UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2022/107

Aid, taxes, and government spending

A heterogeneous co-integrated panel analysis

Abrams M.E. Tagem*

September 2022

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: A substantial amount of aid to developing countries is given to the government, or goes through the budget, meaning it should have an impact on government fiscal behaviour (particularly on government spending). The few existing empirical studies on the effects of aid on government spending neglect variable time-series properties, cross-country (recipient) heterogeneity, and the potential for cross-country correlation. This paper examines the impact of foreign aid and taxes on government spending for a sample of 69 developing countries over 1980–2013, taking account of dynamics characterizing fiscal data, cross-country heterogeneity, and the distorting impact of cross-section dependence. Our econometric approach addresses these problems by applying the Pesaran (2006) common correlated effects mean group estimator. We show that spending, net aid (as well as variants including grants and loans), and taxes comprise an equilibrium (cointegrated) relation. Our results provide robust evidence of a positive, long-run (as well as short-run) association between aid and spending. On average, the aid coefficients are positive but smaller than the tax coefficients, indicating that, in the long run and short run, taxes have a stronger association with expenditures than aid.

Key words: aid, cross-section dependence, heterogeneity, tax revenue

JEL classification: C23, E62, F35

Acknowledgements: The author is grateful to Oliver Morrissey, Simon Appleton, Patrick Marsh, Markus Eberhardt, Tim Lloyd, Joerg Breitung, and Zakari Mumuni for useful comments. The author also thanks seminar participants at the GDRI International Development Economics Conference 2016. This work was supported by the Commonwealth Scholarship Commission [grant number 2015111]. The usual disclaimer applies.

This is an abridged version of CREDIT Research Paper 17/02, available at https://www.nottingham.ac.uk/credit/news/papers/1702.aspx. All citations have been updated to reflect the most recent versions of the papers.

* UNU-WIDER, tagem@wider.unu.edu

This study is published within the UNU-WIDER project GRD – Government Revenue Dataset. It is part of UNU-WIDER's Domestic Revenue Mobilization (DRM) programme, which is financed through specific contributions by the Norwegian Agency for Development Cooperation (Norad).

ISSN 1798-7237 ISBN 978-92-9267-241-6

https://doi.org/10.35188/UNU-WIDER/2022/241-6

Typescript prepared by Siméon Rapin.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Aid still represents a vital source of revenue in many developing countries as the poorest of these countries do not have sufficient resources to finance their development needs (Herzer and Morrissey 2013). The central premise in giving more aid is that it should spur economic growth and development in recipient countries, and that is usually the basis from which the effectiveness of aid is judged. Morrissey (2012) states that, as most aid given to a country goes to the government or finances services that would otherwise be a demand on the budget, it should impact on government spending. Substantial amounts of aid do not actually go through the budget: aid may be in the form of donor-funded projects or technical assistance (Morrissey 2012; Van de Sijpe 2013) which does not even leave the donor country. Nonetheless, the amount of aid going through recipients' budgets is large, which should have a direct effect on the level and composition (allocation to different expenditure headings) of government spending (Morrissey 2015a). This proposes a prior hypothesis to investigate: is there empirical evidence that aid is related to total spending (and its components) over time on average?

That is the fundamental question addressed in this paper, building on insights from multi-factor models in nonstationary panels (Kapetanios et al. 2011; Pesaran 2006). We employ common factor models by Pesaran (2006) to estimate the average long-run effect of aid and taxes on spending in a sample of 69 countries using annual data covering the period 1980 to 2013. The common factor approach allows for cross-section correlations in the data, *created* by global shocks that affect countries to varying degrees, and *represented* by unobserved common factors. Our focus in the empirical part of this paper is on obtaining average estimates of the impact of aid and taxes on spending using these common factor models that allow for heterogeneous fiscal impact across countries and variable dynamics. They are also robust to endogeneity created by unobserved common factors, as well as omitted variable bias.

Most studies on the effects of aid on spending (Remmer 2004; Morrissey et al. 2011) and a few early fiscal response models (Heller 1975; Mosley et al. 1987) have been cross-country models in which the dynamics between aid and fiscal aggregates (tax/GDP and spending/GDP ratios) are assumed to be the same across countries. The general pattern of findings in these studies is that over time, aid increases total spending in recipient countries, albeit spending not increasing by the full amount of aid (i.e., aid is not fully additional). Additionality is difficult to establish (Morrissey 2015a), not least because recipient governments are not aware of all aid available to finance their public spending. While donors can implement stringent measures to ensure their aid is fully spent on intended sectors, recipients may respond by reducing their own domestic revenue allocated to spending in those sectors, resulting in aid not being fully additional (spending may not increase at all if the increase in aid is completely offset by a reduction in domestic funds allocated to the specific sector). It is also possible that spending in a particular sector increases by more than the amount of aid, especially in pro-poor social sectors (for example health and education) where aid-funded investments create claims on future government spending. As our focus is on average effects across countries, we do not delve further into the contentious issue that is aid additionality. Nevertheless, we argue that the studies are restrictive in modelling the impact of aid on spending and base our arguments on three key aspects.

First, these studies apply standard panel approaches which typically ignore the time-series properties of the data, thus ignoring potential long-run (levels) relationships between aid and fiscal variables. Söderbom et al. (2014: 394) state that 'erroneously' assuming variable stationarity has more severe implications for empirical inference than assuming variable non-stationarity. Fiscal variables are usually trending (nonstationary) in the long run, hence in a levels specification a

mixture of stationary and nonstationary variables, or a mixture of nonstationary variables of different order may lead to unreliable results (Herzer and Morrissey 2013). Our approach considers the time series properties and dynamics of the data, permitting us to make credible claims about long-run (equilibrium) and short-run relationships between aid and fiscal aggregates. Specifically, we test for the existence of a long-run equilibrium (cointegrating) relationship between aid, taxes and spending.

Second, as Roger (2015) emphasizes, these cross-country estimates are based on the stringent *homogeneity* assumption, that the effects of aid on government spending are the same for all countries in the respective samples. *Homogeneity* imposes the coefficients determining how aid impacts on government spending, the data generating process (DGP hereafter), to be the same across countries. Countries may be in the same region, and even be at the same stage of economic development but have fundamental institutional differences that influence the fiscal impact aid would have. Incorporating such heterogeneity is fundamental in estimating any potential effects aid might have on government spending. Barring the few cross-country studies mentioned above, all fiscal response models (FRMs hereafter) are country-specific (Franco-Rodriguez et al. 1998; Osei et al. 2005). In this paper we build on this country-specific, time series analysis by estimating the effect of aid and taxes on spending (and its components), over time on average, in a panel of 69 developing countries, allowing for those effects to differ across countries.

Third, cross-country approaches suffer from the potential endogeneity of aid. Conditional on the level of spending that can be domestically financed (which itself can be influenced by external or internal shocks), aid tends to go to those countries with *ex ante* lower expenditure (hence greater need for aid). This creates an endogeneity problem which standard instrumental variable methods can address. Nonetheless, Temple (1999) documents the difficulty in finding variables that qualify as instruments, alongside the spuriousness of estimates when the instruments are weak or invalid.

Closely related to the issue of endogeneity is cross-section dependence. Internal and external factors that influence recipients' spending and taxation capacities, and potentially the amount of aid they receive, create interdependencies across countries. This means in standard panel data approaches the country variable series, as well as residuals from country-specific regressions, will be correlated. Ignoring such correlations results in inconsistent and biased estimates (Chudik and Pesaran 2015). Hence in this study we employ nonstationary panel methods that allow for cross-section correlation and test for the existence of a long-run (equilibrium) relationship between aid, taxes and spending. We then test for exogeneity in domestic fiscal variables and aid, providing evidence on how recipients and donors react to deviations from the budgetary equilibrium.

Specifically, our findings are three-fold. First, the results provide evidence of a long-run equilibrium (cointegrating) relationship between aid, taxes and spending, akin to a domestic budget equilibrium. The average long-run effect of aid on spending is positive, and robust to variable and residual correlation, outliers, and omitted variables. This long-run impact is quite small, considerably smaller than the average long-run impact of taxes. This suggests that in the long-run, expenditure patterns are driven mainly by taxes, consistent with results in the fiscal effects literature (Mascagni and Timmis 2017). Second, aid has a significant long-run and short-run impact on spending in least developed countries (particularly sub-Saharan African countries). These are the most aid-dependent countries where the political costs of tax outweigh the political costs of aid. Third, aid is weakly exogenous, meaning there is no donor 'disbursement rule' that considers recipients' budget situation. The level of aid to recipients is independent of their fiscal situation.

The rest of the paper is organized as follows: Section 2 provides a conceptual framework from which hypotheses will be tested; emphasizing theoretical and econometric issues faced in the empirical analysis, as well as the importance of cross-section dependence and recipient

heterogeneity in econometric modelling. Section 3 provides a review of the literature on aid and spending, and fiscal response modelling in general. Section 4 presents a brief discussion on the data used for this analysis, and some charts while Section 5 sets out the empirical model specifications. Section 6 presents results for tests for cross-section dependence and unit roots, estimates for the impact of aid and taxes on spending, as well as results from exploratory analysis and the analysis of weak exogeneity. Section 7 concludes.

2 Conceptual framework: dynamics and hypotheses

2.1 Theoretical issues

There has been growing interest in modelling the dynamic relationship between aid and domestic fiscal aggregates; studies referred to as fiscal response models (FRMs), which draw heavily on the work of Heller (1975). Basically, governments raise revenue from different sources (for example, taxes, aid and sometimes domestic borrowing) and allocate them to different expenditures (for example investment or recurrent expenditures) in a bid to meet some revenue and expenditure targets (Lloyd et al. 2009). Governments have targets for expenditures and revenue (including aid) and their fiscal behaviour is an attempt to meet these targets, subject to a budget constraint. Thus, the decision-makers are assumed to act in a rational, utility-maximizing manner (McGillivray and Morrissey 2001). While FRMs have been fundamental in providing the basic motivation and intuition on the interplay between aid and domestic fiscal variables, they are fraught with theoretical and empirical limitations.

