University of Sussex
Economics

Working Paper Series

No. 73-2014

Aid and Taxation: Evidence from Ethiopia

Giulia Mascagni
Institute of Development Studies
Department of Economics, University of Sussex
G.Mascagni@sussex.ac.uk

Abstract: The relation between aid and tax has been largely debated in the literature, given its far-reaching consequences: the presence of a crowding-out effect of aid on domestic revenue would seriously impair the sustainability of the development process. This paper explores this relation by adopting a case-study approach, which overcomes some of the common limits of the cross-country literature. I use time series data for Ethiopia for 1960-2009, a longer time series than most country studies of this kind. The estimation is based on an error correction model that allows separating long-run equilibrium relations and short-run dynamics. The analysis shows that both foreign grants and loans have a positive relation with tax revenue in Ethiopia. This effect seems to be robust to endogeneity and to structural breaks, although clearly establishing causality remains a challenge. The results show that aid has a beneficial effect on tax revenue, which may be due to its role in supporting fiscal reforms and improvements in tax administration.

Key words: Foreign Aid, Taxation, Grants, Loans, Ethiopia

JEL classification: F35, O23

Contents

List of Tables

List of Figures

1 Introduction

This paper provides an empirical analysis of the relation between aid and taxation in Ethiopia, using a unique dataset of 50 annual observations. This is a longer series than those used in most of the literature[1]. The main hypothesis explored is the existence of a crowding out effect of aid on tax revenues. In addition to the aid-tax relation, this paper offers insights into the role of tax determinants such as the manufacturing sector, agriculture and trade.

I fully exploit the case study approach by complementing the econometric analysis with an in depth knowledge of the country specific context, grounded on the qualitative analysis in Mascagni (2014). Not only the qualitative analysis allows the interpretation of the results in a more informed and relevant way. It also informs the econometric exercise by offering a deeper understanding of issues such as endogeneity and structural breaks. Given the limited attention given to the country context in the majority of the econometric literature on this topic, this approach represents a novelty.

Ethiopia has received increasing amounts of aid in recent years and it is often considered an 'aid darling'. While per capita aid is still below the African average, the share of aid to GDP is relatively high compared to Sub-Sarahan Africa. Perhaps most importantly, Ethiopia has been identified by some donors as a focus country for increased aid flows in coming years. For example DFID identified Ethiopia as one of its focus countries in its aid review of 2011, as it ranked in the first 5% of countries with the highest 'need-effectiveness' index (Department for International Development, 2011). At the same time tax revenues amount to only about 11% of GDP (in the last year of the sample), a share that is well below the average for low income African countries (and for any other grouping of African countries). This situation, coupled with the centrality of domestic revenue in financing the Growth and Transformation Plan (GTP, the current five-year development plan), makes the analysis of the relation between tax and aid central and relevant for the policy debate in Ethiopia. This analysis provides evidence for a positive effect of aid on tax revenues, that is mostly explained by the role of aid in building capacity at the national level.

To gain a deeper understanding on the aid-tax relation, this paper explores the role of aid heterogeneity in terms of grants and loans. These are sometimes thought to have different effects due to the need for repayment for the latter, while the former would create fiscal space without giving rise to future liabilities (Gupta et al., 2004; Morrissey et al., 2006; Benedek et al., 2012). I find some support for this hypothsis in the long run, where loans have a stronger positive effect than grants. However there is still no evidence of a negative effect of grants in the long run, only of a weaker but still positive one. In the short run both grants and loans have a positive and significant effect.

Moreover by including other tax determinants in the analysis I find that manufacturing has a strong positive effect on tax revenues, probably due to its role as tax collector besides that of tax payer[2]. In addition agriculture does not

[1] This is true for the literature focussing on case studies and using annual observations.

[2] Firms collect collect taxes on wages and on consumption that are then transfered to the

appear to have a significant effect on aggregate tax, but it affects negatively domestic taxes. Trade has the expected positive effect. These results, while somewhat secondary with respect to those on aid, offer insights into the possible challenges and opportunities for a stronger tax revenue mobilisation, which is a top policy priority under the GTP.

Therefore this paper is driven by the following three research questions, that are reported in order of priority:

- Is aid a disincentive to tax revenue mobilisation?
- Does aid heterogeneity matter (i.e. grants and loans)?
- What are the determinants of the tax share in Ethiopia?

2 Literature review

The literature on tax effort is based on the analysis of the determinants of the tax share (sometimes also referred to as tax effort), defined as total tax revenue collection as a share of GDP. It is mostly based on cross country analyses using data from international sources (e.g. Government Finance Statistics by the IMF or the World Development Indicators from the WB) to ensure comparability across countries. Part of this literature has been specifically focussed on the relation between tax and aid, without finding a consensus on the sign and significance of that effect.

The tax effort literature has its roots in the early studies conducted within the International Monetary Fund (Chelliah, 1971; Chelliah et al., 1975; Tait et al., 1979). They identify the core tax determinants that are still largely used in this literature today, and namely: mining share, trade openness (export + imports), GDP per capita and agricultural share. They use mainly cross-section estimation techniques often applied to averaged data on different periods. Typically a tax effort index is computed as the ratio between the actual tax share and the predicted one from the econometric model. This index is used for guidance in evaluating the fiscal performance of the countries considered. Tanzi (1992) largely confirms these early results by using cross section time series data for 88 countries. He finds that half the variation in the tax share can be attributed to those core variables, with the addition of the foreign debt share of GDP.

The most important theoretical contribution to this literature certainly is the article by Heller (Heller, 1975), that develops the framework largely used today in most of the literature. Heller's framework originally focussed on variables from the public budget alone including tax, aid, borrowing, and expenditure. His empirical estimation of the model on cross section-time series data from 11 African countries underlines a negative effect of aid on tax effort. This framework is further tested in the fiscal response literature that includes mostly country case studies such as Franco-Rodriguez et al. (1998), Machado (2009),

state.

Mavrotas and Ouattara (2006), Franco-Rodriguez (2000), and Osei et al. (2003). The results on aid in this litreature are contrasting, with the former three articles finding a negative effect on tax revenue and the latter two a positive one.

The seminal work of Heller (1975) is further developed and expanded by Leuthold (1991) and Ghura (1998) to provide the theoretical basis for tax effort models, including also the quality of institutions and macroeconomic policies along with aid. The former study provides an empirical application on a panel of eight African countries observed over 9 years. The author disaggregates tax revenue in direct and indirect taxes and finds a positive effect of foreign grants on the former and a negative one on the latter. Ghura (1998) also focuses on African countries, again over a relatively short period between 1985 and 1996. He finds that aid has an adverse effect on tax revenue. However he underlines that endogeneity may be a problem if grants are directed towards less performing countries (therefore with higher needs) where the tax share is lower. The author also finds that corruption and macroeconomic policies are important determinants of the tax share. Mahdavi (2008) further expands the theoretical famework by adding new sets of explanatory variables related to demographics and labor force characteristics; administrative and enforcement costs; the macroeconomic environment; and the political environment and corruption. This rich specification is estimated using an unbalanced panel dataset of 43 developing countries observed over the period 1973-2002. The paper finds that aid still has a negative and significant (at the 10% level) effect even after all the other variables are included in the model.

Providing a more positive picture on aid, Gupta (2007) suggests that aid has a positive effect on tax revenue and that this result is even stronger for low income countries. He uses a bigger sample than previous studies, including 150 developing countries observed over 25 years. Stotsky and WoldeMariam (1997) focus particularly on the effect of IMF programs on tax effort in the receiving country, failing to find strong support for a beneficial effect. Other studies in this tradition have focussed also on corruption and the institutional environment (Bird et al., 2008; Imam and Jacobs, 2007), or on natural resources (Bornhorst et al., 2008). As far as other tax determinants are concerned, these studies generally find a negative effect of agricultural share in GDP, a positive effect of trade openness, a weak positive effect of manufacturing, and contrasting results on GDP per capita.

Some studies have looked specifically at aid heterogeneity, particularly disaggregating it in grants and loans. As underlined in Gupta et al. (2004), loans may have a positive effect on tax effort because of the need for repayment while grants may instead be expected to crowd out domestic revenue, therefore having an expected negative effect. This hypothesis is tested on a dataset of 107 low and middle income countries from 1970 to 2000. The results on the variables related to the tax base (namely agriculture, industry, trade and income) are in line with the literature. The hypothesis on grants and loans is confirmed by finding the expected coefficients to be significant. Morrissey et al. (2006) estimated a similar specification using a comparable sample over the period 1975-2000 and confirmed these results: loans have a positive and significant effect on tax rev-

enue, grants have a smaller negative and significant effect, so that the combined effect (when the two variables are combined in total aid) is a positive but non-significant coefficient. When lagged aid is used however, to account for possible endogeneity, the negative effect of grants becomes non-significant and total aid is found to have a positive effect on tax effort. Teera and Hudson (2004) reach a similar result, finding a positive but not significant result on the aid variable. By using interaction terms the authors also find that the importance of trade and manufacturing in influencing the tax share decreases as countries get richer.

More recently there has been a renewed interest in the estimation of the aid-tax relation, perhaps also sparked by a reflection on the existing evidence. In particular Carter (2013) provides a critical summary of the literature on the aid-tax relation underlining its methodological problems and other limitations. He provides new estimates, partly addressing the methodological concerns, that provide no evidence of a negative effect of aid on tax revenues perhaps also due to the increased attention of donors to domestic revenue mobilisation. Benedek et al. (2012) try to address specifically some of these concerns, particularly as regards endogeneity. They use a panel of 118 countries over 1980-2009 and a number of econometric methodologies including GMM. The results show a negative coefficient on grants that however is weakening over time, particularly when compared to the results of Gupta et al. (2004) that found a larger coefficient. They also estimate separate regressions for disaggregated tax types. While aid has a negative effect on most tax types, it has a positive relation with trade taxes. Clist and Morrissey (2011) use a large sample of developing countries and find that aid has a positive effect on tax revenue after the mid-80s when there appears to be a break in the relation. This break corresponds to a stronger emphasis on conditional lending in the international debate and it may indicate the success of conditionality in stimulating fiscal reforms. To address endogeneity, the authors use lagged aid (in addition to current aid in a separate specification) and find a positive result on both grants and loans after 1985. Aid lagged two years is also used as an instrument, yielding similar results. Mkandawire (2011) focuses on the effect of the colonial heritage on tax effort in a cross section of African countries, using averaged data over four-year intervals in the period 1984-2004. The author argues that the colonial status of African countries has implications in terms of taxation that can still be observed in today's tax systems. While aid it is not the focus of the analysis, it is still included in the model and it is found to have a positive, although weak, effect on the tax share. This result is confirmed even when its lagged value is used as an instrument to account for endogeneity. Other recent papers largely confirm the results on the tax determinants related to the tax base, although they find no evidence of a significant relation between aid and taxation (Drummond et al., 2012; Le et al., 2012).

While cross country studies are prevalent in this literature, they also suffer from limitations due to the heterogeneity in the effect of aid amongst countries. Carter (2013) argues that due to this heterogeneity, estimates of averaged effects are unstable and they are of limited policy interest. One possible way forward indicated in the paper is the increased reliance on country case studies that can

also include interviews with government officials. This is the approach adopted in this research and it is in this niche that I aim to contribute.

So far there are only few studies of tax effort that take a case study approach. Amongst these, Ezemenari et al. (2008) analyses Rwanda and it finds a small negative effect of aid. The paper also contributes to the literature by providing a theoretical framework alternative to Heller (1975). Other country studies have focussed mostly on the fiscal variables alone, following more the fiscal response tradition than the tax determinants literature (Hisali and Ddumba-Ssentamu, 2013; Martins, 2010; Osei et al., 2005; Bwire, 2013). They generally provide evidence of a positive effect of aid.

Contribution

This paper contributes to the literature by providing a country study on Ethiopia, in a literature largely dominated by cross-country studies. I provide evidence for a positive relation between aid and taxation, therefore contributing to a debate that is still far from settled.

By complementing econometric evidence with qualitative information, I show the advantages of a case study approach. This is reinforced by in depth interviews of government and donor agencies officials that are specifically aimed at understanding the aid-tax relation. Since this approach is rather rare in the literature, it represents an original contribution aimed at overcoming the limits of the cross-country literature.

3 Data and empirical methodology

The data used to investigate the research questions outlined in section 1 is Ethiopian time series from 1960 to 2009[3]. The characteristics of this data, as well as the data generating environment are discussed in the next two sections. The last section describes the empirical model used to analyse this data and it discusses some econometric challenges.

3.1 The data

The Ethiopian dataset spans 50 years, from 1960 to 2009. By relying on 50 annual observations, this exercise can count on a larger sample than other country-level studies working with annual data. In a small sample context, even adding a few observations adds robustness to the results. This dataset is therefore a crucial asset for this analysis.

The dataset was compiled exclusively from national sources, and particularly from the National Accounts data originally compiled by the Ministry of Finance and Economic Development (MOFED). Using a single national source presents two advantages. First, it allows consistency in the data that cannot be fully

[3]These are more specifically 1960/61 and 2009/2010, that are the Gregorian calendar years that correspond to the relevant Ethiopian years. The Ethiopian calendar is 7 years behind the Gregorian one and it starts in September.

ensured when mixing data sources. This is particularly the case of aid figures that may present huge discrepancies across different datasets, as discussed in the next paragraphs. Secondly the national data is used for policy making, thus making it more relevant than international datasets.

The use of national data is particularly important in the case of aid measures, namely grants and loans. The national figures used in this exercise consist only of aid that flows through the government budget (i.e. the Treasury). This includes, but it is not limited to, budget support. The use of this data for aid is not only sensible in terms of data consistency but it is also relevant to the research questions presented in section 1. Indeed the component of aid flowing to the budget is the closest substitute to tax revenue and it is therefore the one that is more likely represent a disincentive or substitution effect. Of course there are issues related to aid fungibility, but nonetheless it seems reasonable to consider this component of aid as the most relevant one for the analysis. This is reinforced by the fact that budget data is the one used by policymakers when taking fiscal decisions and it is the portion of aid that they are fully aware of.

The national aid figures are smaller than the ones reported by OECD-DAC[4]. This discrepacy is due to items within grants that are not disbursed through the government system, mainly humanitarian and emergency aid that in Ethiopia can be rather substantial due to the constant threat of drought and famine. Moreover items that are not included in the budget grant figure are, for example, projects managed by donors and grants to NGOs.

