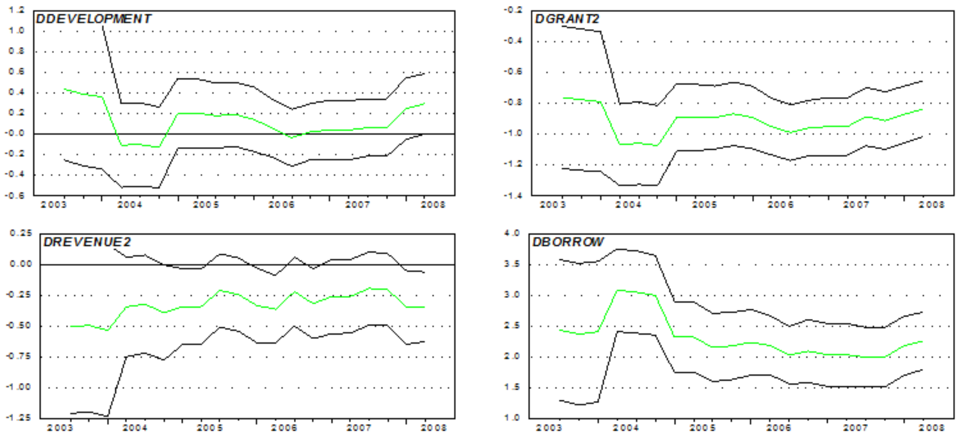
Parameter Constancy

Finally, the following graphs 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).

Figure 16: Constancy of the Coefficients of Adjustment



Alpha 2 (R1-model)



Alpha 3 (R1-model)

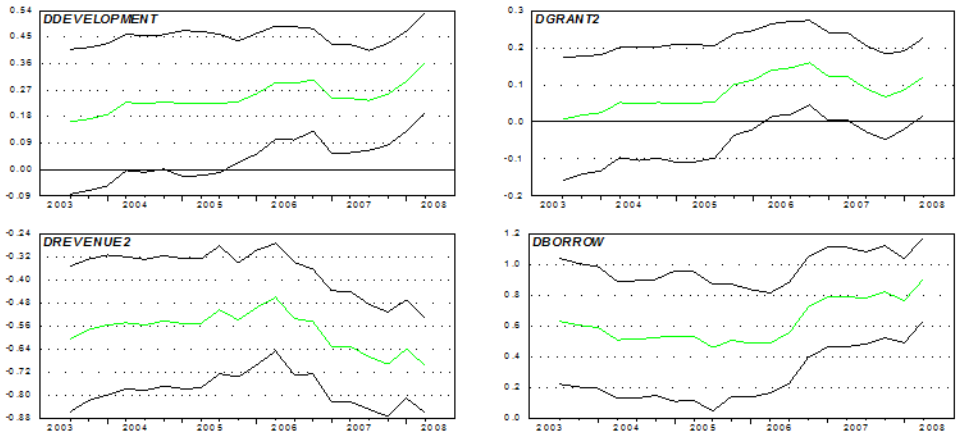
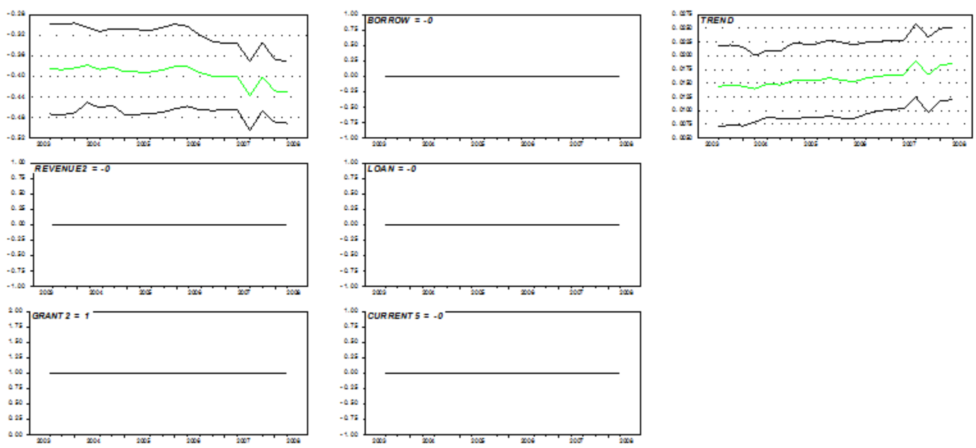
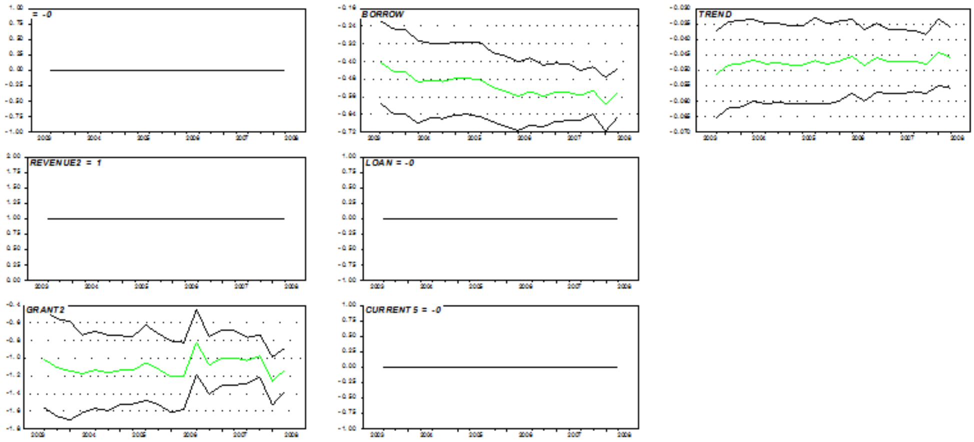


Figure 17: Constancy of the Long-Run Coefficients

Beta 2 (R1-model)



Beta 3 (R1-model)



Overall, the stability tests do not seem to provide strong evidence of model misspecification, since there were no strong rejections.