First, on theoretical grounds, there is no consensus on the precise form the public-sector decision-maker's utility function should take (Lloyd et al. 2009). As a result, FRMs assume that government's utility function takes the form of a perfectly symmetric loss function. Adopting a perfectly symmetric loss function implies undershooting and overshooting revenue and expenditure targets equally lead to a reduction in utility. This is counterintuitive because decision-makers' preferences are potentially asymmetric. One would expect the government to suffer utility losses only when they are unable to attain their pre-set targets; but instead maximize utility in situations where they surpass their expected targets. This is an inherent symmetry problem which FRMs cannot address.[1]

Second, FRMs are estimated using three-stage least squares (3SLS) and so are difficult to estimate, interpret and are highly sensitive to the data used (McGillivray and Morrissey 2001). Third, for FRMs to be estimated it is necessary to estimate budgetary targets (revenue and expenditure targets), but there is no accepted theory regarding how governments form revenue and expenditure targets (Morrissey 2012). Fourth, the FRM is inherently static and thus empirical applications make no attempt to distinguish long-run and short-run relationships involving aid and domestic fiscal variables. They ignore the intertemporal changes in macroeconomic characteristics of recipient countries which are unlikely to be stable over time (Lloyd et al. 2009); such characteristics influencing the impact aid might have on other fiscal aggregates. Moreover, aid can play a dual role of relaxing the budget constraint in the short run, while also forming part of the long-run budgetary

[1] Feeny (2006) proposes a utility function that allows for asymmetries—overshooting equally as bad as undershooting all targets, overshooting some targets being preferable to undershooting others, and undershooting some being preferable to overshooting—and shows that incorporating asymmetries has no major implications for empirical inferences as the reduced form and structural equations are like those derived from perfectly symmetric loss functions.

planning process. Such distinguishable long-run and short-run effects are not accommodated in the FRM.

While the theoretical FRM is the underlying motivation for looking at spending, taxes and aid together, we just provide a conceptual framework for the dynamics between foreign aid, taxes and spending; based on a government budget identity which could form the basis for testing hypotheses. In the underlying budget identity, all revenues and borrowing must equal all expenditures:

$$Revenue + Aid + Borrowing = Expenditures \quad (1)$$

Where revenue includes tax and non-tax revenues (for example, receipts from central banks), borrowing includes domestic and foreign borrowing (i.e., from international private markets), aid includes grants and loans while expenditures consist of government capital and recurrent expenditures.[2] Equation (1) is based on the underlying accounting identity, which is not predictive of the effects aid might have on domestic fiscal variables (particularly expenditures). Aid is posited to affect domestic fiscal variables, in a manner that can only be determined empirically (Lloyd et al. 2009). At the risk of stating the obvious, Equation (1) is not a theory that states how expenditures are determined; just that all revenues and borrowing must equal all expenditures.

To generate any testable hypotheses about the direction of the effects of aid and taxes on spending, we impose a structural relationship, which can be interpreted as a behavioural representation.[3] Hence we adopt a structural relationship of the form:

$$exp_{it} = \alpha_i + \beta_{i1}(aid)_{it} + \beta_{i2}(tax)_{it} + u_{it} \quad (2)$$

where exp_{it}, aid_{it} and tax_{it} are respectively total government spending, net aid and non-resource tax revenue (all as percentages of GDP). Based on Equation (2) we can explicitly test the hypotheses of aid leading to an increase or a reduction in spending, as well as the impact of tax revenue on spending. The β_i coefficients in Equation (2) represent the cross-country average effects of aid and taxes on spending respectively and they vary by country; a point we will emphasize in Section 2.4. We consider Equation (2) to be the long-run (equilibrium) relationship, the relationship of primary interest in most empirical studies. This equation can be made to include dynamics (for example the inclusion of a lagged dependent variable as well as lags of explanatory variables) and factor structures (created by common shocks which cause cross-section dependence) but with implications for estimation.

As aid is still a considerable share of GDP and accounts for large portions of government spending in most developing countries, it should have a direct financing impact on the level and composition of government spending (Morrissey 2012). Aid can also have an indirect impact on spending through donors' policy conditions (for instance in the 1980s, structural adjustment loans were promised to countries that reduced their government expenditures). Taxes have a direct effect on spending, determined by changes in tax rates and the tax base (Lloyd et al. 2009).

---

[2] We abstract from seigniorage in this analysis. While printing money generates finance for domestic expenditures, some countries in the sample (for example, CFA franc countries) have their exchange rates pegged to more developed countries, making them unable to print more money.

[3] In estimating any structural relationship of the effects of aid and taxes on spending, we omit non-tax revenue and borrowing. This is so we do not end up estimating an accounting identity.

While we expect a positive relationship, it is conceivable that there might be a negative relationship between aid and spending for two reasons. First, it may be a result of reverse causality (as described in the introduction). Second, donors' conditionality may result in reductions in spending. Most developing countries had very large public sectors in the late 1980s and early 1990s so the structural adjustment programmes (hereafter, SAPs) of the World Bank and IMF emphasized the retrenchment of public sectors as one of the prerequisites for increased aid (Williams 1994). structural adjustment aid (SAPs aid) was promised to countries that cut (or at least made attempts to cut) their public sector, notwithstanding the fact that there was no consensus on the expected size of the retrenched public sector in recipient countries. In this way donor conditionality may lead to reductions in expenditure in the short run. This intuition, however, applies only to countries that were stagnating and struggling to adjust, and maintained large and inefficient public sectors (for example, Kenya and Côte d'Ivoire). Contrarily, countries like Uganda that suffered economic collapse and were then recovering from disaster were expected to increase their expenditures after receiving more aid; with a view to rebuilding the economy.

Of equal importance is incorporating off-budget aid (proxied here by technical cooperation) into the analysis. Van de Sijpe (2013) finds that technical cooperation takes up a big share of education and health aid to developing countries (and is also a huge share of total aid). Significant amounts of aid are in the form of donor-funded projects or technical assistance (Morrissey 2012; Van de Sijpe 2013) which does not even leave the donor country. Such aid does not go through recipients' budgets but can still generate fiscal responses from recipients. Insofar as public-sector decision-makers implement the policies stipulated by technical assistance consultants, technical cooperation can generate a fiscal response from recipients.

2.2 Econometric issues

In some countries, fiscal variables like government expenditures and tax revenue show high degrees of persistence, albeit with positive and/or negative trends, meaning in the long run they are nonstationary processes (Lloyd et al. 2009; Eberhardt and Teal 2011). The concept of non-stationarity is closely linked to the order of integration of variable series, which determines the number of times the variable series needs to be differenced to achieve stationarity. If after first differencing a variable series becomes stationary, that variable series is said to be integrated of order one, i.e. I(1). This means the first differenced series will itself be I(0). While variable non-stationarity is a property characterizing macroeconomic data for most countries, the precise order of integration of the variable series should be a feature of the sample, not a global property (Pedroni 2007). This means variable (non)stationarity should be investigated for each specific sample; and not assumed to always exist in every dataset.

In the fiscal response context, the order of integration of variables is particularly salient. Aid, for instance, performs two distinct, but not mutually exclusive, fiscal roles in developing countries. First, aid may form part of recipients' domestic budget planning processes, in which case it will be nonstationary, and enter the long-run I(1) relationship (cointegrating relationship like Equation (2)). Second, it may be used to just relax the budget constraint, in which case it will enter the I(0) short-run relationship (Lloyd et al. 2009). In this case, aid substitutes for borrowing from private markets. This distinction, then, corresponds to the econometric notions of long-run equilibrium (representing the budgetary planning process of which aid plays an important role) and the short-run dynamic adjustments (represented by aid simply relaxing the budget constraint).

In statistical terminology, a cointegrating relationship (long-run equilibrium) between fiscal variables (expenditures and taxes) and aid exists when a linear combination of such I(1) variables have errors that are stationary, I(0). That is, regressing nonstationary fiscal variables like aid and taxes on nonstationary expenditures is valid if and only if the resulting error terms are stationary

(i.e., u_{it} should be stationary). Nonstationary errors from Equation (2) result in spurious regressions (Granger and Newbold 1974).[4] Occasionally, the observed long-run evolution deviates from its path but short-run 'error corrections' in the system ensure that it returns to its long-run equilibrium path (Hendry 1995). Eberhardt and Teal (2011) state that this long-run relationship can be the same for all countries in the sample (i.e., homogenous cointegration); or alternatively, each country may follow its own long-run trajectory (i.e., heterogeneous cointegration). Furthermore, for countries with stationary variable series, the problem of non-cointegration and spurious regression does not arise (Eberhardt and Teal 2011).

Discussion of non-stationarity in fiscal variables, as well as accommodating dynamics in the relationship between fiscal variables, relates to the notion of exogeneity. Once dynamics are incorporated into Equation (2), particularly the inclusion of a lagged dependent variable, the strict exogeneity assumption may no longer hold. This allows for the possibility of feedback between variables in the fiscal equilibrium, meaning past fiscal imbalances determine the amount of aid recipients receive and the amount of tax revenue they raise. In statistical terminology such reverse causality translates to the concept of weak exogeneity, which is of great importance in the fiscal response literature. In a time-series context, if the variables are nonstationary and form an equilibrium relationship (i.e., they are cointegrated), then exogeneity of aid and other variables can be tested.

When weak exogeneity tests are applied to domestic fiscal variables, the tests indicate which of the fiscal variables adjust in light of fiscal disequilibria (deficit/surplus) to restore equilibrium. When such tests are applied to the level of aid, they offer insights into the disbursement behaviour of donors. If donors respond to domestic fiscal imbalances when allocating aid, this will imply aid is endogenously determined. If on the other hand, donors do not respond to fiscal imbalances in their allocation decisions but aid influences other aggregates in the system, aid is said to be weakly exogenous or 'long-run forcing' (Lloyd et al. 2009).

Likewise, the (non)stationarity of tax revenue indicates its potential effects on expenditures. Despite the important fiscal roles aid plays in developing countries, government spending is driven mainly by domestic revenues in most of those countries (with tax revenue being the most important). Indeed, there is country-specific evidence that tax revenue is more important for government spending in developing countries (Mascagni and Timmis 2017 for Ethiopia; Bwire et al. 2017 for Uganda). As domestic revenues (represented solely by tax revenue in this study) are typically trending, they are expected to impact positively on (and sustain) domestic expenditures in the long run. This implies a cointegrating relationship between taxes and spending. While non-stationarity is a necessary condition for tax revenue to have a long-term impact on spending, it is, however, not a sufficient condition. Tax revenues are dependent on several recipient-specific factors[5] that might make it highly unpredictable, with such variations making it approximate a stationary series. In such instances, revenues will merely have a short-run impact on spending.