The national aid figure today is largely dominated by the Protection of Basic Services (PBS) project. Other programs disbursed through the Treasury are for example the Public Sector Capacity Building Program (PSCAP), the Productive Safety Net Program (PSNP) and other funds aimed at specific sectors such as infrastructure.

All variables are measured as a share of GDP, except for GDP per capita (log of constant GDP per capita) and the GDP gap (percentage deviation of GDP from trend). A summary of the variables used in this analysis is reported in appendix A.

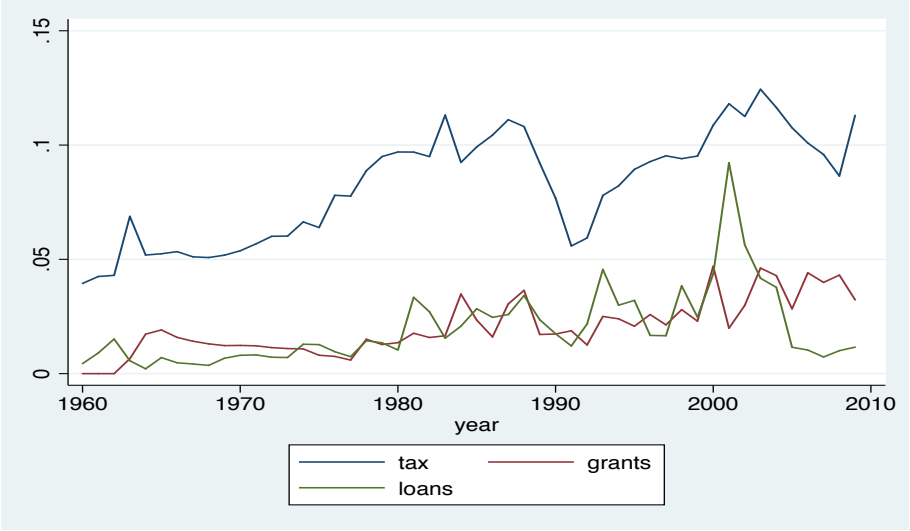
3.2 Ethiopian context and some descriptives

This section highlights briefly a few qualitative elements that are particularly relevent when carrying out the econometric exercise. It therefore does not aim at providing an extensive account of the political, hisotrical, and economic context of Ethiopia. A full qualitative analysis of the Ethiopian historical and economic context can be found in Mascagni (2014), which relies on data and informtion collected in Ethiopia, including interviews to stakeholders such as donors, the government and independent experts.

First of all, it is important to recognise that Ethiopia has a history of independence that is unique in the African context. This history translates in fiscal terms in the importance assigned to tax revenue mobilisation by all governments

[4]For example the grants figure is about a fourth of the corresponding OECD-DAC one.

Figure 1: Plot of tax, grants and loans



Note: all variables are expressed as a percentage of current GDP. Source: author's calculations using data from the Ministry of Finance and Economic Development.

that have ruled Ethiopia over the period considered.

As far as aid is concerned, Ethiopia has received both budget aid and technical assistance throughout the whole period analysed. The latter in particular was initially provided on the basis of the early sectoral and multi-year development plans of Haile Selassie. Even in the first imperial period, missions from the UN, the WB and the US provided advice in the field of taxation, amongst others. Still today taxation is one of the areas of greatest agreement between the government and donors, in a relation that is often characterized by confrontation and contrasting views.

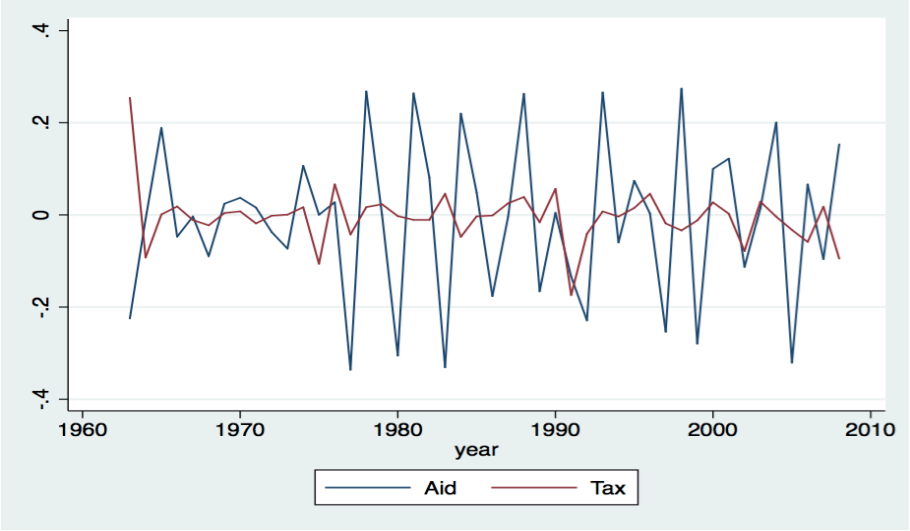
Figure 1 shows a plot of tax, grants and loans over the period, all measured as a share of GDP. The first message from the picture is that tax is a much larger source of revenue in the government's budget than grants and loans. Secondly the figure does not immediately suggest a negative relation between tax revenue and the aid variables.

Moreover figure 2 reports the 3-year moving average of aid (aggregated grants and loans, for the sake of clarity) and tax, showing that the former is much more volatile than the latter thus making it a more unpredictable source of revenue.

3.3 Empirical framework and challenges

The Ethiopian time series are used to estimate a standard tax effort equation, stemming from the literature. Equation 1 is largely in line with other studies in this literature, where A is aid, later disaggregated in grants and loans, NT is

Figure 2: Deviations from three-year moving average of tax and aid



Note: the figure reports deviations from the 3-years moving average of the grants and tax series in levels. Source: author's calculations using data from the Ministry of Finance and Economic Development

non tax revenue, $Agri$ is agriculture, $Manuf$ is manufacturing, $Trade$ is trade openness, and $GDPpc$ is GDP per capita. Appendix A reports a more detailed descrition of these variables. Some specifications add to these core variables a cycle variable (GDPgap) and dummies to reflect changes in regime.

$$T = b_0 + b_1 A + b_2 NT + b_3 Agri + b_4 Manuf + b_5 Trade + b_6 GDPpc \quad (1)$$

In equation 1 tax revenue is determined by variables that are proxy measures of components of the tax base, so the equation is largely in line with the cross-country tax effort literature. The theoretical basis for equation 1 is provided by Heller-type models, particularly as developed in the literature for example by Leuthold (1991) and Ghura (1998) (see section 2). The Heller-type fiscal response models focus largely on the effects of aid on spending, and revenue (including aid and tax) is required to finance spending. In Heller (1975) and other early papers aid does not enter the government utility function, but it is later introduced by Franco-Rodriguez et al. (1998) where aid and tax are allowed to affect differently government utility.

In a time series setting, variables may present trends. If the variables of interest move together in the long run, they may be cointegrated and this issue is discussed below. As far as short run dynamics are concerned, trends in the data become irrelevant since first differences are used in the empirical estimation. In particular equation 1 can be estimated using a cointegration model that allows

separating the long run (LR) and short run (SR) effects. I use the two-step procedure suggested by Engle and Granger (1987). The Engle-Granger (EG) procedure requires the series to be integrated of the same order in levels (first step) and to be stationary in first difference (second step). To this end three tests for stationarity are used: the Augmented Dickey-Fuller test (*dfgls*), the Kwiatkowski-Phillips-Schmidt-Shin test (*kpss*), and the Clemente, Montanes, Reyes unit root test (*clem*). The last one in particular allows for the presence of up to two structural breaks[5] and it is therefore useful in the Ethiopian case where breaks may be expected in correspondence with regime changes[6]. It is important to remember however that the *clem* test is not a test for structural breaks as such but instead it only analyses the properties of a single series (Perron and Vogelsang, 1992). The results of the stationarity tests show that all variables are I(1)[7]. They are reported and discussed in appendix B.

The first step of the EG procedure identifies long run relations and it involves the variables in level, and more specifically as shares of GDP. Cointegration occurs when the relation between those non-stationary variables produces stationary residuals. If this is the case the series move together in time and they form a long-run equilibrium. The test for cointegration proposed by Engle and Granger is therefore esentially a stationarity test on the residuals similar to the Dickey-Fuller test[8]. Since this statistic does not follow the standard distribution, the MacKinnon critical values are used (MacKinnon, 2010). When residuals are not serially independent the test includes lags of the first difference of the residual and this version of the test is known as the Augmented Engle Granger (AEG) test.

The second step of the procedure identifies short run effects. Here the variables are included in first difference and they are therefore stationary, provided that they are I(1). The lagged residual from the first step is also included as an error correction term (ECT), indicating the adjustment to LR equilibrium.

As far as econometric challenges are concerned, two issues are of particular interest: endogeneity and structural breaks.

Endogeneity regards in particular the aid variables and the short run equations since LR cointegrating relation is superconsistent (Engle and Granger, 1987). Concerns are raised in the literature based on the idea that aid may be given to countries with higher needs, which may happen to be also those countries with a lower potential to raise taxes due to low levels of income. If

[5]The Zivot-Andrews test can only take into account one break (Zivot and Andrews, 1992).

[6]In the *clem* test the breaks are not imposed a priori but are estimated by the test. Using this test is particularly important because Dickey-Fuller style tests may confuse non-stationarity with structural breaks. In other words, in presence of structural breaks, the residuals will display a clear pattern and the classic unit root tests are not able to reject the null hypothesis of non-stationarity (Ghosh, 1999).

[7]The *grants* variable shows some stationarity in levels that however disappears in the *dfgls* test with one lag and when treated in conjunction with *loans* in the *aid* variable. The series is treated as I(1), also to ensure consistency with other variables.

[8]The test regresses the first difference of the residuals from the first step (LR) equation on the lagged level of these residuals, without a constant. The test statistic is a simple OLS t-statistic on the lagged residual.

this is the case, a negative spurious relation would indeed be observed between aid and tax. This however is a concern more in relation to cross-country studies than in a time series setting. In time series analysis endogeneity may stem from a possible detrimental effect of aid on domestic institutions that would therefore undermine the administrative basis for an effective fiscal system, resulting again in a spurious negative relation. Moreover aid flows may respond to government's decisions in terms of fiscal and economic policies, rewarding 'good' policies with increased financial assistance. If donors believe that increasing tax revenue is a 'good' policy and reward government that succeed in doing so, a spurious positive relation would occur. However this case seems particularly ill-grounded in the Ethiopian case. The following quote (Furtado and Smith, 2007) summarizes the argument for this, which is also exposed in more detail in section 5.1.2.

> "Changes in the level of donor assistance to Ethiopia have been driven overwhelmingly by political and geopolitical considerations [...]. These factors, rather than anything the government has consciously done to manage the aid agenda, have largely determined the level of aid inflows."

Endogeneity is discussed in detail in section 5.1 and an instrumental variable approach is proposed. Moreover the role of the institutional environment and governance are explicitely taken into account in section 4.3.

Turning to structural breaks, they are particularly expected in correspondence with the regime changes of 1974 and 1991. Two shift dummies are therefore included in the regressions presented in section 4 and this issue is explored in more detail in section 5.2. The dummies capture the Derg (1975-1990) and EPRDF (1991-2009) regimes, with the Imperial period (1960-1974) being excluded and thus representing the reference period. While breaks need to be taken into account, it is important to keep in mind the elements of continuity underlined in more deatil in Mascagni (2014). This is particularly relevant in the case of the administrative apparatus, that was largely inherited by successive regimes. While breaks can be expected in 1974 and 1991, section 5.2 explores the possibility of breaks other than these without however finding any evidence for it.

4 Results

This section presents the main results from the estimation of equation 1 using the Engle-Granger two step procedure. It is divided in two sections that report respectively the long run and short run results.

4.1 First step: the long run

Table 1 reports the long run (LR) results for different specifications of the empirical model, using all variables in levels[9]. Cointegration amongst these variables is crucial to ensure consistent estimates and to proceed to the next step of the Engle-Granger procedure. Therefore table 1 also includes the AEG test statistics at the bottom, for testing cointegration. These test statistics are to be compared with the MacKinnon critical values reported in table 2 and calculated using the Stata program module devised by Schaffer (2010). The null hypothesis is no cointegration, so rejection would provide evidence that the series are indeed cointegrated.

The first column of table 1 reports the full specification excluding regime dummies. However the AEG test statistic shows that this regression does not produce stationary residuals, therefore indicating the lack of cointegration. Since structural breaks may produce a pattern in the residuals, therefore making them non-stationary, column two includes regime dummies and with these the system is cointegrated. Indeed this regression passes the AEG test with 0 lags. Further lags in this case are not needed because the residuals do not show evidence of autocorrelation[10]. The regime dummies are therefore included in all remaining specifications.

The third column excludes the GDP gap variable over concerns of a possible built-in negative correlation with the dependent variable due to GDP being at the denominator of the tax share. This may also explain the negative coefficient on GDP gap that indicates a countercyclical behavior of tax revenue and that may be somewhat puzzling. Dropping the GDP gap has little impact on the results but it does allow grants to reach the 10% significance threshold. Note that this specification passes the cointegration test also when one lag is included, although zero lags would be sufficient since autocorrelation is rejected (see table C.1 in appendix C for autocorrelation tests on the LR equation). Given its superior cointegration properties and the little impact of dropping the potentially problematic GDPgap variable, column three represents the preferred specification[11].

Finally specifications 4 and 5 take multicollinearity into account by including alternatively *agri* and *trade* that are found to be the variables most affected by this issue[12].

For all specifications a test for homoskedasticity (White and Breusch-Pagan tests) is run and it provides no evidence of a problem. The residuals also pass the normality tests. The R^2 for all LR equations is rather large, ranging from 0.86 to 0.92, which may be expected in presence of cointegration.

The results on the aid variables are consistently positive across specifications, with *loans* seemingly driving the positive effect of aggreagate aid as it is always

[9] Mostly as a percentage of GDP, except GDP per capita and the GDP gap.

[10] Table C.1 in appendix C shows autocorrelation tests for the LR equation. They also show that autocorrelation disappears once regime dummies are included.

[11] This specification is used to compute the error correction term (ECT) for the second step of the Engle-Granger procedure and it is the basis of robustness checks.

[12] More details on multicollinearity are available from the author

Table 1: First step: LR results from tax equation

	(1)	(2)	(3)	(4)	(5)
grants	0.116	0.293	0.325*	0.325	0.248
	(0.52)	(1.65)	(1.78)	(1.48)	(1.56)
loans	0.215	0.305***	0.413***	0.428***	0.397***
	(1.62)	(2.81)	(4.31)	(3.74)	(4.24)
non-tax	0.260	0.194	0.163	0.304*	0.154
	(1.60)	(1.53)	(1.26)	(2.02)	(1.19)
agri	0.053	0.064	0.052	-0.038	
	(0.88)	(1.09)	(0.87)	(-0.56)	
manuf	1.366***	0.766**	0.736**	0.167	0.548**
	(4.00)	(2.48)	(2.31)	(0.48)	(2.35)
trade	0.087**	0.116***	0.144***		0.134***
	(2.09)	(3.29)	(4.34)		(4.32)
GDPpc	0.023	0.016	-0.007	0.039**	-0.011
	(1.09)	(0.79)	(-0.39)	(2.46)	(-0.66)
GDP gap	-0.028*	-0.025*			
	(-1.88)	(-1.91)			
Derg		0.019***	0.017***	0.017***	0.015***
		(4.10)	(3.61)	(3.06)	(3.58)
EPRDF		0.004	-0.003	0.003	-0.007
		(0.51)	(-0.42)	(0.41)	(-1.25)
Constant	-0.204	-0.152	0.012	-0.207	0.087
	(-1.41)	(-0.95)	(0.08)	(-1.34)	(0.80)
Observations	50	50	50	50	50
r2	0.86	0.92	0.91	0.87	0.91
Augmented Engle-Granger test for cointegration					
0 lags	-4.086	-6.363	-6.706	-6.687	-6.582
1 lag	-3.638	-5.734	-6.326	-5.884	-6.199
2 lags	-2.563	-4.745	-5.017	-4.708	-4.726

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All variables are included in levels, i.e. as a share of GDP except GDPpc and GDPgap. The dependent variable in all columns is the tax share.

significant except for column 1. *Grants* are significant only in specfication 3 at the 10% level. In particular a percentage point increase in the share of grants to GDP is associated with a 0.3 percentage points increase in the tax share, and the corresponding result for loans is 0.4 (specification 3). These findings are largely in line with the literature using the disaggregation between grants and loans, where the latter is found to have a stronger and positive effect.

Agriculture displays its expected negative coefficient only once multicollinearity is tackled, in specification 4, although it remains non-significant. Manufacturing and trade display the expected positive and significant coefficients which are confirmed in all specifications. In particular manufacturing has a large coefficient, implying that a percentage point increase in the manufacturing share of GDP is associated with a 0.7 increase in the tax share. This may be due to the fact that firms act as tax collectors, both for wage taxes and for indirect taxes. GDP per capita has a mostly positive coefficient that however is only significant in specification 4, suggesting that a 1% increase in GDP per capita increases the tax share by 0.04 percentage points.

Table 2: MacKinnon critical values

k (excl. constant)	10%	5%	1%
k=8	-5.328	-5.708	-6.478
k=9	-5.623	-6.011	-6.800
k=10	-5.907	-6.304	-7.112

Note: Critical values valid also for AEG with lags.

4.2 Second step: the short run

Having found cointegration in the first step, it is now possible to proceed to the second step (SR)[13]. The variables are now made stationary by first differencing[14].

Table 3 reports the SR results, starting in column 1 from the same specification as the LR one in column 3 of table 1. Let us note at the outset that overall the SR results are largely in line with the LR ones. Both aid variables still display positive coefficients in all specifications. Moreover in the short run both grants and loans are significant in all specifications, confirming that they both have a beneficial effect on tax effort. The effect of grants in the SR is larger in magnitude in all specifications. In particular, a percentage point increase in the share of grants in GDP is associated with a 0.4 percentage points increase in the tax share and the corresponding figure for loans is 0.25 (in the specification of column 1). Both coefficients are robust to changes in specification.

Column 2 considers a dynamic structure by including lags of grants and loans. The lags are not significant and their inclusion has only a small impact

[13]Specification 3 of table 1 is used to compute the ECT.

[14]For stationarity test on the variables in first difference, see appendix B.

on the aid variables. Their coefficients are still positive although smaller than the contemporaneous ones, as it may be expected. In column 3 the GDP gap is brought in the equation, with a negative and significant coefficient consistent with the LR result. Although multicollinearity is not flagged as a problem by the tests performed[15], columns 4 and 5 replicate the last two specifications of table 1 for completeness. Therefore *trade* and *agri* are included alternatively. Tests of autocorrelation and heteroskedasticity are carried out on all specifications and provide no evidence of the presence of these problems[16].

Trade is still significant and positive in all specifications, indicating that a percentage point increase in trade openness is associated with 0.2 percentage points increase in the tax share. Manufacturing in the short run is only significant at the 10% level in the last two specifications, while agriculture never reaches significance. This may imply that the structure of the economy, that is partly captured by these two variables, influences the tax share mostly in the long run. The regime dummies are not significant in any of the specifications, again indicating that they matter mostly in the LR. Finally, the error correction term (ECT) has the expected negative sign in all specifications and it indicates a quick adjustment to the LR equilibrium within one year.

4.3 The usual omitted 'suspect': governance

Governance is certainly the most obvious candidate as an omitted variable, that may result in endogenity. Indeed the relation between governance and taxation is well established in the literature on taxation and state building (OECD, 2008; Brautigam et al., 2008). Governance may also be thought to be related with aid, as donors can reasonably be worried about the effect of corruption and other bad administrative practices on aid spending.

Governance can be included using the International Country Risk Guide (ICRG) data on political risk which is available from 1985 to 2012. To the best of my knowledge, this is the longest available series of governance indicators. Given that its inclusion results in halving an already small sample, this variable was not included in the previous results. However it can be used to get a sense of the bias that its omission may induce. While the resulting sample is smaller, it is still sufficient to obtain indicative estimates.

The ICRG political risk variable is a measure of governance comprising the following 12 indicators (weights in brackets): government stability (12), socioeconomic conditions (12), investment profile (12), internal conflict (12), external conflict (12), corruption (6), military in politics (6), religious tensions (6), law and order (6), ethnic tensions (6), democratic accountability (6), bureaucracy quality (4). The maximum rating for each country is 100, that indicates very low political risk and the highest level of governance.

In particular two different variables from the ICRG dataset are used: the full indicator (*icrg*) and the indicator excluding the two conflict components

[15] Details available from the author.

[16] See autocorrelation tests for the SR equation in table C.2, appendix C.

Table 3: Short run results from tax equation (2nd step)

	(1)	(2)	(3)	(4)	(5)
non-tax	-0.043	-0.053	-0.005	0.060	-0.035
	(-0.34)	(-0.37)	(-0.04)	(0.38)	(-0.27)
grants	0.434***	0.491***	0.434***	0.186	0.426***
	(3.32)	(3.02)	(3.83)	(1.25)	(3.30)
L.grants		0.110			
		(0.72)			
loans	0.254***	0.242**	0.179**	0.176*	0.262***
	(2.98)	(2.64)	(2.34)	(1.70)	(3.16)
L.loans		0.016			
		(0.15)			
agri	0.066	0.047	-0.074	0.122	
	(0.60)	(0.40)	(-0.72)	(0.91)	
manuf	0.838	0.711	0.347	1.251*	0.595*
	(1.61)	(1.27)	(0.74)	(1.97)	(1.83)
trade	0.207***	0.212***	0.173***		0.210***
	(4.71)	(4.62)	(4.40)		(4.85)
GDPpc	-0.005	-0.005	0.057*	-0.001	0.007
	(-0.16)	(-0.16)	(1.86)	(-0.02)	(0.29)
Derg	0.001	0.001	0.003	0.000	0.002
	(0.45)	(0.35)	(1.15)	(0.14)	(0.59)
EPRDF	-0.002	-0.003	-0.002	0.001	-0.003
	(-0.86)	(-0.97)	(-0.85)	(0.38)	(-1.12)
ECM	-0.955***	-0.992***	-0.963***	-0.705***	-0.973***
	(-5.56)	(-5.02)	(-6.47)	(-3.48)	(-5.81)
GDP gap			-0.059***		
			(-3.68)		
Constant	0.000	0.000	-0.002	0.001	-0.000
	(0.06)	(0.10)	(-0.79)	(0.33)	(-0.18)
Obs.	49	48	49	49	49
r2	0.61	0.62	0.71	0.38	0.60

t statistics in parentheses

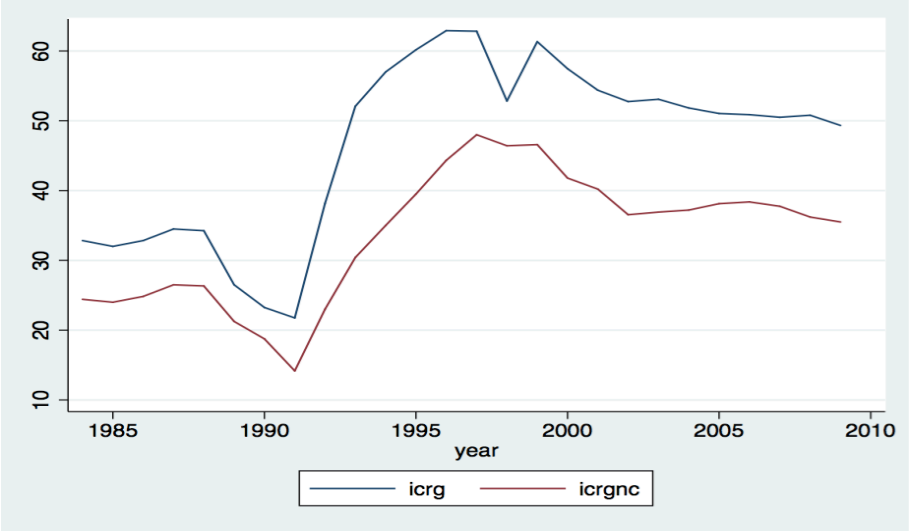
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All variables are included in first difference an the dependent variable is the tax share in all columns.

(*icrgnc*). These indicators are rescaled to take values from 0 to 1. The reason for excluding conflict is that it may not be strictly related to governance and it may have a different effect on tax than governance. For instance if taxes are increased at times of war because of increased financing needs, a higher ICRG indicator (i.e. less conflict) would be associated with a lower tax share. On the contrary the effect of governance on tax postulated in the literature is generally positive: better working institutions are better able to collect taxes and to overcome compliance problems. Having established in theory the relation between governance and taxation, endogeneity may arise when the former is also related with aid.

Figure 3 plots these two series and shows that indeed there is some variability in the ICRG indicators for Ethiopia over the period considered.

Figure 3: Plot of ICRG indicators



Source: International Country Risk Guide

Table 4 reports estimates using the full sample in column 1, to allow comparison with previous results. Column 2 reports the same estimates, still excluding the ICRG indicators, on the period from 1985 onwards to allow a direct comparison in the same time period. Note that only one shift dummy is included in this sub-sample since only one regime change occurs. Columns 3 and 4 report the results of two regressions including respectively *icrg* and *icrgnc*.

Both governance variables display the expected positive and significant coefficients. This result confirms the importance of governance in determining the tax share. Most importantly however the results on aid variables are only marginally changed by the inclusion of governance. The coefficients on grants and loans remain significant and of comparable magnitude with those of column

2, thus confirming the existence of a positive effect. In the case of loans the new estimates including ICRG data are both smaller, thus suggesting the presence of a small positive bias. However in the case of grants it is not clear whether the small bias would be positive (as indicated when using *icrg*) or negative (as when including *icrg*).

It is worth noting that the positive effect of the Derg dummy becomes significant when the governance variables are included. This is due to a purely statistical effect, where the dummy captures the jump shown in figure 3 while the coefficients on governance capture variations within the periods (i.e. Derg and EPRDF).

A further confirmation of the existence of a small positive bias due to the omission of governance can be obtained using the following standard formula (Greene, 2008), applied to the aggregate aid figure.

$$E[b|x,z] = \beta + \frac{cov(x,z)}{var(z)}\gamma \tag{2}$$

In equation 2, b is the estimated coefficient, x is the potentially endogenous variable (i.e. aid), z is the omitted variable (i.e. governance) and γ is the coefficient that z would have if it was not omitted. So the sign of the bias is determined by two elements: the effect of z on the dependent variable and the sign of the covariance between the x and z. All these variables are available from the previous estimation and using ICRG data. For the sake of simplicity the calculation is done using the aggregate aid figure, also considering that the bias on the disaggregated grants and loans coefficients is already clear from table 4. The true coefficient of aid and the size of the bias can therefore be calculated by substituting as follows:

$$b = \beta + \frac{0.093}{160.004}(0.001) \tag{3}$$

The equation provides further evidence on the presence of a small and positive bias (i.e. the second term on the LHS) due to the omission of governance, consistent with the results in table 4. However this bias is negligible and it does not substantially affect the aid estimates. Indeed the results are essentially unaltered when the ICRG indicators are included in the equation.

A caveat of this analysis using ICRG data is that any result only refer to the period 1985-2009 and they may therefore not be fully generalizable to the whole period.

5 Robustness

Having presented the main results, this section checks their robustness to a number of econometric and data issues. Despite having taken into account the effect of governance as a possible omitted factor in section 4, endogeneity remains a possible econometric issue and it is therefore discussed in further detail in section (5.1). In addition structural breaks may also be a concern,

Table 4: SR results with governance

	(1) Full sample	(2) 1985-on	(3) 1985-on	(4) 1985-on
non-tax	-0.043	0.011	-0.099	-0.072
	(-0.34)	(0.06)	(-0.75)	(-0.49)
grants	0.434***	0.469**	0.424***	0.478***
	(3.32)	(2.82)	(3.42)	(3.46)
loans	0.254***	0.350***	0.314***	0.314***
	(2.98)	(3.28)	(3.94)	(3.50)
agri	0.066	-0.077	0.059	0.062
	(0.60)	(-0.49)	(0.49)	(0.44)
manuf	0.838	0.081	0.691	0.752
	(1.61)	(0.11)	(1.25)	(1.17)
trade	0.207***	0.180***	0.176***	0.191***
	(4.71)	(3.28)	(4.33)	(4.16)
GDPpc	-0.005	0.036	-0.004	-0.003
	(-0.16)	(0.91)	(-0.14)	(-0.09)
Derg	0.001	0.006	0.016***	0.013**
	(0.45)	(1.06)	(3.28)	(2.45)
EPRDF	-0.002	.	.	.
	(-0.86)	.	.	.
ECT	-0.955***	-0.830**	-0.821***	-0.953***
	(-5.56)	(-2.94)	(-3.92)	(-3.98)
icrg			0.001***	
			(3.64)	
icrgnc				0.001**
				(2.77)
Constant	0.000	-0.004	-0.033***	-0.025***
	(0.06)	(-1.33)	(-4.01)	(-3.15)
Observations	49	25	25	25
r2	0.61	0.65	0.82	0.77

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All variables are included in first difference and the dependent variable is the tax share in all columns. The first column uses the full sample, the others use a sub-sample starting in 1985.

given the presence of three different political regimes throughout the period. While it is reasonable to assume breaks in correspondence of regime changes, tests are also performed to check other possible breaks in the relation (section 5.2). Finally, section 5.3 checks the robustness of the results to the choice of using variables as a share of GDP and it presents results using logs instead.