Though aid impacts on government spending in recipient countries, aid also has an impact on tax revenue (and domestic revenue mobilization in general), which, as emphasized, impacts on

[4] Non-stationarity of the error term leads to inconsistency in estimation. The covariance between the nonstationary error term and the nonstationary regressor(s) does not converge to zero even as $T \to \infty$ (Coakley et al. 2006). Thus, the estimator does not converge to the true parameter value but to a random number (Coakley et al. 2006).

[5] This is essentially what tax effort models aim to uncover: the factors determining the tax/GDP ratio in developing countries. Typical variables included in such regressions are the share of agriculture in recipients' GDP and exports (proportion of GDP). These two variables are particularly susceptible to volatility in developing countries, resulting in tax revenue volatility as well.

spending. Some aid is given for tax administration and public sector management while some is given for physical capital and human capital development. Data on aid given for specific 'observable or quantifiable' purposes like human capital and physical capital development, public finance management and public sector policy (obtainable from the OECD's Creditor Reporting System database) are now available. However, such data are unsuitable for time series analysis as they are available only for the recent past. Notwithstanding the novelty of aid disaggregation, there is a huge body of research on the impact of aid on tax revenue which uncovers many channels through which aid can influence taxes. However, there is no consensus and results are sensitive to empirical strategy.

Nonetheless, interrelations between aid and taxes should be accounted for in any empirical estimation of the effects of aid and taxes on spending. Omitting tax revenue from Equation (2) will attribute all the effects on spending to aid, ignoring the effects of taxes and biasing the coefficient on aid upwards. This justifies a parsimonious approach of including both aid and taxes. Besides, taxes can act as a proxy for any observed or unobserved time-varying omitted variables that influence spending (for instance exports, imports and recipient policy) and if cointegration is found in Equation (2), it will justify the inclusion of both aid and taxes in the model (implying no important nonstationary variables have been omitted).

2.3 Cross-section dependence in the data

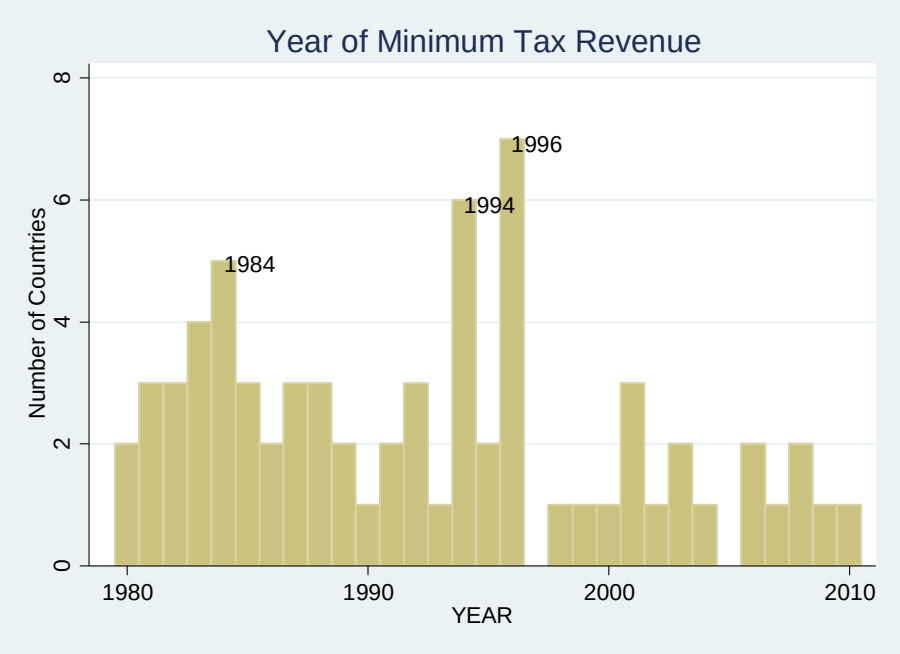
Global events like the oil price shocks of the 1970s and the 2008-09 financial crisis potentially affect all countries, albeit to varying degrees (Coakley et al. 2006). These common shocks induce unobserved time-varying heterogeneity across countries, which in turn introduces cross-section dependence between regression error terms and variable series. Such time-varying heterogeneity can lead to inconsistency in standard panel estimators (Eberhardt and Teal 2011; Pesaran 2006). 'Common' here refers to the fact that the shocks potentially affect all countries but in varying degrees. The common factors giving rise to these shocks can be *strong* factors with more widespread effects; like the oil shocks of the 1970s or the more recent financial crisis. Alternatively, the source of the shock can be *weak* factors like devaluation of the CFA franc in 1994 and the Arab Spring in 2011, which simply represent local spill-over effects. Cross-section dependence may be error cross-section dependence (in which case the error terms of different units of data are correlated) and/or variable cross-section dependence (in which case the shocks affect variable series of different countries).

Over the years, there has been an increase in economic and financial integration of countries; resulting in strong interdependencies across such countries. It is not a far-fetched idea, then, to assume cross-section dependence across countries as *common* shocks affect variable series across these countries differently. The economic mechanisms creating such time-varying unobserved heterogeneity are important as they are indicative of the kind of cross-section dependence in the data (strong vs. weak cross-section dependence; error vs. variable cross-section dependence).

Strong shocks like the global recession of the 1980s or the 2008-09 financial crisis had adverse effects on the amounts of aid donors could disburse. Jones (2015) states that in the wake of the financial crisis of 2008, donors faced severe economic slowdowns, with *small* donors (like Norway and Sweden) suffering most, hence reducing the amount of aid they disbursed. Nevertheless, *large* donors like the US and the UK still recorded increased disbursements during and after the financial crisis. Aid funds still represent a huge source of income in developing countries, especially in sub-Saharan Africa (SSA hereafter); and almost all these countries receive aid funds from both *small* and *large* donors. Hence shocks affecting the disbursement capacity of these donors will invariably affect the amount of aid developing countries receive.

Weaker shocks like the Arab Spring in 2011 and the devaluation of the CFA franc in 1994 represent local spill-over effects. For example, the devaluation affected mainly the countries of the CFA franc zone and their neighbouring countries (through geographic proximity and interactions between economic agents); so, they are merely representative of local spillover effects. One of the consequences of the devaluation was the retrenchment of the public sector in CFA countries, resulting in a reduction in their tax bases (as the formal sector became smaller in these countries), hence a reduction in tax revenues. Such reductions in tax revenues lead to reductions in total government revenues, hence government spending.

Figure 1: Minimum tax/GDP distribution (1980–2013)



Source: author's calculations based on GRD data.

Figure 1 highlights cross-section dependence in tax/GDP *troughs* across the 69 countries in the sample. The histogram shows the years in which each country recorded its lowest tax/GDP rate. More than one quarter of the countries in the sample reached their minimum tax revenues in three years; 1984, 1994 and 1996. Closer inspection of the data revealed that the drop in tax revenues in the aforementioned years was as a result of currency devaluations (for the CFA franc zone countries), economic (trade) liberalization and political instability (internal and external conflict).

Guillaumont et al. (1999) state that SSA has a higher level of *primary* instability (political and economic) than other developing regions. Such instabilities can easily spill over to other countries in the region, depending on the proximity of these neighbouring countries to the country of primary instability. Proximity here refers, but is not strictly restricted, to the actual distance between the two countries. It can also be a common colonial heritage, cultural affinity (language and religion), as well as bilateral trade volumes. Murdoch and Sandler (2002) discuss spatial war spillovers, whereby civil wars in a particular country may lead to the destruction of infrastructure and capital in neighbouring countries, as well as creating disincentives for foreign direct investment (FDI) in those neighbouring countries. Gyimah-Brempong and Traynor (1999) find that political instability negatively affects growth, particularly through its effect in discouraging investment in such unstable countries (and potentially in neighbouring countries as well). Moreover, resources which can be spent on development and infrastructure projects are instead spent on national

security and defence, resulting in stagnant growth in the fighting countries (and possibly their neighbours too). These all result in cross-section dependence in the data.

In addition to the economic mechanisms creating the cross-section dependence, the econometric implications of cross-section dependence are of equal importance. Plausibly, shocks creating cross-section dependence may create factors that need differencing to achieve stationarity (i.e., the unobserved shocks create common factors that may be nonstationary). Such I(1) common factors cause the variables not to cointegrate, creating spurious regressions (Coakley et al. 2006). If the factors are stationary (i.e., they are time-specific or relatively stable over time), then unbiased estimates can be obtained simply by augmenting the OLS regressions with a full set of time dummies or with a linear time trend. This, however, implicitly assumes that the shocks have identical effects on each country (Eberhardt and Teal 2011; Coakley et al. 2006), an assumption this paper aims to further investigate. If the factors are nonstationary then they become part of the cointegrating relationship (Banerjee and Carrion-i-Silvestre, 2017), and a full set of time dummies and a linear trend can capture parameter heterogeneity but may misspecify the true nature of the evolution of the common factors (Eberhardt and Teal 2011). Therefore, methods that allow for the possibility of nonstationary common factors should be used.

2.4 Allowing for heterogeneity in the fiscal impact of aid

Lloyd et al. (2009) state that there is a high degree of heterogeneity in government fiscal behaviour among developing countries, with fiscal effects of aid differing from one country to another (Franco-Rodriguez et al. 1998). While countries in a particular region (for example, sub-Saharan Africa) may have similar structural characteristics (relating primarily to their level of political and economic development), they have fundamental institutional differences that influence the impact aid will have on spending in those countries. Moreover, they are also heterogeneous in terms of their resource endowments, country size and population. Hence it becomes misleading to pool together all countries in a regression equation, assuming common dynamics. Such institutional differences and country-specific characteristics will ensure that each country has its own long-run equilibrium trajectory (hence country-specific fiscal equilibrium) and its own short-run dynamics.

Eberhardt and Teal (2011) argue that neglecting parameter heterogeneity in econometric analysis has more severe implications if observable variables (in our case, aid, taxes and spending) and unobservable (time variant and/or time-invariant) variables are nonstationary, resulting in the breakdown of the cointegrating relationship between aid, taxes and spending. Generally, a pooled regression equation (with common parameters for all countries) in levels will have nonstationary errors if the actual DGP differs by country (i.e., the true coefficients determining how aid and tax revenue influence spending differs across countries), and also if observable variables are nonstationary (Eberhardt and Teal 2011). With reference to Equation (2), a pooled regression equation will have error terms that contain one or more of:[6]

$$(\beta_{i1} - \mu_1)aid_{it} \qquad (\beta_{i2} - \mu_2)tax_{it} \tag{3}$$

where μ_1 and μ_2 are the common (the same for all units of data) regression coefficients for aid and taxes respectively, while β_{i1} and β_{i2} are the *true* country-specific parameters. Evidently, each of the terms in Equation (3) is a linear combination of nonstationary processes, hence the nonstationary errors themselves. Eberhardt and Teal (2011: 140) state that even in cases where there is heterogeneous cointegration (i.e., aid, taxes and spending cointegrating in each country

[6] This exposition draws heavily on Eberhardt and Teal (2011).

regression), the pooled equation does not cointegrate. Hence pooled estimation will not yield the mean of the cointegrating parameters across countries.

Heterogeneity in the time-varying unobservables (i.e., the true nature of evolution of the unobservables) should also be incorporated in analysis. Analogous to the above point, a pooled regression equation, even after augmenting with a full set of *T-1* dummies, imposes common evolution of unobservables in all countries. If the DGP of unobservables is indeed heterogeneous and nonstationary, the errors again become nonstationary as well, resulting in the breakdown of the cointegrating relationship (Eberhardt and Teal 2011). If the true DGP process of unobservables is nonstationary, country regressions with linear trends or time dummies capture heterogeneity but lead to misspecification of the true evolution of unobservables, resulting in non-stationarity in errors (Bai et al. 2009). We will formally introduce the concepts of stationarity, parameter heterogeneity and cross-section dependence into our econometric model in Section 5.

3 Data

Annual data on 69 developing countries covering the period 1980 to 2013 are used in this analysis. All variables are in logarithms. Data on gross official development assistance (ODA hereafter) disbursements, net ODA disbursements, gross ODA loans, gross ODA grants and technical cooperation (off-budget aid) are obtained from the OECD's Development Assistance Committee (DAC) database. The main independent variable of interest is net ODA disbursements, which is the sum of ODA grants, capital subscriptions and net loans. As the OECD DAC database comprises aid data provided by donors, it tends to overstate the amount of aid that goes through recipients' budgets. Hence to get a measure of net ODA 'close' to that which goes through recipients' budgets, we deduct technical assistance from grants. This gives us a new measure of grants which we then add to net loans to get net ODA figures for the econometric analysis. We also estimate variants of the main model with grants, loans and technical assistance (the proxy for off-budget aid) as regressors of primary interest.

Total government spending data are obtained from the World Development Indicators database. As total government spending is the sum of government consumption (recurrent) expenditure and domestic public investment (capital expenditure), we also collect data for these two expenditure headings (both from the World Development Indicators database as well).

Data on total revenue, total tax and non-resource tax revenues, all excluding grants and social contributions, are obtained from the Global Revenue Dataset (GRD) developed by the International Centre for Tax and Development (ICTD) and currently hosted by UNU-WIDER.[7] Non-resource tax revenue excludes royalties and natural resource taxes (Prichard et al. 2014). Total tax revenue comprises all direct and indirect tax revenues while total government revenue is a composite of all tax and non-tax revenues (for instance, central bank receipts, asset sales and retrievals, and sales of telecommunications licenses). The GRD data are compiled by combining data from the major international databases, individual country IMF Article IV reports and national budgets, as well as private tax datasets (for example, Keen and Mansour 2009). The data distinguish between resource and non-resource components of each tax type, with the non-resource components of particular interest in econometric research. The data are scaled by recipients' GDP series (sourced from World Economic Outlook) in local currency units, ensuring

[7] The 2015 version was used for this paper.

consistency across countries and data sources. By combining datasets and still ensuring consistency across sources, the GRD dataset achieves improved data coverage (Prichard et al. 2014).

Figure 2 shows the distribution and evolution of net aid/GDP, spending/GDP and tax/GDP rates for all 69 countries in the sample covering the period 1980 to 2013. The lower and upper ends of each box plot show the 25th and 75th percentiles, respectively, with the horizontal line in the middle of the box indicating the median. The interquartile range and median of net aid/GDP rates increase consistently until 1988, then drop slightly in 1992 and continue falling until 2000. They pick up after 2000 until 2008, then drop again in 2013. Despite these fluctuations, aid still represents a large part of recipients' GDP over time on average (28% in Guinea-Bissau, 20% in Mozambique, 17% in Burundi and Malawi, and 14% in Rwanda). Spending/GDP ratios increase in 1984 but reduce gradually thereafter, with a noticeable reduction 1992 and 2000. The rates pick up gradually after 2000. Tax/GDP ratios show a consistent pattern of evolution over time, with the interquartile range as well as the median increasing steadily over time (tax/GDP ratios are consistently between 9% and 20%) albeit with some minor fluctuations.

Figure 3 shows the composition of aid (grants vs. loans) to countries at different stages of their economic development (based on the World Bank's income classification). Least developed countries (LDCs) received more aid in grants, rather than in loans, over time. This is intuitive as grants are not expected to be repaid. It is possible that these countries receive more grants simply because their tax/GDP ratios are low. Lloyd et al. (2009) state that the fiscal impact of aid may be conditional on the level of development (income) of the country. A growing economy produces a larger tax base and more efficient tax collection mechanisms so that the tax/GDP ratio increases while the aid/GDP ratio declines (with aid representing a lower share of government spending). Least developed countries are those with slow (sometimes stagnant) growth, low tax/GDP ratios and for whom aid still represents a huge share of government spending. It follows then, that they would receive more aid in the form of grants, as their tax collection capacities are constrained, and they are not able to tax as much as may be necessary for development.

As countries become more developed, they tend to receive less aid (proportionately). Aid to upper middle-income countries (UIMCs), regardless of its composition, has reduced over time, with grants and loans representing less than 3% of recipients' GDP in 2013 (for the UMICs in the sample). Some countries currently in the 'upper middle income' bracket were probably 'less developed' in the late 1970s and early 1980s, which is why the amount and composition of their aid is higher in the 1980s than it is now.

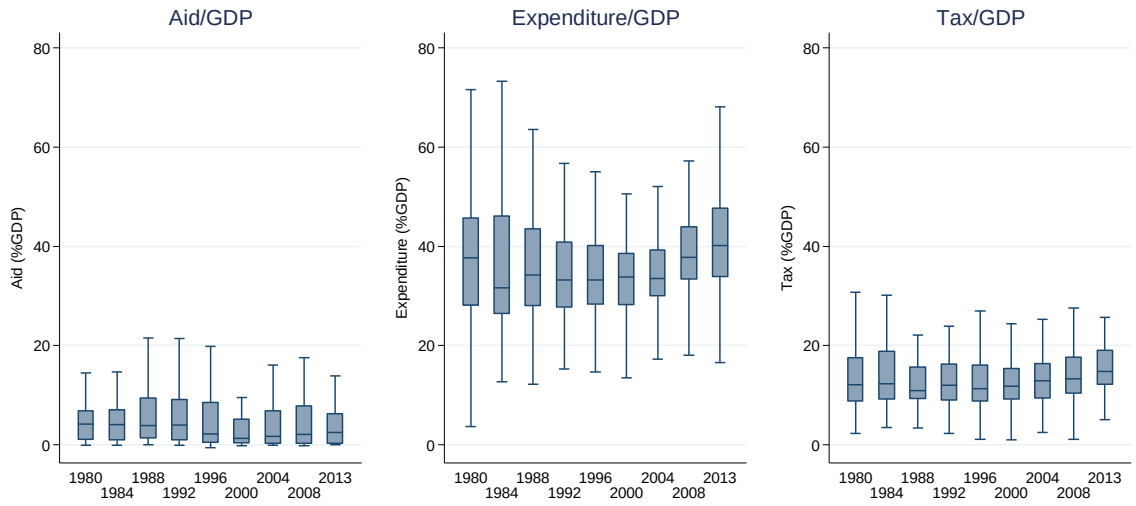
Figure 4a shows that least developed countries have the lowest tax/GDP ratios in the sample. While this may indicate fiscal nonchalance on the part of governments in LDCs, it may also reflect the fact that least developed countries are constrained in their ability to raise taxes. Indeed, Keen and Simone (2004) find that the poorest countries face greatest difficulty in raising tax revenue. Given the tax base, they may be taxing as much as they can, but it still is insufficient to generate economic and political gains (Osei et al. 2005; Morrissey 2015b). Even in LDCs with more efficient tax collection mechanisms than their peers at the same stage of economic development, the tax base is still small as the informal sector is very large in such countries (Morrissey 2012). All these effects translate into lower tax/GDP ratios, slow (or even non-existent) growth and high dependence on foreign aid flows. As alluded to earlier, such countries are the ones that receive more aid in the form of grants.

As mentioned in Section 2.1, Van de Sijpe (2013) finds that technical assistance takes up a big share of education and health aid to developing countries (and is also a huge share of total aid). Such aid does not go through recipients' budgets but can still generate fiscal responses from recipients. Figure 4b shows that UMICs receive more in technical assistance than low- and middle-

income countries (LMICs) and LDCs. As UMICs are in more advanced stages of development, they have the basic (necessary) stock of human and physical capital, as well as the required institutional capacity to benefit fully from the transfer of knowledge from donors through technical assistance. Fiscal institutions are weaker in LDCs, and policies designed to improve capacity building are at their nascent stage of implementation. As such, LDCs cannot fully reap the benefits of donors' technical assistance, reason why they receive relatively less in technical cooperation. In absolute terms the amount of technical cooperation LDCs receive is large but relative to middle-income countries, the amount is considerably less.