Additional problems that may affect the results are influential observations and non-linearities. The former in particular may be an issue in presence of exceptional events that indeed occurred in Ethiopia in the period considered, such as droughts and famines. These events may cause disruptions in tax payments and large inflows of aid, therefore potentially affecting the relation. This issue however does not appear to determine the results, that are robust to dropping the observations that are potentially problematic (for more details see results and discussion in appendix D).

The second problem may rise from the possible non-linearities in the aid-tax relation. To tackle this, quadratic terms are included in both in the LR and SR equations. Only mild evidence is found of a decreasing effect of grants, which would be consistent with the idea that the marginal benefit of aid is decreasing with the amount of aid[17].

The remainder of this section focuses specifically on endogenity and structural breaks.

5.1 Endogeneity and reverse causality

The endogeneity of aid is a common concern in the tax effort literature, as discussed in sections 2 and 3. In the cointegration framework endogeneity would be a problem only in the SR, since the LR equilibrium relation is characterized by superconsistency (Engle and Granger, 1987). In presence of endogeneity, the SR coefficients could pick up spurious effects due to omitted variables or simple correlations in which causality cannot be established.

A common way to test for endogeneity in a time series setting is the Granger causality test (Granger, 1969). When applied to the Ethiopian series, with particular attention to the aid variables, the test shows that both grants and loans Granger-cause tax. However this test suffers from several limitations such as its purely statistical nature, therefore ignoring suggestions from economic theory, and its focus on the temporal dimension only. Appendix E reports the test results and it discusses them.

This section starts by discussing possible instrumenting strategies and by providing 2SLS results. It continues by exploring the nature of endogeneity in the specific case of Ethiopia, using qualitative information including interviews to donors and government officials.

5.1.1 IV estimation

Finding a valid instrumenting strategy for aid is generally a difficult task, and it is an even harder one in a time series setting. Not only the proposed instrument

[17] These results are available from the author.

should be measured with available data for the whole period considered, but also it is required to vary in time. This difficulty is further exacerbated when the analysis focuses on a developing country, where data availability is usually more problematic and relatively long time series are rare.

Given this difficulty in finding an appropriate instrument for aid, the literature commonly uses its lagged value to decrease concerns over endogeneity (Gupta et al., 2004; Morrissey et al., 2006; Clist and Morrissey, 2011; Mkandawire, 2011). This is also the main instrumenting strategy used here and particularly two lags are used as instruments.

Table 5 reports the results of the first stage and second stage regressions respectively in the first two and last two columns. As regards the first stage, the results show clearly that the instruments are better fit to explain grants than loans. In the case of loans the proposed instruments may not be associated with the problem variable, therefore violating one of the standard conditions for a relevant IV, (Cameron and Trivedi, 2009). The weaker the association of the IV with the problem variable, the weaker the identification.

Instruments that are marginally relevant are weak instruments and they make estimation much less precise, thus inflating standard errors and resulting in less satisfactory t-statistics. The partial R^2 provides a test for weak instruments as it captures the explanatory power of the instruments once all other variables are controlled for. The partial R^2 in the grants equation is 0.3 whereas for the loans equation is 0.1, thus confirming the low validity and relevance of the IV in the latter case. A more formal test for weak instruments (Stock and Yogo, 2005) confirms that the proposed instruments are weak for loans but not for grants[18].

Having established that the IV may not be valid for loans, columns three and four instrument respectively both variables and only grants, for which the IV appears valid. When both variables are instrumented, loans becomes largely non-significant with a large standard error (0.3, almost ten times larger than the coefficient) which could be expected. The coefficient on grants is positive but also non-significant, with a more precise estimate and a smaller standard error (0.24). When only grants is instruemnted (column 4), its coefficient is positive and significant with a magnitude that is relatively similar to the OLS estimate of 0.434[19].

The 2SLS results therefore provide evidence for the robustness of the positive coefficient of grants in the SR to endogeneity, while the IV estimates are not satisfactory for loans. However the Granger causality test did not flag any

[18]The hypothesis of weak instruments cannot be rejected for the former but it can for the latter at the 10% level, with F-statistic of 1.5 for loans and 8.7 for grants, to be compared with a 10% critical value of 7.5.

[19]The coefficient estimated with 2SLS is larger than the OLS one, thus indicating a small negative bias. While the IV estimate should not be taken as a precise point estimate, it may be compared with the results obtained when including governance in section 4.3. As far as grants are concerned, the exercise with ICRG data did not show a clear indication on the sign of the bias. Therefore the ICRG and IV results are largely consistent. The fact that all estimates (OLS, IV and including ICRG) are of comparable in magnitude supports the robustness of a positive effect of grants.

Table 5: IV estimation with lagged grants and loans

	(1)	(2)	(3)	(4)
	1st stage	1st stage	2nd stage-IV	2nd stage-IV
non-tax	0.258*	-0.163	-0.090	-0.052
	(1.99)	(-0.63)	(-0.66)	(-0.45)
grants		0.109	0.378	0.545***
		(0.34)	(1.56)	(3.28)
loans	0.031		-0.035	0.274***
	(0.34)		(-0.12)	(3.55)
agri	-0.196*	0.223	0.131	0.068
	(-1.75)	(1.02)	(1.04)	(0.69)
manuf	-0.354	1.416	1.306*	0.810*
	(-0.66)	(1.42)	(1.88)	(1.73)
trade	-0.087**	-0.070	0.180***	0.226***
	(-2.19)	(-0.89)	(2.84)	(5.30)
GDPpc	0.037	-0.080	-0.018	-0.004
	(1.15)	(-1.34)	(-0.55)	(-0.17)
Derg	0.001	0.000	0.001	0.001
	(0.26)	(0.03)	(0.19)	(0.20)
EPRDF	0.000	0.004	-0.002	-0.004
	(0.14)	(0.69)	(-0.56)	(-1.35)
L.grants	-0.461***	0.409		
	(-3.01)	(1.28)		
L.loans	-0.212**	-0.164		
	(-2.36)	(-0.91)		
L2.grants	-0.308**	0.100		
	(-2.14)	(0.35)		
L2.loans	0.164*	-0.267		
	(1.78)	(-1.51)		
ECT			-0.848***	-1.019***
			(-3.54)	(-6.32)
Constant	-0.000	0.001	0.001	0.001
	(-0.15)	(0.15)	(0.46)	(0.42)
Observations	47	47	47	47
r2	0.58	0.25	0.49	0.61

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The first column reports results for the first stage regression on grants; the second one is the first stage regression with loans as a dependent variable; the third reports the second stage 2SLS regression where both aid variables are instrumented; the fourth column reports the 2SLS results when only grants is instrumented. All variables are included in first difference.

concern for loans while it indicated the possibility of reverse causality for grants (see appendix E). The combined evidence therefore suggests that estimates are fairly robust to endogeneity. Biases may occur, but they appear to be small. However it should be noted that the instruments used here suffer from limitations on validity, and therefore the issue of endogeneity is not fully tackled. These concerns can be addressed more systematically using a CVAR approach (see Mascagni and Timmis (ming)).

The IV estimates can be compared to the OLS ones (those showed in table 3) to assess the extent of endogeneity. Two tests are performed and they provide consistent results: the Hausman test and the Durbin-Wu-Hausman test[20]. The former finds that the difference in coefficients between the two methods is not systematic (p-value: 0.99), therefore supporting the OLS estimates on efficiency grounds. The latter confirms this result, failing to reject the null hypothesis of exogeneity of the aid variables with p-values of 0.29 (Durbin score chi2) and 0.41 (Wu-Hausman F statistic).

5.1.2 Qualitative evidence

In addition to econometric tests and techniques, qualitative evidence may help explaining the nature of endogeneity in Ethiopia and assessing the extent to which it is indeed a problem. This qualitative analysis is supported by interviews to government officials and donor agencies, therefore allowing for great insight into the working of the aid-tax relation in the Ethiopian context.

Mascagni (2014) shows that Ethiopia since 1960 has received aid for reasons largely other than economic ones. To the extent that aid is driven by political and strategic reasons in donor countries, it is exogenous to the Ethiopian government's decision making process that determines the tax share conditional on a number of variables.

This discussion does not aim at arguing that donors disburse money regardless of the country level conditions. Clearly they want to see 'value for money' and thus they do care about basic economic indicators as well as fiscal policy. However conditionality in Ethiopia is rather mild, due to weak enforceability and a relatively low bargaining power of donors. Strong ownership and leadership on the government's side means that policy objectives and conditions need to be in line with its priorities to be acceptable. This is largely recognised by donors as well, as confirmed in the interviews. So for example fiscal discipline and relatively low corruption are part of the Ethiopian institutional tradition and not a response to donor preferences, perhaps with the aim of receiving more aid. This situation is not particularly surprising in a country with a long institutional tradition and a history of independence from colonial powers.

An example of mild conditionality is PBS, which indeed does condition aid on service delivery, something of which the government is very enthusiastic about. Indeed this focus on services not only is in line with government's objectives, but also it allows it to gain legitimacy and popularity amongst citizens as well

[20]The first test requires homoskedasticity while the second one allows for heteroskedasticity.

as largely wiping off the discussion table broader economic policies.

Still using the example of PBS, let us consider conditionality on the revenue side. In principle PBS imposes additionality, as foreign funds have to be matched by government resources. I asked donors about the enforcement of this principle and in particular what would happen if the government does not adhere to it. They confirmed that there is no actual way to enforce additionality. However this has not been a problem in Ethiopia so far, as the Government has consistently overperformed on the revenue targets. Indeed increasing revenue is first and foremost a priority of the government thus making strict conditionality on revenue irrelevant.

As far as economic factors are concerned, they do not seem to influence aid either. For example Ethiopia has experienced very high inflation in the past 5 years and while donors have been critical of this, aid has increased. Similarly the case of liberalizations is emblematic. Although the governemnt has systematically refused to open up to donors requests in this area, there has been no consequence on aid flows.

So aid does not seem to be driven by economic factors in Ethiopia. It is interesting that even the most technical of the donor agencies involved in the interviews confirmed that it is indeed political and stretegic reasons that really matter in Ethiopia, while donors can always turn a blind eye on economic issues. Ethiopia benefits from a strategic geographical position that allows easy access to the Middle East as well as a good base for communications and military operations. For example the American Kagnew communications station, inaugurated well before the period considered here, was a crucial listening post in the Korean war, World War Two and during the Cold War (see Wrong (2005)). The strategic importance of Ethiopia was then reinforced during the Cold War and as a result of the increased terrorist threat following the attacks to US embassies in Nairobi and Dar es Salaam in 1998 and the New York attacks of September 11^{th} 2001.

To the strategic considerations is to be added the fact that Ethiopia is home to millions of African poor, being a large country and still one of the poorest in the world. All this explains why donors are reluctant to leave Ethiopia even when economic and political conditions deteriorate.

One relevant example is the aftermath of the 2005 elections. While many donors pointed to violations of human rights as well as irregularities in the election process, aid was only temporarily suspended. One immediate 'punishment' was the end of budget support that however was promptly replaced by PBS. Under this project (PBS is a project although some similarities exist with tied budget support) aid could not only be restored quickly after the interruption, but also it increased in the following years.

In conclusion, domestic economic and political factors do not seem to play a role in determining variations in aid flows. The qualitative elements discussed here indeed suggest that aid in Ethiopia seems to be only weakly or not at all determined by economic or political governance (thus potentially raising concerns about the omitted variable bias) or by the tax share itself (thus resulting in reverse causality). Instead it is more plausible to explain aid flows with the po-

litical and strategic needs of donor countries, therefore making them exogenous to the domestic decision making process.

5.2 Structural breaks

As anticipated in section 3, Ethiopian history over the period considered suggests the presence of two obvious breaks in correspondence to the changes of regime in 1974 and 1991. These 'expected' breaks are already taken into account in the main exercise (section 4) and this section justifies their inclusion. Moreover it explores them in more detail by using interaction terms between the regime dummies and other explanatory variables. In addition to the regime changes, other events may result in breaks. For example Clist and Morrissey (2011) provide cross-country evidence for a break in the aid-tax relation in the mid-80s, coinciding with the increased adoption of structural adjustment programs and conditional aid. The CUSUM and CUSUMSQ tests[21] explore the possibility of additional breaks by looking at the data rather than the context information. However no additional breaks are found except the change of regime in 1991 in the LR, and no breaks are detected in the SR. The positive results on the aid variables are largely confirmed.

Finally rolling and recursive estimation techniques[22] are used to check for parameter stability. While the parameter on grants is shown to be consistently positive both in the SR and the LR, the recursive and rolling estimations do provide negative coefficients for loans. However after the mid-80s, and surely during the EPRDF period, the estimated loans coefficients are consistently positive[23]. While these results certainly raise concerns about parameter stability, they are to be taken with caution since they are based on sub-samples of an already small sample.

The small sample also prevents a meaningful application of the Chow test[24], that is commonly used to test for the presence of breaks. Since the test requires estimation on the sub-samples, it would rely on estimates calculated on 15, 16 and 19 observations respectively for the Imperial, Derg and EPRDF periods. Besides technical concerns related to the small sub-samples, such few observations are not really compatible with the long run nature of the first step equation.

[21] The former plots the time sequence of the cumulative sum of residuals divided by the standard error of the regression. The latter plots the cumulative sum of squared residuals rescaled by the partial sum of the residual sum of squares so the last value will always be one.

[22] The recursive technique estimates the equation on the first 15 years and it then repeats the estimation gradually adding each year until the full sample is included. The recursive parameter typically shows instability in the first years, estimated with fewer observations, before stabilizing. The rolling method estimates the equation over a 15 year window starting from the first 15 years (i.e. 1960 to 1974) then progressively rolling that window year by year over the sample.

[23] While it would be tempting to take this as evidence in favor of the Clist and Morrissey (2011) argument, Ethiopia only started structural adjustment in the early 90s.

[24] The Chow test tests the null hypothesis of no structural breaks by comparing the residual sum of squares of the restricted model (i.e. whole sample) against unrestricted models estimated on the separate sub-samples.

Figure 4: CUSUM graph (LR)

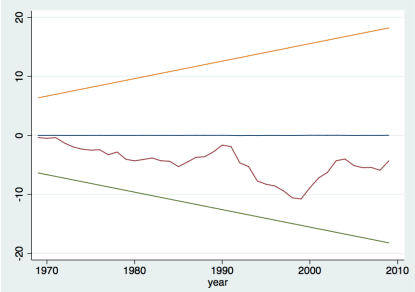
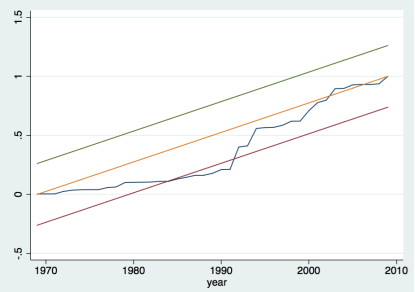


Figure 5: CUSUMSQ graph (LR)



The following sections look at the details on robustness to structural breaks respectively for the LR and SR. While they show that structural breaks need to be taken into account, which is done in the econometric exercise, they also support the robustness of the main results and particularly for grants. This may be due to the large degree of continuity in Ethiopian culture of power an administrative apparatus, despite the regime changes (see section 3 and Mascagni (2014) for a more detailed discussion).

5.2.1 1st step: the long run

A first symptom of structural breaks in the LR equation is the presence of autocorrelation in the residuals when the regime dummies are not included. Since structural breaks result in a clear pattern in the residuals, they may induce autocorrelation. Indeed when regime dummies are included the problem disappears (see results of autocorrelation tests in appendix C).

The CUSUMSQ tests confirm that the second regime shift in particular, from the Derg to the EPRDF, produces a break in the late 80s/early 90s[25]. Figures 4 and 5 show the results of the two tests, with 95% confidence bands. The exit of the residual plot from those bands is taken as evidence of the presence of structural breaks. The calculation of cumulated residuals only starts at time $k+1$ where k is the number of parameters to be estimated in the equation. It may therefore be that the failure to detect the first break in 1974 is due to the fact that only a few years of the Haile Selassie empire are included in the calculation. No other breaks are detected by these tests.

Having confirmed that structural breaks are corresponding to the expected dates (i.e. to regime changes), they can be taken into account using dummies both as independent variables and in interaction terms. The latter in particular is useful to identify the effect of breaks on the parameters of interest, namley those on the aid variables[26]. Table 6 reports the LR regression without and with

[25]This break is however not detected in the CUSUM test.

[26]This exercise is repeated for all variables in the model but the results are not reported for simplicity. The Derg dummy and its interaction terms are generally the most significant

regime shift dummies respectively in columns 1 and 2. Columns 3 and 4 add the interaction terms with the aid variables, respectively without and with the shift dummies. In columns 3 and 4 therefore the parameter on the variable refers to the reference period, i.e. Haile Selassie's regime, while the coefficients for other regimes are obtained by summing the coefficients on the original variable and on the relevant interaction term.

The results in table 6 show that both grants and loans seem to have negative, though largely non-significant coefficients in the first period and positive ones later. While loans and its interaction terms are non-significant individually, they are jointly significant ($F(3, 38) = 5.25$, p-value: 0.004). However this is not the case for grants and its interaction terms, for which the test of joint significance cannot reject the null ($F(3, 38) = 1.07$, p-value: 0.373).

Figure 6: Grant rolling estimates (LR)

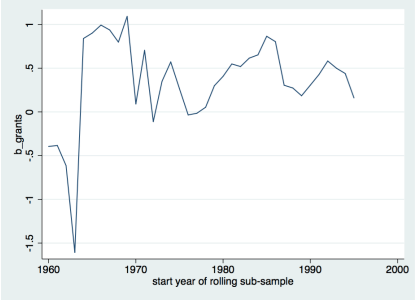
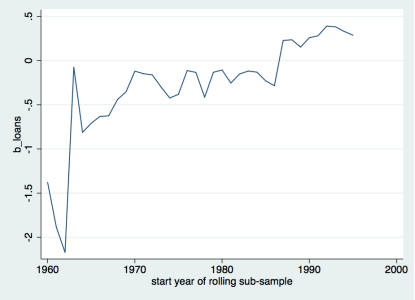


Figure 7: Loan rolling estimates (LR)



Having identified the presence of structural breaks, parameter stability is assessed across the period using a specification that includes regime dummies[27]. Rolling estimates are obtained using rolling 15-years windows starting from the period 1960-1974. The year on the horizontal axis is therefore the starting year of the rolling estimation. These estimates are then plotted in figures 6 and 7 to assess the stability of parameters. While informative, these results are to be taken with caution since each estimate is only relying on 15 observations. Both parameters show a period of instability at the beginning of the period. The coefficient on grants then settles on mostly positive values, while the one on loans becomes positive from the late 80s.

The plots of coefficients obtained using recursive estimation (not reported here for simplicity) are largely consistent with the results from the rolling estimation.

ones. The previous results are largely confirmed.

[27] The specification used as a basis for rolling and recursive estimation corresponds to column 3 in table 1.

Table 6: LR equation with structural breaks

	(1)	(2)	(3)	(4)
grants	0.199	0.325*	-0.131	0.145
	(0.90)	(1.78)	(-0.39)	(0.31)
loans	0.332***	0.413***	-0.436	-0.031
	(2.74)	(4.31)	(-0.83)	(-0.04)
non-tax	0.222	0.163	0.085	0.249*
	(1.34)	(1.26)	(0.55)	(1.71)
agri	0.081	0.052	0.104	0.036
	(1.36)	(0.87)	(1.68)	(0.54)
manuf	1.480***	0.736**	1.092***	0.792**
	(4.27)	(2.31)	(3.00)	(2.33)
trade	0.109**	0.144***	0.131***	0.131***
	(2.67)	(4.34)	(3.36)	(3.71)
GDPpc	0.006	-0.007	0.011	-0.008
	(0.30)	(-0.39)	(0.54)	(-0.39)
Derg		0.017***		0.017
		(3.61)		(1.55)
EPRDF		-0.003		-0.017
		(-0.42)		(-1.23)
grants_derg			0.807	0.014
			(1.56)	(0.03)
grants_eprdf			0.309	0.333
			(0.81)	(0.67)
loans_derg			0.879	0.095
			(1.54)	(0.12)
loans_eprdf			0.851	0.456
			(1.59)	(0.62)
Constant	-0.115	0.012	-0.146	0.033
	(-0.81)	(0.08)	(-0.94)	(0.21)
Observations	50	50	50	50
r2	0.85	0.91	0.90	0.92

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All variables are included in levels, i.e. as a share of GDP except GDPpc and GDPgap. The dependent variable is the tax share.

5.2.2 2nd step: the short run

Contrary to the LR case, in the SR no autocorerlation is detected regardless of whether the regime dummies are included or not. This suggests that structural breaks are not present in the SR and indeed the CUSUM and CUSUMSQ tests confirm it. The test results, reported in figures 8 and 9, show no evidence of structural breaks in the short run as the plotted residuals lie inside the 95% confidence interval.

Although the analysis of residuals offers no evidence of the presence of structural breaks, it may still be useful to check for the effect of interaction terms between regime dummies and the aid variables[28]. Table 7 reports the results using the regime dummies both as independent variables and in interactions terms. The coefficient on grants is positive in all periods, generally with a larger magnitude in the Haile Selassie period. Loans instead have a negative coefficient in the first period that turns positive in the remaining two periods, with the EPRDF interaction term being always significant at the 10% level. A test of joint significance of grants and loans and their respective interaction terms shows that they are both jointly significant[29].

Parameter stability is explored using rolling estimation (in this section) and recursive estimates (results not reported for simplicity) on a 15 years window[30]. Consistently with LR results, the coefficient on grants is always positive as shown in figure 10. Its magnitude decreases as the window moves to more recent years. The coefficient on loans (figure 11), also in line with the long run, shows some negative estimates between the late 60s and the early 80s when it becomes consistently positive. The coefficient of loans therefore shows more instability than the one on grants.

Figure 8: CUSUM graph(SR)

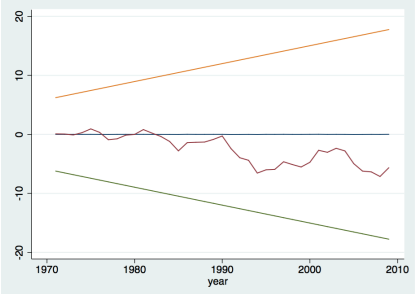
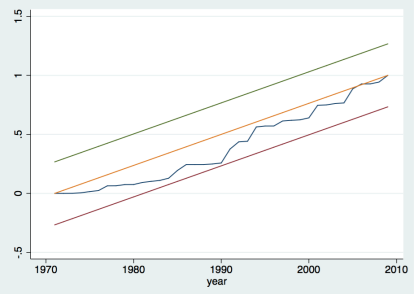


Figure 9: CUSUMSQ graph (SR)



[28] As before, interaction terms are also included for other explanatory variables although the results are not reported here. The previous results are largely confirmed and the interaction terms, as well as the dummies included as independent variables are never significant.

[29] Grants: F(3, 34) = 4.02, p-value: 0.015; loans: F(3, 34) = 4.15, p-value: 0.013.

[30] Estimates are based on the SR specification reported in column 1 of table 3.

Table 7: SR equation with structural breaks

	(1)	(2)	(3)	(4)
non-tax	-0.042	-0.043	-0.022	-0.012
	(-0.33)	(-0.34)	(-0.17)	(-0.09)
grants	0.415***	0.434***	0.637	0.490
	(3.21)	(3.32)	(1.01)	(0.74)
loans	0.241***	0.254***	-0.576	-0.640
	(2.88)	(2.98)	(-1.14)	(-1.22)
agri	0.127	0.066	0.163	0.090
	(1.32)	(0.60)	(1.55)	(0.73)
manuf	1.064**	0.838	1.177**	0.903
	(2.23)	(1.61)	(2.37)	(1.63)
trade	0.190***	0.207***	0.188***	0.207***
	(4.61)	(4.71)	(4.63)	(4.69)
GDPpc	-0.023	-0.005	-0.025	-0.006
	(-0.92)	(-0.16)	(-0.98)	(-0.21)
Derg		0.001		0.001
		(0.45)		(0.24)
EPRDF		-0.002		-0.003
		(-0.86)		(-0.95)
grants_derg			-0.016	0.088
			(-0.02)	(0.13)
grants_eprdf			-0.278	-0.082
			(-0.43)	(-0.12)
loans_derg			0.569	0.645
			(1.01)	(1.10)
loans_eprdf			0.869*	0.953*
			(1.70)	(1.78)
ECT	-0.928***	-0.955***	-0.967***	-0.973***
	(-5.49)	(-5.56)	(-5.44)	(-5.41)
Constant	0.000	0.000	0.001	0.001
	(0.27)	(0.06)	(0.53)	(0.35)
Observations	49	49	49	49
r2	0.59	0.61	0.65	0.66

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All variables are included in first difference and the dependent variable is the tax share.

Figure 10: Grant rolling estimates (SR)

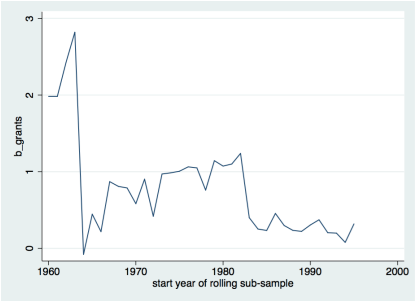
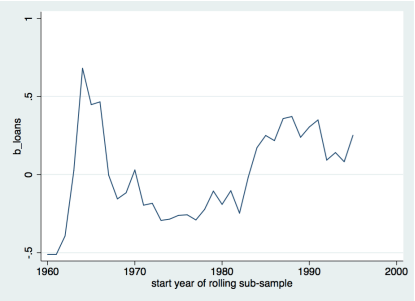


Figure 11: Loan rolling estimates (SR)



5.3 Estimation in logs

A further robustness check regards the way variables are measured, namely as shares of GDP. GDP figures are based on surveys and other estimates of economic activities. They may therefore be imprecise or change in time because of better estimation techniques. In addition GDP figures have been at the center of a heated debate in recent years, with many forign observers doubting their validity (Dercon and Zeitlin, 2009). The analysis is therefore repeated with the variables in logs instead of shares of GDP, to check whether it is indeed the denominator that drives the results.

Stationarity is tested for all variables and they all turn out to be I(1). It is therefore possible to proceed with the cointegration procedure as in the previous analysis. Table F.1 in appendix F reports the results for both the LR and SR.

The LR specification in logs still passes the Engle-Granger test for cointegration and it is therefore possible to carry out the standard two-step procedure. The results are robust to a different variable definition that excludes GDP at the denominator. Indeed previous results are largely confirmed when repeating the estimation in logs and particularly so for the aid variables that are positive and significant in all specifications. In particular a 1% increase in grants is associated with a 0.1% increase in tax revenue in the short run, while a 1% increase in loans would have a smaller effect of 0.05% on tax revenue.

Note that now the coefficient on the GDP gap is non-significant and positive, confirming the hypothesis that the negative and significant coefficient found in section 4 may be due to a built-in correlation with GDP at the denominator of the dependent variable. Indeed when tax is measured as a log, rather than a share of GDP, the GDP gap is not significant anymore.

The coefficients on other variables are largely similar to previous estimates. The only noteworthy change is the positive coefficient on agriculture, that is now significant whereas manufacturing fails to reach significance in all specifications.