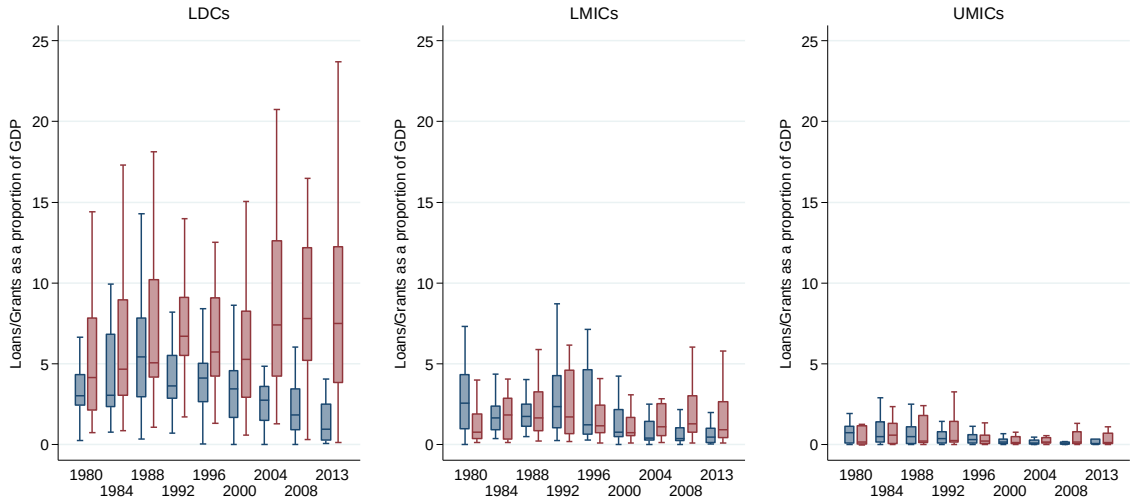
Figure 5 shows the aid/spending ratio for countries in the three income groups. On average the aid/spending ratio is substantially larger for least developed countries, with aid accounting for 80% of spending in Guinea-Bissau, over 50% of spending in Burundi, Mozambique, and Rwanda and over 40% in Malawi, Chad and Uganda. These figures indicate the importance of aid to the poorest countries, and aid is expected to have a financing impact on spending in those countries. The average aid/spending ratio is much smaller for middle-income countries (0.09% in Venezuela, 0.14% in Argentina, 0.18% in Chile). While these descriptive statistics are informative on the distribution of fiscal variables across different groups, they are also highly stylized. Much cannot be inferred from the diagrammes, except that countries receive significant amounts of aid (disbursed in varying ways), with aid/GDP ratios much higher for low-income countries. In addition, expenditures are higher than aid and taxes, indicating that expenditure patterns cannot be sustained solely by either tax revenue or aid. Hence aid and taxes must be complementary, and the observable impact (and direction of effects) of aid and taxes on spending can only be determined empirically.

Figure 2: Distribution and evolution of fiscal variables over time (1980–2013)



Source: author's calculations based on data from OECD DAC, World Development Indicators, and GRD.

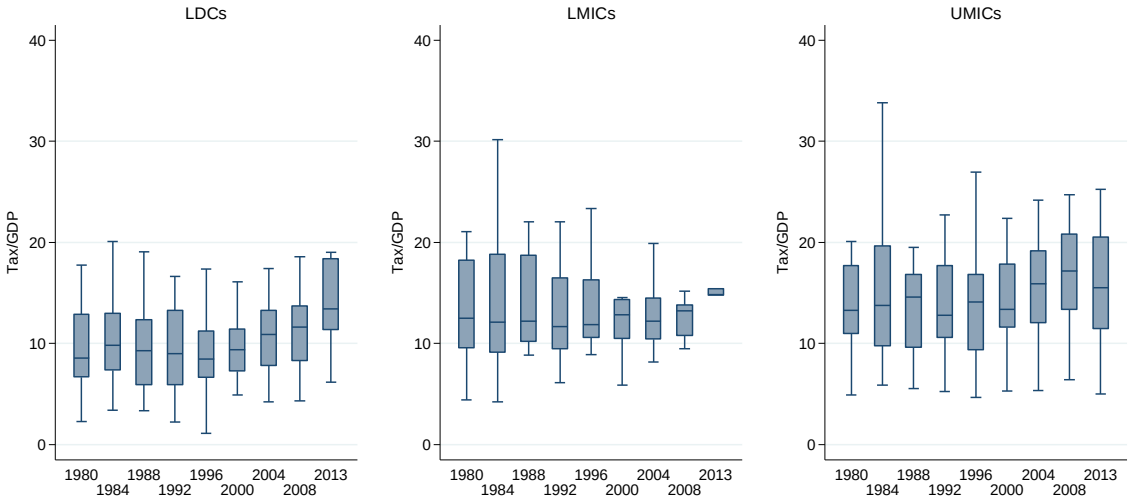
Figure 3: ODA loans and grants by income group (1980–2013)



Note: the blue boxes represent loans while the red ones represent grants. Countries are classified according to their level of economic development; based on the World Bank's income classification. LDCs are low-income countries; LMICs are lower middle-income countries and UMICs are upper middle-income countries. See Table A2 for details of the sample for each income group.

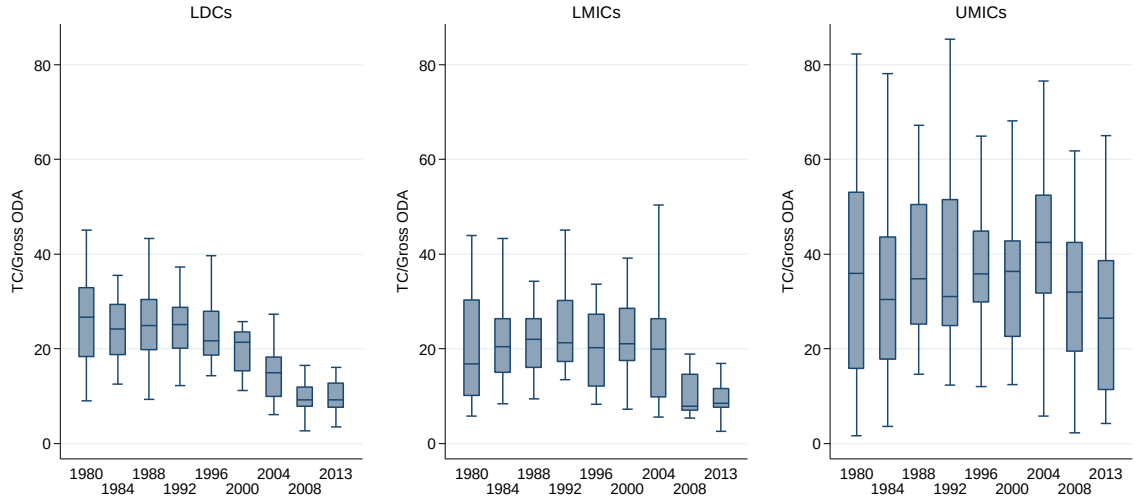
Source: author's calculations based on data from OECD DAC.

Figure 4a: tax revenue by income group (1980–2013)



Source: author's calculations based on data from GRD.

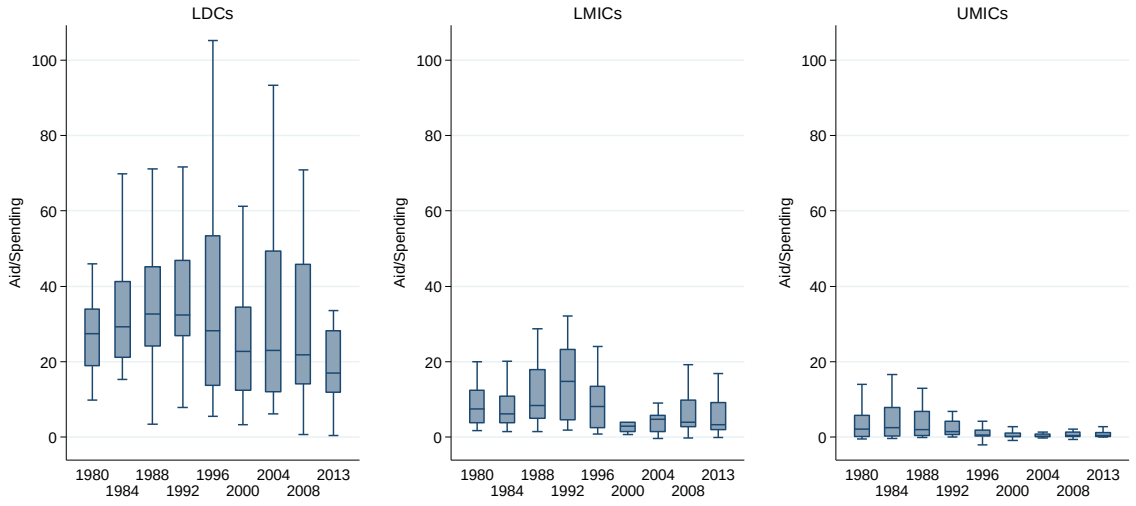
Figure 4b: Technical cooperation by income group (1980–2013)



Note: TC/Gross ODA refers to technical cooperation as a share of gross aid disbursements.

Source: author's calculations based on OECD DAC.

Figure 5: Aid to spending ratio by income group (1980–2013)



Source: author's calculations based on data from OECD DAC and World Development Indicators.

4 Empirical model specification

4.1 Cross-section dependence and unit roots

As mentioned in Section 2.3, it is important to investigate for cross-section dependence in the data. This can be done by taking a variable series for country *i* (or residuals from an estimating equation for country *i*) and correlating it with the variable series (or residual) for the other *N-1* countries. Doing that for all countries in the sample, we end up with *N(N-1)* correlation coefficients from which we can obtain the average correlation and the average absolute correlation coefficients. Alternatively, these *N(N-1)* correlation coefficients can be used to obtain a more formal test statistic (for example, the Pesaran *CD* statistic). The Pesaran (2015) cross-section dependence (*CD*) test is based on the pairwise correlations of variable series or residuals, and the

statistic is approximately normally distributed as *N* and *T* get sufficiently large (De Hoyos and Sarafidis 2006). For a balanced panel, the statistic can be defined as:

$$CD = \sqrt{\frac{2T}{N(N-1)}}\left(\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\rho}_{ij}\right) \quad (4)$$

where $\hat{\rho}_{ij}$ is the average pairwise correlation of the variable series (or residuals) and under of null of cross-section independence, *CD* is distributed *N(0,1)* for sufficiently large *T* and $N \to \infty$. De Hoyos and Sarafidis (2006) explain, theoretically and empirically, the most used tests of cross-sectional dependence. Hence, we will provide the *CD* statistic of the variable series, the average of the *N(N-1)* correlation coefficients, as well as the average of their absolute values.