6 Unpacking the effects: tax types and tax composition

Having carried out the analysis using the aggregate tax share, this section turns to the individual tax types to check if the identified effects run through a particular one of these. There are potentially various channels through which aid relates to tax revenue. For example aid may provide foreign exchange to purchase imports that, if taxed, contribute to increasing tax revenue. Aid may also provide support to tax reform aimed at increasing revenue collections from 'hard to tax' bases. By disaggregating tax revenue in tax types I am able to shed light on some of these channels. The aggregate tax share is therefore disaggregated into domestic indirect, direct and trade taxes, all of which are I(1). All three variables are taken in shares of GDP and are used as dependent variables. They are plotted in figure 12 to give a sense of their dynamics.

The LR results are reported in table 8 along with the Engle-Granger test

Figure 12: Plot of tax types

Source: author's calculations using data from the Ministry of Finance and Economic Development

statistic for cointegration, at the bottom of the table, to be compared with the MacKinnon critical values in table 2. Only the version of the test without lags is reported because no autocorrelation is detected. The direct tax and indirect tax equations pass the cointegration test in both specifications while the trade tax one does not. Table 9 therefore reports SR results, using the EG procedure, only for direct and indirect taxes. In addition results for trade taxes are reported for completeness, based on an ARIMA(0,1,0) model.

As far as the aid variables are concerned, the results generally show that their positive and significant effect on tax is mainly due to direct and trade taxes. In the long run, only loans is significant in the direct and trade tax equations, whereas grants never reach significance although the coefficients are always positive. Note that grants was significant only at the 10% level in the aggregate tax equation. In the short run however grants also becomes significant, and still positive, in both the direct and trade tax equations while loans is significant only in the former.

Focussing particularly on domestic taxes, these results are interesting because they indicate that the effect of aid may occur through stimulating the collection of 'hard to tax' handles, such as income and profits, whereas 'easy to tax' ones (i.e. consumption through indirect taxes) are already tapped[31]. This

[31] Aizenman and Jinjarak (2009) provide a cross country analysis of the effects of globalization on 'easy to collect' and 'difficult to collect' taxes. In their classification, the former category includes trade taxes and seignorage, whereas the second one includes domestic direct (income, profits) and indirect (sales, VAT) taxes. However in the case of Ethiopia indirect

Table 8: Tax types LR equations: direct, indirect and trade tax

	Direct tax		Indirect tax		Trade tax	
	(1)	(2)	(3)	(4)	(5)	(6)
grants	0.091	0.077	0.026	0.016	0.208	0.161
	(1.07)	(0.91)	(0.26)	(0.17)	(1.54)	(1.16)
loans	0.175***	0.177***	0.000	0.002	0.238***	0.246***
	(3.89)	(3.95)	(0.00)	(0.03)	(3.37)	(3.35)
non-tax	0.248***	0.267***	0.087	0.099	-0.172*	-0.112
	(4.07)	(4.60)	(1.22)	(1.47)	(-1.79)	(-1.18)
agri	-0.038	-0.058***	-0.038	-0.051**	0.129***	0.066*
	(-1.35)	(-2.79)	(-1.16)	(-2.12)	(2.89)	(1.95)
manuf	0.153		0.097		0.486**	
	(1.02)		(0.56)		(2.06)	
trade	0.009	0.003	-0.001	-0.006	0.136***	0.115***
	(0.59)	(0.19)	(-0.08)	(-0.34)	(5.56)	(4.97)
GDPpc	0.002	0.002	-0.001	-0.001	-0.008	-0.007
	(0.28)	(0.30)	(-0.11)	(-0.10)	(-0.62)	(-0.54)
Derg	0.004*	0.004*	0.002	0.002	0.011***	0.012***
	(1.85)	(1.92)	(0.64)	(0.68)	(3.24)	(3.24)
EPRDF	-0.006*	-0.007**	-0.007*	-0.008**	0.010*	0.006
	(-1.72)	(-2.19)	(-1.73)	(-2.06)	(1.81)	(1.13)
Constant	0.015	0.033	0.047	0.059	-0.049	0.010
	(0.22)	(0.53)	(0.62)	(0.81)	(-0.48)	(0.10)
Obs	50	50	50	50	50	50
r2	0.92	0.92	0.63	0.63	0.79	0.77
E-G test	-5.691	-5.883	-6.097	-6.185	-3.949	-3.840

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All variables are in levels, i.e. as shares of GDP except GDPpc and GDPgap. The dependent variables are: direct taxes (columns 1 and 2), indirect taxes (columns 3 and 4), and trade taxes (columns 5 and 6).

may also explain the non-significant coefficient in the indirect tax equation. In this context the introduction of the VAT, that was strongly supported by the international financial institutions, may play a minor role because it largely replaced pre-existing sales taxes. Figure 12 shows that in the early 90s indirect and trade taxes (both considered 'easy to collect') increased substantially while direct taxation lagged behind. This would support the view that income taxes are more difficult to raise, because of administrative capacity but also political constraints. In this context the stimulus effect of aid may happen both through technical assistance aimed at increasing local capacity; and through the interantional backing of donors and international financial institutions that may give legitimacy to politically difficult reforms.

Since trade is certainly considered 'easy to tax', this argument would however not be informative in explaining the positive coefficient of loans in the trade tax equation. That posive effect may instead be due to the increased availability of foreign exchange allowed through aid, that may fuel imports and in turn trade taxes. Note that the positive effect of aid on trade taxes is documented also elsewhere in the literature (Benedek et al., 2012).

Another interesting result is the coefficient on agriculture, that was consistently non-significant in all specifications notwithstanding its expected negative effect. Agriculture is generally found to have a negative effect on domestic direct and indirect taxes, as expected, while it has a positive effect on trade taxes, consistently with other studies in the literature (Aizenman and Jinjarak, 2009). The negative effect becomes significant once manufacturing is dropped both in the LR and SR equations, thus improving precision in the estimates[32]. In the aggregate tax equation this expected negative effect of agriculture is probably offset by the positive one through trade taxes, thus making the coefficient largely non-significant. The positive sign on agriculture in the trade tax equation may be due to export taxes that have been largely reliant on agricultural exports, most notably coffee. Indeed Ethiopian exports are largely agricultural products and up to 2002/2003, when export taxes were eliminated, they represented an important source of revenue.

Finally, an alternative approach to gaining insights into tax composition is to use a variable that indicates the share of domestic taxes (i.e. direct and domestic indirect) to total revenue. Tax composition, as defined here, is a central issue in developing countries where trade taxes still represent the lion's share of tax revenue. This is indeed the case in Ethiopia where in the last year of the sample trade taxes still contributed over 40% to total tax revenue, and over the whole period that share peaked at 56%. The international community has been concerned with the possibility of a shift towards domestic taxes and this idea has been at the basis also of the adoption of VAT by an increasing number of countries (Keen and Ligthart, 2002; Keen and Simone, 2004). Tax composition is used as a dependent variable, however this exercise did not yield

taxes have been always considered 'easy to collect'. Therefore it seems legitimate to consider only direct taxes on income and profits as 'hard to tax' in the discussion here.

[32]The equation excluding agriculture is not reported here for simplicity but it yields coefficients of the same sign and significance for other variables.

Table 9: Tax types SR equations: direct, indirect and trade tax

	Direct tax		Indirect tax		Trade tax	
	(1)	(2)	(3)	(4)	(5)	(6)
non-tax	0.110**	0.117**	0.036	0.028	-0.201**	-0.140
	(2.08)	(2.35)	(0.49)	(0.40)	(-2.15)	(-1.45)
grants	0.089*	0.088*	0.020	0.021	0.209**	0.188
	(1.74)	(1.74)	(0.27)	(0.30)	(2.12)	(1.37)
loans	0.140***	0.144***	-0.011	-0.015	0.033	0.064
	(4.11)	(4.43)	(-0.23)	(-0.33)	(0.56)	(0.88)
agri	-0.033	-0.047*	-0.089	-0.070*	0.157***	0.039
	(-0.73)	(-1.69)	(-1.41)	(-1.78)	(2.94)	(1.00)
manuf	0.084		-0.118		0.750*	
	(0.40)		(-0.39)		(1.87)	
trade	0.039**	0.040**	0.027	0.026	0.094***	0.096***
	(2.32)	(2.43)	(1.14)	(1.11)	(4.62)	(4.24)
GDPpc	0.012	0.014	0.018	0.014	-0.025	-0.005
	(1.00)	(1.38)	(1.03)	(0.98)	(-1.50)	(-0.35)
ECM	-0.726***	-0.739***	-0.927***	-0.916***		
	(-4.85)	(-5.11)	(-5.57)	(-5.65)		
Constant	-0.000	-0.000	-0.000	-0.000	0.001	0.000
	(-0.23)	(-0.31)	(-0.23)	(-0.17)	(0.93)	(0.06)
Dummies	yes	yes	yes	yes	yes	yes
Obs	49	49	49	49	49	49
r2	0.58	0.58	0.50	0.49	-	-

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All variables are included in first difference. The dependent variables are: direct taxes (columns 1 and 2), indirect taxes (columns 3 and 4), and trade taxes (columns 5 and 6). In the case of direct and indirect taxes, the results are fromt he second step of the EG procedure. For trade taxes the results are obtained using and ARIMA(0,1,0) model. Regime dummies for the Derg and EPRDF period are included in all equations.

satisfactory results or additional insight. It is therefore not reported here.

7 The positive aid-tax relation: interpretation

The main result of the analysis is a strong rejection of the disincentive effect that is sometimes argued in the context of the aid effectiveness debate. I show that aid is instead positively associated with the tax share. The positive effect is generally confirmed for both grants and loans, with some differences that are discussed later in this section.

This postive relation implies that aid and tax are not treated as substitutes but that instead increases in aid are associated with increases in taxes. Indeed aid may influence tax collection by contributing to improving the tax administration (thus decreasing administrative costs, for example by promoting the introduction of IT systems) and by providing international legitimacy for tax reform (thus decreasing political costs). Its coefficient in the reduced form equation is therefore the result of a negative effect due to possible substitution and a positive one through the reduction of the costliness of raising tax. The empirical results from Ethiopia suggest that the latter positive effect is larger than the possible negative one.

Turning to aid heterogeneity, while grants and loans appear to have both a positive effect on tax revenue, some differences emerge from the analysis. The idea developed in the literature (Gupta et al., 2004) that loans may stimulate tax mobilisation more than grants because of repayment, is partly confirmed. In the long run loans are always significant while grants only reach the 10% level in one specification. Despite this, no evidence can be provided for a LR detrimental effect of grants on tax effort. The coefficient on grants is consistently positive, also when rolling estimates are provided in section 5.2 both for the LR and SR.

In the short run both aid variables are significant above the 5% level and still positive, with grants having a larger coefficient than loans. This may suggest that grants are used mainly to release short run constraints in the administration, thus allowing a contemporaneous increase in tax revenue. Moreover both coefficients (grants and loans) are robust to the inclusion of governance in the equation, to capture institutional factors that are excluded from the main equations. The coefficient of grants in particular is also robust to taking into account endogeneity with IV estimation, while a satisfactory instrument cannot be found for loans.

A disaggregation of tax revenue into direct, indirect and trade taxes, shows that the positive effect of aid occurs primarily through direct and trade taxes. In the case of direct taxes, the explanation may be related to the beneficial effect of aid in supporting the increased mobilisation of 'hard to collect' taxes such as those on income and profits. Instead 'easy to collect' taxes such as those on consumption may already be tapped, thus not benefiting as much from aid. This explanation would not be appropriate in the case of trade taxes however, where the positive effect may be due to the provision of foreign exchange through aid.

In the Ethiopian context, the positive effect can be explained by two ele-

ments in particular: capacity building and matching[33]. Indeed capacity is the most widely cited constraint to tax revenue mobilisation in interviews, both with government officials and with donors. Aid is likely to have an effect on this capacity constraint thanks to technical assistance, additional resources for tax administration and external advisors. Ethiopia has received external advice and support in the area of taxation across the whole period, including assistance from the US, UN and WB during the Imperial period. More recently taxation has remained one of the few policy areas of strong cooperation between donors and the government, while on other macroeconomic policies disagreement often occurs. For this reason the positive effect of aid through various capacity building activities is fully consistent with the Ethiopian context. By making the administration more efficient, aid may also contribute to making taxation less costly thus resulting in an increase in tax revenue.

At least three examples exist today in Ethiopia of this foreign role in the area of taxation. The first one is the Public Sector Capacity Building program (PSCAP), one of the 6 themes of which is precisely taxation. Secondly PBS funds, amongst other things, projects aimed at raising awareness on tax issues, with the final objective of increasing compliance. Indeed at least two interviewees mentioned that a basic knowledge about the budget process and what taxes are used for can help overcoming compliance and evasion problems. Finally the IMF has had a crucial advisory role both in the 2002 reform, that amongst other things introduced the VAT, and in the 2008 reforms.

The second possible explanation is that the government is matching foreign resources with internal ones. This may happen because of two reasons. The most obvious one is that donors may require the governemnt to do so. While this is indeed the case, for example under PBS that includes an additionality clause, it does not seem that donors are really driving the matching effect. This is confirmed by the consistent over-performance of the Government of Ethiopia with respect to the revenue targets under PBS.

The second reason for matching is related to the specific historical and cultural characteristics of Ethiopia. Ethiopia is the only country in Africa to have never been colonized and this is deeply rooted in the national culture and pride. Not only does this translate in great policy ownership and a heavy weight attached to aid conditions and external influence; but also it implies a tradition of independence and a feeling that the country can 'stand on its own feet'. In this sense, as emerged also in a few interviews, having to rely on foreign funding to deal with domestic development challenges is a shame for Ethiopians. It means that the country is not able to 'feed its population' and to respond to its basic needs. This sentiment is clear in some of the speeches of the late Prime Minister Meles Zenawi who ruled Ethiopia for 20 years, and it is at the basis of the willingness of the Government to match foreign resources with domestic ones.

[33]Other possible explanations can be provided for a positive effect of aid on tax revenue, which I however believe being secondary to capacity building and matching. For example aid may increase marketable surplus due to better infrastructure and thus increase revenue through indirect taxes on goods. It could also imply an increase in aid-related imports that, while largely duty-free, may still imply an increase in trade tax revenues.

In fact aid may even set an incentive for tax revenue mobilisation not only because of the possible dependence and shame coming with it, but also because it makes the ambitious development project happen. Again, this emerged in many interviews. Ethiopia has ambitious development projects that would be hard to implement only with domestic resources. The fact that aid is available makes the project realistic and feasible, thus giving momentum to domestic revenue mobilisation efforts.

The positive relation between aid and tax is reinforced by the fact that they are not treated as substitutes in Ethiopia but rather as complements. In addition to the elements discussed previously (i.e. independence, national pride, external influence) aid volatility and unpredictability also prevent them to be substitutes. A plot of the three-year moving average of the tax and aid series shows very clearly that the latter is a much more volatile source of revenue than tax (see figure 2). Many interviewed government officials argue that it is not a sustainable source of funding because its little reliability makes it difficult to plan ahead.

8 Conclusions and policy implications

This paper provided evidence on a positive relation between aid and tax in Ethiopia. The use of in depth qualitative information, that represents a novelty in the literature, is used to explain the results in the specific context of Ethiopia. Given the country-specific historical and political context, a positive relation between aid and tax seems particularly well grounded in Ethiopia and examples are provided throughout the paper.

The main conclusion of the analysis is the presence of a strong and positive relation between aid and taxation, that is particularly robust for grants (particulalry to structural breaks and endogeneity). The positive results on grants and loans are robust to the inclusion of governance, that is often omitted from time series analyses of tax effort. As far as aid heterogeneity is concerned, grants and loans both have a positive sign although they present some differences in their LR and SR effects that are discussed in section 7.

The first policy implication is therefore that in the case of Ethiopia a crowding out effect of aid should not be a source of concern for donors. I find no evidence in Ethiopia of such a negative effect of aid on tax effort or of the opportunistic behavior that would result in a substitution between tax and aid. Instead the results support the idea that aid has been effective in supporting the administration and in strengthening institutions in the field of taxation.

In addition, the inclusion of other tax determinants in the analysis allows drawing a few broader conclusions and policy implications.

Firstly the effect of manufacturing is very large and significant, particularly in the long run equation. Its coefficient exceeds unity in some specifications, implying that a percentage point increase in the share of manufacturing in GDP results in an even higher increase in the tax share. For example the coefficient of 0.7 (in the preferred LR equation, column 3 in table 1) suggests that for

each percentage point increase in the manufacturing share, 70% of it goes into increased tax. Besides arguments related to the high visibility and accessibility of manufacturing firms, this large effect may be due to the fact that firms act as tax collectors, as they transfer to the state taxed levied on employees' salaries and on consumption. This result, by indicating the strong revenue generating potential of the manufacturing sector, underlines that industrial development has fiscal advantages in addition to the better documented benefits in terms of employment, formality and structural transformation.

Secondly trade displays the expected positive and significant coefficient in virtually all specifications. This is a confirmation of the major role that trade plays as a tax base in many developing countries, and particulalry in Ethiopia. Indeed Ethiopia is one of the African countries with the highest share of trade taxes in total revenue (OECD and African Development Bank Group, 2010). This clearly has implications in terms of trade liberalization. While the increase in trade flows would bring about more revenue, a decrease in tax rates may have an adverse revenue impact at least in the short term. It is therefore important to understand the balance between these two opposite effects and to provide for alternative sources of revenue in order to alleviate the possible negative effects of liberalization, particularly in countries like Ethiopia that are highly reliant on trade taxes.

Thirdly agriculture does not seem to have a significant effect on the aggregate tax share due to its contrasting effects on different tax types. While a positive effect is found for trade taxes, the expected negative one occurs for domestic direct and indirect taxes. This result is particularly worrying in the context of the efforts of developing countries to switch away from trade taxes towards domestic revenue (Keen and Simone, 2004). The agricultural sector may represent a possible opportunity for future reform to increase its revenue generation capacity, especially as commercial agriculture develops. However the subsistance and remoteness that characterize this sector are likely to remain major constraints to tax revenue mobilisation in Ethiopia.

Finally, also in line with much of the literature, the analysis largely fails to find a significant effect of GDP per capita.

Appendices

A Summary of variables

Table A.1: Summary of variables

Name	Definition	Mean	St. dev.	Min	Max
tax	Total tax revenue as a share of GDP	.082	.024	.039	.124
dirtax	Direct tax revenue as a share of GDP	.028	.012	.010	.048
indtax	Indirect tax revenue as a share of GDP	.022	.006	.010	.036
tradetax	Trade tax revenue as a share of GDP	.032	.011	.014	.061
grants	Foreign grants as a share of GDP	.020	.012	0	.047
loans	Foreign loans as a share of GDP	.020	.016	.002	.092
trade	Imports and export as a share of GDP	.187	.111	.085	.504
agri	Agriculture share of GDP	.591	.099	.415	.784
manuf	Manufacturing share of GDP	.048	.010	.026	.065
GDPpc	Log of constant GDP per capita	6.946	.154	6.733	7.494
GDPgap	Percentage deviation of GDP from trend (Hodrick-Prescott filter)	0	.134	-.312	.416
taxcomp	Direct and indirect domestic taxes as a share of total tax revenue	.600	.085	.439	.780

B Stationarity

Table B.1: Results of Dickey-Fuller GLS test for trend-stationarity: levels

Variable	1 lag	2 lags	3 lags	4 lags	5 lags
tax	-2.39	-2.79	-2.68	-2.24	-2.01
dirtax	-2.48	-3.23	-2.43	-1.78	-2.04
indtax	-2.08	-1.91	-1.67	-1.62	-1.67
tradetax	-2.20	-2.35	-2.59	-2.70	-2.02
grants	-4.34	-2.88	-2.25	-2.80	-2.58
loans	-3.06	-2.42	-3.69	-2.57	-2.67
aid	-3.06	-2.59	-3.62	-2.78	-2.64
trade	-1.14	-1.19	-0.41	-1.38	-1.27
agri	-2.90	-1.97	-2.17	-2.05	-3.01
manuf	-1.63	-1.39	-1.29	-1.29	-1.42
GDPpc	-0.77	-0.37	-1.19	-1.14	-1.10

H_0: non trend-stationarity

Critical values (5%): -3.202 (1 lag), -3.159 (2 lag), -3.108 (3 lag), -3.052 (4 lag), -2.992 (5 lag).

Table B.2: Results of Dickey-Fuller GLS test for stationarity: 1st difference

Variable	1 lag	2 lags	3 lags	4 lags	5 lags
tax	-3.60	-3.40	-3.53	-3.26	-2.62
dirtax	-2.97	-3.70	-4.32	-3.08	-2.26
indtax	-4.52	-4.30	-3.57	-2.93	-2.81
tradetax	-4.17	-3.19	-2.77	-3.29	-2.58
grants	-8.90	-6.50	-3.90	-3.67	-2.94
loans	-6.06	-3.42	-4.19	-3.28	-3.67
aid	-6.25	-3.66	-4.28	-3.73	-3.56
trade	-4.27	-5.13	-2.65	-2.49	-2.44
agri	-6.32	-3.97	-3.43	-2.12	-2.70
manuf	-4.74	-4.07	-3.43	-2.79	-2.67
GDPpc	-4.65	-2.05	-1.95	-1.80	-0.80

H_0: non stationarity

Critical values (5%): -2.285 (1 lag), -2.259 (2 lag), -2.230 (3 lag), -2.199 (4 lag), -2.167 (5 lag).

Table B.3: Results of Kwiatkowski-Phillips-Schmidt-Shin test for trend-stationarity: levels

Variable	0 lags	1 lag	2 lags	3 lags	4 lags	5 lags
tax	0.34	0.19	0.14	0.12	0.11	0.10
dirtax	0.56	0.30	0.21	0.17	0.15	0.14
indtax	0.50	0.32	0.25	0.21	0.18	0.16
tradetax	0.33	0.18	0.13	0.11	0.09	0.09
grants	0.13	0.11	0.10	0.09	0.08	0.08
loans	0.17	0.11	0.09	0.08	0.08	0.08
aid	0.11	0.07	0.06	0.06	0.05	0.06
trade	0.90	0.49	0.35	0.27	0.23	0.20
agri	0.28	0.17	0.13	0.11	0.10	0.09
manuf	0.85	0.46	0.33	0.27	0.23	0.20
GDPpc	0.67	0.37	0.27	0.21	0.18	0.16

H_0: series is trend-stationary

Critical values: 10%: 0.119, 5% : 0.146, 1% : 0.216.

Table B.4: Results of Kwiatkowski-Phillips-Schmidt-Shin test for stationarity: 1st difference

Variable	0 lags	1 lag	2 lags	3 lags	4 lags	5 lags
tax	0.05	0.06	0.05	0.06	0.06	0.07
dirtax	0.10	0.08	0.07	0.06	0.07	0.08
indtax	0.04	0.06	0.07	0.09	0.10	0.11
tradetax	0.05	0.05	0.05	0.05	0.05	0.06
grants	0.02	0.03	0.05	0.06	0.06	0.07
loans	0.05	0.05	0.07	0.07	0.08	0.11
aid	0.05	0.06	0.08	0.08	0.10	0.12
trade	0.30	0.31	0.31	0.38	0.38	0.39
agri	0.05	0.05	0.06	0.08	0.08	0.07
manuf	0.31	0.26	0.26	0.27	0.29	0.29
GDPpc	0.58	0.46	0.46	0.42	0.37	0.36

H_0: series is stationary

Critical values: 10%: 0.347, 5% : 0.463, 1% : 0.739.

Table B.5: Results of Clemente, Montanes, Reyes unit root test with two structural breaks: levels

variable	break 1	break 2	test stat
tax	1974	1990	-3.89
	0.01	0.39	.
dirtax	1976	1990	-3.85
	0.00	0.47	.
indtax	1977	1989	-6.76
	0.00	0.00	.
tradetax	1987	1991	-3.55
	0.06	0.00	.
grants	1982	1998	-6.69
	0.00	0.00	.
loans	1998	2003	-2.12
	0.00	0.00	.
aid	1979	1998	-4.99
	0.00	0.00	.
trade	1990	2002	-3.65
	0.00	0.00	.
agri	1976	1995	-2.68
	0.10	0.01	.
manuf	1976	1989	-3.58
	0.01	0.02	.
GDPpc	1981	2002	-2.76
	0.10	0.00	.

H_0: presence of unit root, i.e. non stationarity
Critical value for last column statistic: -5.49 (5%)
P-values reported under break year

Table B.6: Results of Clemente, Montanes, Reyes unit root test with two structural breaks: 1st diff

variable	break 1	break 2	test stat
tax	1987	1990	-8.61
	0.00	0.00	.
dirtax	1987	1991	-5.63
	0.00	0.00	.
indtax	1964	1990	-9.81
	0.05	0.55	.
tradetax	1975	1992	-8.87
	0.22	0.12	.
grants	1966	1999	-7.70
	0.01	0.62	.
loans	1992	2000	-8.11
	0.86	0.00	.
aid	1998	2002	-7.01
	0.00	0.00	.
trade	1990	2001	-8.51
	0.00	0.20	.
agri	1983	1989	-3.54
	0.01	0.01	.
manuf	1982	1989	-9.71
	0.08	0.61	.
GDPpc	1983	2001	-3.76
	0.30	0.00	.

H_0: presence of unit root, i.e. non stationarity
Critical value for last column statistic: -5.49 (5%)
P-values reported under break year

C Autocorrelation tests

Table C.1 reports results for autocorrelation tests and particularly the Durbin-Watson test, that require all variables to be strictly exogenous, as well as the alternative Durbin and Breusch-Godfrey tests that instead do not require strict exogeneity. The tests are run for the first three specifications of table 1, and they show that indeed autocorrelation is not detected once structural breaks are modeled in specifications 2 and 3.

Table C.1: Autocorrelation tests (first step)

specification	test	stat	p-value	result
(1)	Durbin-Watson	1.00	n.a.	reject
	Durbin's alternative	11.102	0.001	reject
	Breusch-Godfrey	10.862	0.001	reject
(2)	Durbin-Watson	1.837	n.a.	inconclusive
	Durbin's alternative	2.26	0.790	can't reject
	Breusch-Godfrey	0.093	0.760	can't reject
(3)	Durbin-Watson	1.927	n.a.	inconclusive
	Durbin's alternative	0.008	0.930	can't reject
	Breusch-Godfrey	0.010	0.920	can't reject

H_0 for all tests = no serial correlation.

Table C.2: Autocorrelation tests (second step)

dummies	test	stat	p-value	result
no	Durbin-Watson	1.641	n.a.	inconclusive
	Durbin's alternative	2.469	0.116	cannot reject
	Breusch-Godfrey	2.918	0.088	cannot reject
yes	Durbin-Watson	1.665	n.a.	inconclusive
	Durbin's alternative	2.070	0.150	cannot reject
	Breusch-Godfrey	2.596	0.107	cannot reject

H_0 for all tests = no serial correlation.

'Dummies' in the first column refers to the regime shift dummies.

D Influential observations

Two methods are used to identify outliers and influential observations: deviations from the three-year moving average of aid (main variable of interest) and standardized residuals. The first method allows the identification of the three highest negative deviations in 1977, 1983, and 2005; and the three highest positive ones in 1998, 1978 and 1993. These dates can easily be matched with relevant events in Ethiopia and in particular:

- 1977 and 1978 are marked by the total withdrawal of US aid in response to the newly established Derg regime. In 1978 the Ogaden war involved a scale-up of Soviet support that largely substituted US aid.
- 1983 appears as a large negative deviation probably due to the high moving average in that period, due to the famine erupting right after that time, in 1984, that attracted much international support.
- 1993 is the year in which Ethiopia engages in its first Structural Adjustment Program that brought to the country a large amount of loans.
- 1998 is the year in which the war with Eritrea starts.
- 2005 is marked by the elections and particularly the following rupture in relations with donors that brought to the withdrawal of direct budget support.

Secondly, the three lowest and highest standardized residuals from the LR and SR equations are identified as follows:

- In the long run equation the largest standardized residuals are in 1975, 2008 and 1984 (negative); and 1963, 2003 and 1983 (positive).
- The largest standardized residuals produced by the short run equation are in 2005, 1991 and 1984 (negative); and 1963, 2001 and 1997 (positive).