Guided by insights from the conceptual framework we aim to investigate the time-series properties of the data (i.e., testing for the presence of unit roots in the variable series). Consider a p^{th} order augmented Dickey-Fuller (ADF) regression of the form:

$$\Delta y_{it} = \mu_i + \beta_i t + (\rho_i - 1)y_{it-1} + \sum_{j=1}^{p}\delta_{ij}\Delta y_{it-j} + u_{it} \quad (5)$$

where y_{it} can be the logarithm of total government spending, net aid, and tax revenue, or regression residuals. *t* is the country-specific linear time trend. We can use both 'first generation' and 'second generation' panel unit root tests to test for variable and residual (non)stationarity. 'First generation' tests assume cross-section independence, and the tests were developed to increase power from pooling low-powered[8] country-specific unit root tests (Eberhardt and Teal 2011). 'Second generation' tests, on the other hand, assume that the variable series or residual has a single factor creating the cross-section dependence. The null hypothesis for the unit root tests is:

$$H_0: \rho_i - 1 = 0, i = 1, \ldots, N \quad (6)$$

against the alternative that:

$$H_1: \rho_i - 1 < 0, i = 1, \ldots, N_1;\ \rho_i - 1 = 0, i = N_1 + 1, \ldots, N \quad (7)$$

where N is the number of countries in the sample, $N_1 < N$ and N_1/N is nonzero and fixed as $N \to \infty$ (Baltagi and Moscone 2010).

The null hypothesis is that the variable series (or the regression residual) is nonstationary for all countries in the sample, with the alternative of stationarity in at least some countries. Rejection of the panel unit root null hypothesis does not imply panel stationarity, but rather rejection of non-stationarity in all countries (Eberhardt and Teal 2011). Smith and Fuertes (2007: 39) discuss the empirical challenges in testing for unit roots; single time-series unit root tests suffering from low power while panel unit root tests are difficult to interpret.

In practice, there is a mixture of countries in terms of the time-series properties of their variable series; hence, estimation methods that accommodate this mixture should be used. The first-generation test often considered is the Im, Pesaran and Shin (IPS) test. Basically, ADF tests akin to Equation (5) are run separately for each country in the sample, and the panel ADF test statistic

[8] Time-series unit root tests have low power against I(0) alternatives that are close to I(1). That is the unit root tests cannot easily distinguish highly persistent stationary series from nonstationary processes (Smith and Fuertes 2007: 39).

obtained is an average of the *t*-statistics from the *N* country ADF regressions. As the distribution of the average *t*-statistic is non-standard, the critical values must be simulated. The main drawback of the IPS test is it does not allow for the impact of unobserved common factors. Pesaran (2007) proposes a test for Equation (5), referred to as the 'CIPS' test, which includes cross-section averages of the dependent and independent variables to account for cross-section dependence. Hence Equation (5), augmented with cross-section averages of the observables will be:

$$\Delta y_{it} = \mu_i + \beta_i t + (\rho_i - 1) y_{it-1} + \sum_{j=1}^{p} \delta_{ij} \Delta y_{it-j} + \theta_{1i}^{CA} \overline{\Delta y}_{it} + \theta_{2i}^{CA} \bar{y}_{it-1} + \theta_{3i}^{CA} \overline{\Delta y}_{it-p} + u_{it} \quad (8)$$

where in addition to the terms in Equation (5), cross-section averages of the dependent and independent variables are included. The 'CIPS' test can deal with a single common factor which affects countries to varying degrees.

4.2 The econometric model: common factor representation

The equation of primary interest in this analysis is a static structural equation akin to Equation (2), modelling the effects of aid and taxes on spending. Building on a common factor approach, the equation will be of the form:

$$exp_{it} = \alpha_i + \beta_{i1}(aid)_{it} + \beta_{i2}(tax)_{it} + \varepsilon_{it} \qquad \varepsilon_{it} = \gamma_i' \Gamma_t + \mu_{it} \quad (9)$$

where exp_{it} is total government spending, $aid_{it}$is net ODA and $tax_{it}$is tax revenue—all in logarithms. The vector of slope coefficients β_i differs across countries but is constant over time. Equation (9) also includes country-specific intercepts (α_i) and a vector of unobserved common factors Γ_t with country-specific factor loadings (γ_i) to account for the levels and evolution of unobserved total factor productivity (TFP), respectively.[9] These common factors can be a combination of *weak* and *strong* factors (see Section 2.3). Furthermore, these common factors not only drive government expenditures, but also aid and taxes, leading to endogeneity. To elucidate this point, assume any independent variable from Equation (9), x, evolves in the form:

$$X_{it} = \pi_i + \tau_i X_{it-1} + \Phi_{1i} \Gamma_{1t} + \cdots + \Phi_{ni} \Gamma_{nt} + \delta_i \Psi_t + \upsilon_{it} \quad (10)$$

$$\Gamma_t = \eta \Gamma_{t-1} + \epsilon_{\Gamma t} \qquad \Phi_t = \theta \Phi_{t-1} + \epsilon_{\Phi t} \quad (11)$$

for $i = 1, \ldots, N$, and $t = 1, \ldots, T$. As seen in Equation (10) the independent variables (aid and tax revenue) are driven by a set of common factors (Γ_{nt} and Ψ_t),[10] some of which influence expenditures in Equation (9). Hence, if $\gamma_i \neq 0$ and $\Phi_i \neq 0$, the error term and the regressors from Equation (9) are correlated, creating serious problems for estimation (Kapetanios et al. 2011).[11] In addition, Equation (11) indicates that the factors can be nonstationary ($\eta = 1,\ \theta = 1$), again with important implications for estimation and inference. Based on Equations (9), (10) and (11), and guided by the conceptual discussions in Section 2, we are interested in using a model that allows government expenditures to fluctuate to short-run changes in aid and taxes, and simultaneously

[9] 'TFP' here does not necessarily refer to total factor productivity. It represents any variable that is not easily captured (unobservable variables) or variables with little or no time variation.

[10] The n in Γ_{nt} refers to the number of common factors included in the model.

[11] The presence of unobserved factors, as well as the underlying equilibrium relationship differing across countries, makes the standard instrumental variables approach invalid in this setup (Eberhardt and Presbitero 2015).

correspond to the long-run levels of aid and taxes. We thus employ an unconditional error correction model (ECM) of the form:

$$\Delta exp_{it} = \alpha_i + \rho_i[y_{it-1} - \beta_{i1}(aid)_{it-1} - \beta_{i2}(tax)_{it-1} - \gamma_i'\Gamma_{t-1}] + \psi_{i1}\Delta(aid)_{it} + \psi_{i2}\Delta(tax)_{it} + \psi_{i3}\Delta\Gamma_t + \varepsilon_{it} \quad (12)$$

where the β_is represent the long-run relationship between spending, aid and taxes in the model (a potential cointegrating relationship, from which we are interested in getting the average value of β_i), and ψ_is represent the short-run relations. We include the unobservable common factors Γ in the long-run equation as we posit that they are nonstationary, hence forming part of the cointegrating relationship. ρ_i represents the speed with which the economy is returned to its long-run equilibrium, following a deviation from its long-run path (Hendry 1995). The ECM specification above has at least three advantages over static or more restrictive dynamic models that assume parameter homogeneity (for example, Remmer 2004). First, we can distinguish long-run from short-run dynamics, the theoretical premise from which ECMs are derived. Second, the error correction term ρ_i and the long-run coefficients ($\beta_{i,}$s) are useful in determining the exogeneity status of the variables (Lloyd et al. 2009). Third, we can investigate and test for cointegration in the ECM based on the statistical significance of the error correction term in the ECM (Eberhardt and Presbitero 2015). A negative and significant error correction coefficient represents cointegration, indicating that the economy returns to its long-run equilibrium following a deviation from equilibrium.

In the context of aid and spending, Remmer (2004) is closest to our approach. Thus, a few comments on our specification and how it relates to Remmer (2004) are noteworthy. Although both studies draw on insights from political economy and public finance theory, they are different in two aspects. First, the main hypothesis of interest in this study differs from that of Remmer (2004). Guided by the literature on growth in government size, she models the determinants of government size (measured by changes in the expenditure/GDP ratio), including aid as the variable of principal interest. She also includes other economic, institutional and demographic variables (trade openness, central government total revenue, government debts, *per capita* GDP, population and the dependency ratio) that may affect government size.

In this study we investigate if aid forms part of the long-run budgetary equilibrium (i.e., the presence of a cointegrating relationship between aid, taxes and spending) and test which of the sources of finance, aid or taxes, has a stronger association with expenditures. We rely on a political calculus between aid and tax, based on the relative importance of aid and taxes (which can be attributed to how recipients perceive their respective political costs), and how these political costs offset each other. These costs are evaluated according to autonomy, accountability and bureaucratic costs (Morrissey 2015b). Thus, we employ a parsimonious specification by including only aid and tax revenue, arguing that the control variables included in the Remmer (2004) analysis are themselves determinants of tax effort. Hence, in our analysis tax revenue will act as a proxy for all the control variables included in the Remmer (2004) analysis, as well as other unobservable time-varying factors that affect both expenditure and aid allocation decisions. This makes finding cointegration in our analysis very important, as evidence of cointegration between aid, spending and taxes would imply that no nonstationary variables have been omitted from estimation (Herzer and Morrissey 2013; Herzer 2017). Lütkepohl (2007) also states that the cointegration property is invariant to model extensions. This means, if other sources of finance (domestic/foreign borrowing) are included in the model, it may result in further cointegration relations but the initial long-run relationship between aid, taxes and spending will still hold.

This now raises the possibility that there may be more than one cointegrating relationship, as there are other sources of finance that influence recipients' spending. Nevertheless, Lloyd et al. (2009) state that there is no economic justification for more than one long-run equilibrium relationship between aid and domestic fiscal variables. Hence, in this study we treat the sole long-run relationship as a statistical analogue to the domestic budgetary equilibrium. Moreover, including other nonstationary variables into the system may result in spurious relations (Herzer 2017). Specifically, including a nonstationary variable that is not cointegrated with the other variables will lead to non-stationarity in the error term, hence a breakdown of the cointegrating relationship because the coefficient on the recently included variable will not converge to zero as expected of irrelevant variables in a standard regression (Herzer 2017). These considerations, then, justify a parsimonious model such as Equation (9).

Second, the dynamic ECM Remmer (2004) estimates is more restrictive, as the model imposes parameter homogeneity without explicit testing for such homogeneity. Tests for unit roots are conducted, and there is evidence of homogenous cointegration (judged by a negative and significant lagged dependent variable). However, potential levels relationships between the nonstationary variables are ignored. Moreover, no attempts are made to accommodate cross-section dependence into the analysis. Cross-section dependence is a recurrent feature in macro data and should always be accounted for in analysis. Thus, in this paper, we test for the presence of a long-run (equilibrium) relationship between aid, taxes and spending (implying no potentially important nonstationary variables would have been omitted), allowing for this equilibrium to differ across countries (heterogeneous cointegration) and also incorporating cross-section dependence.

The common correlated effects mean group (CCEMG) estimator of Pesaran (2006) will be used to estimate the relationship between aid, taxes and spending. The CCEMG estimator augments each country's OLS regression with cross-sectional averages of the dependent variable and the independent variables, $\bar{y}_t$ and $\bar{x}_t$. These newly included regressors account for the prevalence of unobserved common factors (Pesaran 2006; Coakley et al. 2006; Eberhardt 2012). Basically, cross-section averages for all observable variables in the model are computed and then added as explanatory variables of the *N* regression equations. These estimated coefficients are then averaged across panel members as follows:

$$\hat{\beta}^{CCEMG} = \frac{\sum_{i=1}^{N} \hat{\beta}_i}{N} \tag{13}$$

where $\hat{\beta}^{CCEMG}$ is the country-specific estimate from the CCEMG estimator. Coakley et al. (2006) and Kapetanios et al. (2011) show that the estimator is consistent in the presence of multiple common factors (Γ_t and Ψ_t in Equations (9) and (10)), factor loading dependence ($\gamma_i = \Phi_i$), cointegration or non-cointegration of the model variables, and also in the presence of regressor-specific common factors (for instance Ψ_t in Equation (10)). Additionally, the estimator is robust to nonstationary common factors ($\eta = 1,\ \theta = 1$), structural breaks and outliers. However, Chudik and Pesaran (2015) find that the CCEMG estimator is subject to small sample bias, especially in samples with moderate time series dimensions. Furthermore, they relax the strict exogeneity assumption, allowing for feedback between variables in the system, which poses a challenge for the original Pesaran (2006) estimator. To solve these two problems, Chudik and Pesaran (2015) suggest including lags of cross-section averages, in addition to the cross-section averages of all variables already included in the standard CCEMG estimation equation. Hence, augmenting Equation (12) with cross-section averages (and lags of the cross-section averages) of the dependent and independent variables, we get:

$$\Delta exp_{it} = \pi_{0i} + \pi_i^{EC} y_{it-1} + \pi_i^{aid} aid_{it-1} + \pi_i^{tax} tax_{it-1} + \Psi_i^{aid} \Delta(aid)_{it} + \Psi_i^{tax} \Delta(tax)_{it} + \pi_{1i}^{CA} exp + \pi_{2i}^{CA} \overline{exp}_{t-1} + \pi_{3i}^{CA} \overline{\text{aid}}_{t-1} + \pi_{4i}^{CA} \overline{\text{tax}}_{t-1} + \pi_{5i}^{CA} \overline{\Delta aid}_t + \pi_{6i}^{CA} \overline{\Delta tax}_t + \sum_{l=1}^{p} \pi_{7i}^{CA} \overline{\Delta exp}_{t-p} + \sum_{l=1}^{p} \pi_{8i}^{CA} \overline{\Delta aid}_{t-p} + \sum_{l=1}^{p} \pi_{9i}^{CA} \overline{\Delta tax}_{t-p} + \varepsilon_{it} \quad (14)$$

where the coefficients πs and Ψs represent the long-run and short-run coefficients, respectively, and the coefficients π^{CA}s represent the coefficients on the cross-section averages of the dependent and independent variables (all coefficients yielding the standard CCEMG estimator). The coefficients $\sum_{l=1}^{p} \pi^{CA}$s represent the coefficients on the additional lags of cross-section averages, which Chudik and Pesaran (2015) suggest be added to the standard CCEMG estimator (yielding the dynamic CCEMG estimator). As a rule of thumb, the lags of the cross-section averages to be added to the standard model are chosen by $p = T^{\frac{1}{3}}$ (Chudik and Pesaran 2015). Chudik and Pesaran (2015) show that once the CCEMG estimator has been augmented with the sufficient number of lags, the estimator is unbiased in the presence of dynamics (the lagged dependent variable), and also in the presence of weakly exogenous regressors. We then estimate Equation (14) and obtain the long-run coefficients on aid and taxes in the form:

$$\beta_i^{aid} = -\frac{\pi_i^{aid}}{\pi_i^{EC}} \qquad \beta_i^{tax} = -\frac{\pi_i^{tax}}{\pi_i^{EC}} \quad (14^*)$$

The regression coefficients on the terms in first differences capture the short-run (transitory) effects and can be read off directly from estimation. Inference on π_i^{EC}, the speed of convergence to equilibrium, provides insights into the presence of a long-run (cointegrating) relationship between aid, taxes and spending. If $\pi_i^{EC} = 0$, then there is no cointegration, and the model reduces to one with variables in first differences. If $\pi_i^{EC} \neq 0$, then there is 'error correction' in the model. That is, following a shock the economy returns to its long-run equilibrium path, and therefore there exists a cointegrating relationship between aid, taxes and spending.

5 Empirical results

5.1 Cross-section dependence and unit roots

Results for cross-section dependence (Table 1) provide evidence of the pervasiveness of cross-section correlation in the sample, based on the cross-country correlation coefficients and the Pesaran (2015) *CD* test. These results hold for the individual variables, both in levels and first differences. The correlation coefficients and *CD* statistic are considerably lower for variables in first differences, but cross-section dependence is still pervasive.

Table 1: Cross-section dependence

Panel A	**Variables in levels**			
	Exp_{it}	Aid_{it}	TC_{it}	Tax_{it}
avg $\hat{\rho}_{ij}$	0.507	0.132	0.314	0.572
avg $\lvert\hat{\rho}_{ij}\rvert$	0.566	0.288	0.438	0.628
CD	154.80	41.61	97.95	174.72
p-value	0.00	0.00	0.00	0.00
Panel B	**Variables in first differences**			
	ΔExp_{it}	ΔAid_{it}	ΔTC_{it}	ΔTax_{it}
avg $\hat{\rho}_{ij}$	0.111	0.039	0.222	0.107
avg $\lvert\hat{\rho}_{ij}\rvert$	0.203	0.184	0.278	0.205
CD	33.31	12.09	67.10	32.22
p-value	0.00	0.00	0.00	0.00

Note: we use the stata routine 'xtcd' developed by Markus Eberhardt. We report the average correlation (avg $\hat{\rho}_{ij}$) and average absolute correlation (avg $|\hat{\rho}_{ij}|$) coefficients of the N(N-1) sets of correlations. *CD* is the Pesaran (2015) test for cross-section dependence distributed N(0,1) under the null of cross-section independence. Panels A and B test for cross-section dependence in the variable series for levels and first differences, respectively. Net ODA (Aid), technical cooperation (TC), tax revenue (Tax) and government expenditure (Exp) all in logs.

Source: author's calculations.

Cross-section dependence (CSD) leads to standard panel unit roots tests suffering from significant size distortions, resulting in them over-rejecting the null of non-stationarity (Pesaran 2007; Eberhardt and Presbitero 2015). Thus, panel unit root tests that accommodate such dependence are more appropriate. To that end, we use the 'CIPS' test by Pesaran (2007). Table 2 provides the results from conducting the 'CIPS' test. We report the *Z*t-bar statistic (and its corresponding *p*-value) for H_0 = non-stationarity in all countries' variable series versus H_1 = stationarity in some countries' variable series. The panel statistic is obtained by normalizing the individual country *t*-statistics using simulated values of the mean and variance (Söderbom et al. 2014). Normalization makes the *Z*t-bar statistic distributed *N(0,1)*.

For all variables in levels, non-stationarity cannot be rejected once the ADF equation is augmented with a sufficient number of lags and/or a trend. Non-stationarity is rejected for all variables in first differences. Given all the caveats and problems of individual country and panel unit root tests, we suggest most conservatively that non-stationarity cannot be ruled out in this dataset. Non-stationarity of these fiscal variables implies they potentially form a cointegrating relationship. We now formally test for the presence of this budgetary equilibrium (cointegrating relationship).[12]

[12] Panel unit root tests results as well as tests for cross-section dependence for variables used in exploratory analysis can be found in Tables A4 and A5.

Table 2: Pesaran (2007) unit root test

Levels: CIPS with intercept only								
Variable	Aid		TC		Tax		Exp	
Lags	Ztbar	*p*	Ztbar	*p*	Ztbar	*p*	Ztbar	*p*
0	-8.72	0.00	-4.47	0.00	-2.80	0.00	-2.33	0.01
1	-3.84	0.00	-1.01	0.16	-2.37	0.01	-3.72	0.00
2	-1.90	0.03	0.62	0.73	-0.58	0.28	-1.05	0.15
3	-1.30	0.10	-0.18	0.43	-1.12	0.12	-1.11	0.13
4	2.58	1.00	0.12	0.55	-0.55	0.29	0.37	0.65
Levels: CIPS with intercept & trend								
Variable	Aid		TC		Tax		Exp	
Lags	Ztbar	*p*	Ztbar	*p*	Ztbar	*p*	Ztbar	*p*
0	-8.01	0.00	-3.32	0.00	-0.97	0.17	-2.36	0.01
1	-2.86	0.00	0.60	0.73	-0.81	0.21	-4.44	0.00
2	0.11	0.54	3.79	1.00	2.43	0.99	-1.36	0.09
3	-0.03	0.49	3.00	0.99	1.32	0.91	0.06	0.52
4	6.22	1.00	3.57	1.00	2.01	0.98	1.24	0.89
Differences: CIPS test with drift								
Variable	Aid		TC		Tax		Exp	
Lags	Ztbar	*p*	Ztbar	*p*	Ztbar	*p*	Ztbar	*p*
0	-33.26	0.00	-35.35	0.00	-29.28	0.00	-30.63	0.00
1	-21.70	0.00	-22.13	0.00	-19.55	0.00	-20.72	0.00
2	-12.56	0.00	-11.35	0.00	-8.40	0.00	-12.31	0.00
3	-8.80	0.00	-9.57	0.00	-5.64	0.00	-8.04	0.00
4	-0.27	0.39	-3.90	0.00	-4.45	0.00	-5.34	0.00

Note: net ODA (Aid), technical cooperation (TC), tax revenue (Tax) and government expenditure (Exp) all in logs. 'Lags' denote the number of lags of the differenced dependent variable included to wipe out serial correlation.