The years identified are easily traced back to historical events such as the 1984 famine, the 2005 elections and the reform momentum in the early 60s following the attempted coup. Moreover high residuals are observed in correspondence of the regime change in 1991, the war with Eritrea (1997 and 2001), and the tax reforms of 2008 that were accompanied by a more assertive approach to tax enforcement.
Table D.1 reports the long run and short run results for three different specifications. The first one is the preferred specification including all observations, and the LR and SR results are reported in the columns marked (1). The second specification excludes the years identified using the 3-years MA of aid, and the results are reported in the columns marked (2). Finally the observations with highest standardized residuals are excluded and the results reported in the columns marked (3) are obtained.
The coefficients on the aid variables appear to be robust to the exclusion of these

observations. The coefficient on *grants* remains similar, both in magnitude and significance, both in the SR and LR. The coefficient on loans also seems robust to dropping outliers although it is not significant in the SR equation once observations with high standardized residuals are dropped (SR equation, column 3). The results on other variables are largely in line with the previous results.

Table D.1: Tax regression without influential observations

	(1) LR	(2) LR	(3) LR	(1) SR	(2) SR	(3) SR
grants	0.325*	0.290	0.246*	0.434***	0.460***	0.565***
	(1.78)	(1.60)	(1.87)	(3.32)	(3.49)	(5.12)
loans	0.413***	0.416***	0.400***	0.254***	0.273***	0.073
	(4.31)	(4.34)	(5.90)	(2.98)	(3.09)	(1.04)
non-tax	0.163	0.220	0.254**	-0.043	0.040	-0.046
	(1.26)	(1.51)	(2.50)	(-0.34)	(0.32)	(-0.55)
agri	0.052	0.070	0.013	0.066	0.001	0.105
	(0.87)	(1.20)	(0.30)	(0.60)	(0.01)	(1.12)
manuf	0.736**	0.792**	0.445*	0.838	0.299	0.937**
	(2.31)	(2.43)	(1.93)	(1.61)	(0.61)	(2.33)
trade	0.144***	0.172***	0.080***	0.207***	0.304***	0.260***
	(4.34)	(4.08)	(3.11)	(4.71)	(5.98)	(7.26)
GDPpc	-0.007	-0.015	0.026	-0.005	0.018	-0.007
	(-0.39)	(-0.81)	(1.66)	(-0.16)	(0.61)	(-0.26)
Derg	0.017***	0.015***	0.020***	0.001	0.001	0.003
	(3.61)	(3.05)	(5.90)	(0.45)	(0.34)	(1.62)
EPRDF	-0.003	-0.003	0.002	-0.002	-0.002	-0.002
	(-0.42)	(-0.47)	(0.38)	(-0.86)	(-0.95)	(-0.76)
Constant	0.012	0.049	-0.174	0.000	-0.000	-0.001
	(0.08)	(0.36)	(-1.41)	(0.06)	(-0.20)	(-0.62)
Obs	50	44	44	49	43	43
r2	0.91	0.93	0.96	0.61	0.70	0.80

t statistics in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns marked (1) report the main results for allowing comparison; columns marked (2) report the results obtained by dropping observations with high deviations from the aid 3-year moving average; columns marked (3) report the results on a sample where observations with high standardized residuals were dropped.

E Granger causality test

The test devised by Granger (Granger, 1969) is a common way to test for endogeneity in time series. Originally the test is based on a VAR between two variables but it can be applied to three or more variables simply by including them in the underlying VAR. The test is based on a criterion of incremental predictive power: if Y is better predicted by past values of X and Y together than of Y alone, then X is said to Granger-cause Y. The null hypothesis of non-causality (in the Granger sense) is tested with a Wald test on the joint significance of the coefficients on the lags of X in the equation explaining Y. This test is rather limited for at least two reasons. First, it is purely statistical as it does not take into account any suggestion by social or economic theory. Secondly it only explores endogeneity in terms of its temporal dimension and it is therefore of little help when endogeneity is due to omitted factors. Table E.1 reports the results from the test, including all the variables of the empirical model. The first eight rows of table E.1 report the results for the tax equation (i.e. tax on the LHS). The remaining rows report, for completeness, the results for tax in the equations explaining all other variables. Other results are excluded for the sake of clarity. The test shows that both grants and loans Granger-cause tax, though the latter only at the 10% level. Tax does not Granger cause any of the explanatory variables at the 5% level, but the null of non-causality can be rejected at the 10% level in the grants equation. However a different specification of the model including only tax and the aid variables fails to support Granger-causality from tax to grants even at the 10% level, underlining the weakness of this result.

Table E.1: Results of Granger causality test

Equation	tested	chi2	p
tax	nontax	1.49	0.47
	grants	8.52	0.01
	loans	5.03	0.08
	agri	0.92	0.63
	manuf	2.88	0.24
	trade	0.86	0.65
	gdppc	6.16	0.05
	ALL	37.37	0.00
nontax	tax	2.75	0.25
grants	tax	5.58	0.06
loans	tax	0.66	0.72
agri	tax	2.66	0.26
manuf	tax	0.26	0.88
trade	tax	3.41	0.18
gdpps	tax	3.20	0.20

F Results from estimation in logs

Table F.1: Estimation in logs: LR and SR

	(1) LR	(2) LR	(3) SR	(4) SR
grants	0.068***	0.068***	0.108***	0.108***
	(3.38)	(3.31)	(5.74)	(5.66)
loans	0.084**	0.085**	0.049**	0.047**
	(2.60)	(2.54)	(2.20)	(2.05)
nontax	0.082	0.079	0.027	0.021
	(1.54)	(1.37)	(0.61)	(0.44)
agri	0.741***	0.750**	0.056	0.122
	(2.73)	(2.64)	(0.17)	(0.34)
manuf	0.152	0.156	0.084	0.106
	(1.04)	(1.03)	(0.47)	(0.57)
trade	0.545***	0.548***	0.534***	0.529***
	(8.06)	(7.68)	(5.46)	(5.31)
GDPpc	-0.759*	-0.790	0.000	0.000
	(-1.84)	(-1.65)	(0.68)	(0.36)
GDP gap		0.017		0.047
		(0.13)		(0.45)
ECT			-0.945***	-0.927***
			(-6.09)	(-5.72)
Constant	-1.781	-1.698	-0.005	-0.004
	(-1.14)	(-1.00)	(-0.19)	(-0.19)
Dummies	yes	yes	yes	yes
Obs	50	50	49	49
r2	1.00	1.00	0.77	0.77

t statistics in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Columns 1 and 3 report the preferred specification while columns 2 and 4 add the GDP gap variable. This variable in particular is included to check if a negative coefficient can still be found once GDP is not on the LHS anymore.

References

Aizenman, J. and Y. Jinjarak (2009). Globalization and developing countries - a shrinking tax base? *Journal of development studies*.

Benedek, D., E. Crivelli, S. Gupta, and P. Muthoora (2012). Foreign aid and revenue: still a crowding out effect? IMF Working Paper WP/12/186, International Monetary Fund.

Bird, R. M., J. Martinez-Vazquez, and B. Torgler (2008). Tax effort in developing countries and high income countries: the impact of corruption, voice and accountability. *Economic analysis and policy*.

Bornhorst, F., S. Gupta, and J. Thornton (2008). Natural resource endowments, governance, and the domestic revenue effort: evidence from a panel of countries. IMF Working Paper WP/08/170, International Monetary Fund.

Brautigam, D., O.-H. Fjeldstad, and M. Moore (2008). *Taxation and state building in developing countries: capacity and consent*. Cambridge University Press.

Bwire, T. (2013). *Aid, fiscal policy and macroeconomy of Uganda: a CVAR approach*. Ph. D. thesis, University of Nottingham.

Cameron, A. C. and P. K. Trivedi (2009). *Microeconometrics using STATA*. STATA Press.

Carter, P. (2013). Does foreign aid displace domestic taxation? *Journal of Globalization and Development*.

Chelliah, R. J. (1971). Trends in taxation in developing countries. Staff papers, International Monetary Fund.

Chelliah, R. J., H. J. Baas, and M. R. Kelly (1975). Tax ratios and tax effort in developing countries. Staff papers, International Monetary Fund.

Clist, P. and O. Morrissey (2011). Aid and tax revenue: signs of a positive effect since the 80s. *Journal of international development*.

Department for International Development (2011). Bilateral aid review: technical report. technical report, DFID.

Dercon, S. and A. Zeitlin (2009). Rethinking agriculture and growth in Ethiopia: a conceptual discussion. report, Department for International Development.

Drummond, P., W. Daal, N. Srivastava, and L. E. Oliveira (2012). Mobilizing revenue in Sub-Saharan Africa: empirical norms and key determinants. IMF Working Paper WP/12/108, International Monetary Fund.

Engle, R. F. and C. W. J. Granger (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica*.

Ezemenari, K. E. E. K. K., E. Kebede, and S. Lahiri (2008). The fiscal impact of foreign aid in Rwanda: A theoretical and empirical analysis. Policy research working paper, World Bank.

Franco-Rodriguez, S. (2000). Recent developments in fiscal response with an application to Costa Rica. *Journal of international development*.

Franco-Rodriguez, S., O. Morrissey, and M. McGillivray (1998). Aid and the public sector in Pakistan: Evidence with endogenous aid. *World Development*.

Furtado, X. and W. J. Smith (2007). Ethiopia: aid, ownership and sovreignty. Geg working paper 2007/28, Global Economic Governance Programme.

Ghosh, M. (1999). Structural break and unit root in macroeconomic time series: Evidence from a developing economy. *The Indian Journal of Statistics*.

Ghura, D. (1998). Tax revenue in Sub-Saharan Africa: Effects of economic policies and corruption. Imf working paper, International Monetary Fund.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross spectral methods. *Econometrica*.

Greene, W. H. (2008). *Econometric Analysis*. Pearson Education.

Gupta, A. S. (2007). Determinants of tax revenue efforts in developing countries. IMF Working Paper WP/07/184, International Monetary Fund.

Gupta, S., B. Clemens, A. Pivovarsky, and E. Tiongson (2004). Foreign aid and revenue response: Does the composition of aid matter? In S. Gupta, B. Clemens, and G. Inchauste (Eds.), *Helping countries develop: the role of fiscal policy*. International Monetary Fund.

Heller, P. (1975). A model of public fiscal behaviour in developing countries: Aid, investment and taxation. *American Economic Review*.

Hisali, E. and J. Ddumba-Ssentamu (2013). Foreign aid and tax revenue in uganda. *Economic Modelling*.

Imam, P. A. and D. F. Jacobs (2007). Effect of corruption on tax revenues in the Middle East. IMF Working Paper WP/07/270, International Monetary Fund.

Keen, M. and J. E. Ligthart (2002). Coordinating tariff reduction and domestic tax reform. *Journal of international economics*.

Keen, M. and A. Simone (2004). Tax policy in developing countries: some lessons from the 1990s and some challenges ahead. In S. Gupta, B. Clemens, and G. Inchauste (Eds.), *Helping countries develop: the role of fiscal policy*. International Monetary Fund.

Le, T. M., B. Moreno-Dodson, and N. Bayraktar (2012). Tax capacity and tax effort. extended cross-country analysis from 1994 to 2009. Policy Research Working Paper 6252, World Bank.

Leuthold, J. H. (1991). Tax shares in developing economies. a panel study. *Journal of Development Economics*.

Machado, R. (2009). Aid and fiscal policy in Nicaragua: a fiscal response analysis. *Journal of international development*.

MacKinnon, J. G. (2010). Critical values for cointegration tests. Queen's Economics Department Working Paper 1227, Queen's University.

Mahdavi, S. (2008). The level and composition of tax revenue in developing countries: evidence from unbalanced panel data. *International Review of Economics and Finance*.

Martins, P. M. G. (2010). Fiscal dynamics in Ethiopia: the cointegrated VAR model with quarterly data. Economics Department Working Paper Series 9-2010, University of Sussex.

Mascagni, G. (2014). Tax revenue mobilisation in Ethiopia. Phd thesis, University of Sussex, Department of Economics.

Mascagni, G. and E. Timmis (forthcoming). Fiscal effects of aid in Ethiopia: Evidence from CVAR applications. Research paper, CREDIT.

Mavrotas, G. and B. Ouattara (2006). Aid disaggregation and the public sector in aid-recipient economies: some evidence from Cote D'Ivoire. *Review of Development Economics*.

Mkandawire, T. (2011). On tax effort and colonial heritage in Africa. *Journal of development studies*.

Morrissey, O., O. Islei, and D. M'Amanja (2006). Aid loans versus aid grants: are the effects different? CREDIT Research paper 06/07, CREDIT.

OECD (2008). Governance, taxation and accountability: issues and practices. DAC guidelines and reference series, OECD.

OECD and African Development Bank Group (2010). African Economic Outlook 2010: public resource mobilization and aid. report, OECD and African Development Bank Group.

Osei, R., O. Morrissey, and T. Lloyd (2003). Modelling the fiscal effects of aid: An impulse response analysis for Ghana. CREDIT Research paper 03/10, CREDIT.

Osei, R., O. Morrissey, and T. Lloyd (2005). The fiscal effects of aid in Ghana. *Journal of international development*.

Perron, P. and T. J. Vogelsang (1992). Testing for a unit root in a time series with a changing mean: corrections and extensions. *Journal of business and economic statistics*.

Schaffer, M. (2010). egranger: Engle-Granger (EG) and Augmented Engle-Granger (AEG) cointegration tests and 2-step ECM estimation. Stata command.

Stock, J. H. and M. Yogo (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews and J. H. Stock (Eds.), *Identification and inference for econometric models: essays in honor of Thomas Rothenberg.* Cambridge University press.

Stotsky, J. G. and A. WoldeMariam (1997). Tax effort in Sub-Saharan Africa. IMF Working Paper WP/97/107, International Monetary Fund.

Tait, A., B. Eichengreen, and W. Gratz (1979). International comparisons of taxation for selected developing countries. Staff papers, International Monetary Fund.

Tanzi, V. (1992). Structural factors and tax revenue in developing countries: a decade of evidence. In I. Goldin and A. Winters (Eds.), *Open Economies: Structural Adjustment and Agriculture*. Cambridge University press.

Teera, J. M. and J. Hudson (2004). Tax performance: a comparative study. *Journal of international development*.

Wrong, M. (2005). *I didn't do it for you: how the world used and abused a small African nation.* HarperCollins Publishers.

Zivot, E. and D. Andrews (1992). Further evidence on the Great Crash, the oil-price shocks, and the unit-root hypothesis. *Journal of Business and Economic Statistics*.