Source: author's calculations.

5.2 Heterogeneous baseline estimates

Having confirmed the prevalence of cross-section dependence and established that all the variables are nonstationary in levels, we proceed to estimate the heterogeneous dynamic ECM using a dynamic CCEMG estimator, results of which are reported in Table 3. We report results for the standard CCEMG (Pesaran 2006) as well as variants augmented with one and two lags of cross-section averages, respectively (Chudik and Pesaran 2015). Long-run averages and short-run coefficients of the variables are reported. The coefficient on the lagged dependent variable is reported as well.

The long-run average coefficient is obtained by averaging ECM coefficients, then computing the long-run coefficient with standard errors computed through the Delta method. We employ the robust regression (see Hamilton 1992)—which weighs down outliers in computing the averages—in all estimations. The first column of each model reports the results with net ODA (excluding technical cooperation) as the primary regressor of interest (*Aid1*), while the second column of each model reports the results with net ODA including technical cooperation (*Aid2*). Relevant diagnostics (RMSE, CD test statistic) are reported at the bottom of the table. Non-stationarity of residuals is tested (not reported) in all models and the null of a unit root in residuals is rejected in all cases (i.e., at varying levels of significance and the inclusion of further lags of the dependent variable and/or a trend). Given the small sample bias the standard CCEMG faces, in addition to the *favourable* results and diagnostics from the variant with two additional lags of the cross-section

averages, we only discuss results based on the CCEMG augmented with two lags of cross-section averages.

Table 3: CCEMG estimates

	CCEMG		One-lag CCEMG		Two-lag CCEMG	
Long-run						
Tax	0.760***	0.719***	0.726***	0.735***	0.825***	0.734***
	[0.079]	[0.076]	[0.085]	[0.085]	[0.092]	[0.089]
Aid1	0.041**		0.025*		0.032**	
	[0.016]		[0.013]		[0.015]	
Aid2		0.055**		0.092***		0.066**
		[0.027]		[0.025]		[0.027]
Short-run						
Tax	0.513***	0.573***	0.582***	0.582***	0.586***	0.591***
	[0.044]	[0.045]	[0.048]	[0.050]	[0.055]	[0.056]
Aid1	0.023**		0.017**		0.017**	
	[0.006]		[0.007]		[0.008]	
Aid2		0.014		0.007		0.029**
		[0.011]		[0.007]		[0.012]
EC coefficient						
y_{it-1}	-0.674***	-0.728***	-0.744***	-0.777***	-0.788***	-0.822***
	[0.035]	[0.033]	[0.038]	[0.039]	[0.040]	[0.046]
t-statistic	-19.01	-22.02	-19.73	-19.76	-19.45	-17.78
Diagnostics						
RMSE	0.106	0.104	0.094	0.092	0.079	0.078
CD test	-0.46	-0.57	-0.65	-0.60	-0.07	-0.25
(*p*-value)	0.649	0.565	0.515	0.546	0.943	0.804
Observations	2086	2033	2038	2007	1989	1921

Note: these results are based on an error correction model for all 69 countries in the sample with the first difference of log spending as dependent variable. *Aid1* refers to net ODA excluding technical cooperation, while *Aid2* refers to net ODA including technical cooperation. 'CCEMG' represents the Pesaran (2006) common correlated effects mean group (CCEMG) estimator while 'one-lag' and 'two-lag CCEMG' represent the standard CCEMG augmented with one and two lags of the cross-section averages, respectively. The *long-run* averages (the βs from Equation 14*) are computed from the robust mean estimates of the CCEMG models with standard errors (reported below the averages) computed through the Delta method. The *short-run* coefficients (the Ψs from Equation 14) are read off directly from estimation. All models are augmented with country-specific linear trend terms. The *t*-statistic on the lagged dependent variable is a non-parametric statistic derived from the country-specific coefficients following Pesaran and Smith (1995). RMSE is the root mean square error, while *CD* test is the Pesaran (2015) test distributed *N(0,1)* under the null of cross-section independence (*p*-value in parentheses below). *, ** and *** indicate rejection of H_0 = coefficient equals zero at 10%, 5% and 1%, respectively.

Source: author's calculations.

As the variables in the analysis are all I(1), we can test for cointegration by investigating the statistical significance of the lagged dependent variable as shown in Table 3. The coefficient on the lagged dependent variable is negative, statistically significant and different from zero, indicating that the system reverts to its equilibrium path following a shock. Additionally, the t-statistic[13] is *sufficiently* greater than 10 in the model, providing 'solid' evidence of a long-run equilibrium (cointegrating) relationship between aid, taxes and spending. The results indicate that on average there is a long-run budgetary equilibrium between fiscal variables and aid is an important

[13] This *t*-statistic on the lagged dependent variable does not follow a *t*-distribution, but a large value of, say, 10 is 'solid' evidence of cointegration.

determinant of the equilibrium. Hence, increases in government spending are sustained by movements in taxes and foreign aid.

Tests for panel cointegration have been proposed in the econometric literature, each with their own strengths and flaws. Analogous to panel unit root tests, the 'first generation' of panel cointegration tests assumes cross-section independence, while 'second-generation' tests incorporate cross-section dependence. Again, like panel unit root tests, these cointegration tests are for the null hypothesis of cointegration in all country series against the hypothesis of non-cointegration in *at least* some country series (Söderbom et al. 2014). If the null is rejected, there are complex issues on how to entertain a mixture of countries with cointegrated data and others with no cointegration. Although the CCEMG estimator is consistent in the presence of non-cointegration, we emphasize the difficulties in carrying out 'solid' inference on the existence of cointegration. Hence, our results should be merely indicative.

The long-run average coefficients on aid and taxes are positive and statistically significant, consistent with the fiscal effects literature (Bwire et al. 2017; Mascagni and Timmis 2017). A positive relationship between aid and spending is to be expected, as aid still represents an important source of revenue for most developing countries. In the long run, a one-percentage-point increase in aid is associated with a 0.032% increase in spending. This average effect is small, given the perceived importance of aid to developing countries,[14] but provides insight into the *general tendency* in the panel that on average, there is a positive association between aid and expenditures.

As regards taxes, higher tax revenue is also associated with higher spending in the long run. A one-percentage-point increase in taxes is associated with a 0.825% increase in government expenditures in the long run. The coefficient on taxes is large (and larger than that of aid), indicating that on average, long-run spending plans are driven mainly by tax revenue (or domestic revenue in general). While developing countries still maintain high shares of aid, their spending patterns are dictated mainly by their domestic sources of revenue. This is plausible as spending driven by domestic revenue reduces the risk of fiscal vulnerability because of aid unpredictability.

Turning to the short-run coefficients, again, aid and taxes are positively associated with spending. The short-run coefficient on aid is smaller than the long-run average, but still positive and significant. In the short run, a one-percentage-point increase in aid is associated with a 0.02% increase in spending. Not only is aid a component of the long-run budgetary equilibrium, on average, but it also relaxes the budget constraint in recipient countries. This short-run impact may also reflect the volatility of aid flows to some countries, with such volatility making it impossible for aid to be used for planning in those countries. Hence, it merely relaxes the budget constraint. Regarding taxes, a one-percentage-point increase in taxes is associated with a 0.59% increase in expenditures in the short run.

Overall, the results suggest that aid is important for long-run budgetary planning but is also used as a substitute for private sources of finance in the short run. Spending in recipient countries is driven mainly by tax revenue, though recipients form some expectation of the amount of aid to be received in the long run and incorporate these expectations into their budgetary planning

[14] Using sample data to illustrate, on average aid accounted for about 16% of government spending for all 69 countries. On average, aid was approximately 80 % of spending in Guinea-Bissau, over 50% of spending in Burundi, Mozambique, and Rwanda and over 40% in Malawi, Chad and Uganda. These statistics illustrate the importance of aid in developing countries.

processes. The choice between tax and aid also relates to the political costs associated with tax and aid, evaluated according to autonomy, accountability and bureaucracy.

Recipients that are overly dependent on aid will have to account to donors on how their aid is spent, and also negotiate on conditionality, reducing their policy discretion, since increased aid dependence means they will have to cede some of their policy influence to donors (Morrissey and Torrance 2015; Morrissey 2015b). In that regard, any government that can make independent policy choices is an autonomous one. The desire for greater autonomy encourages governments to expend effort on improving tax collection, reducing their dependence on aid. Morrissey and Torrance (2015) posit that increasing taxes makes the public-sector decision-makers more accountable to their constituencies, but the benefits of autonomy offset such costs of accountability to the domestic constituencies. They also state that domestic revenue mobilization with accountability to constituencies can enhance legitimacy and state building. Therefore, to the extent that recipients dislike ceding policy influence to donors, increasing taxes is more preferable than aid dependence as these recipients cede less influence to donors.

The costs of accountability refer to whom and the extent to which a government must account for its uses of revenue, and these costs are likely to be greater for aid than taxes (Morrissey 2015b). The donor agencies themselves must account to their own governments and parliaments, so they implement measures to monitor the use of aid and minimize fungibility. Donors often provide aid with policy reform conditions (usually relating to public finance management) and exert considerable effort in monitoring the use of their aid. This means recipients must expend effort in negotiating conditions with donors and trying to circumvent those conditions is usually costly. In contrast, accountability to taxpayers is much weaker in least developed countries that are major recipients of aid, which reflects the broader institutional set-up in those countries. Hence, the costs of accountability are higher for aid, especially in these countries with weak institutions and comparatively lower revenues.

In addition to costs of autonomy and accountability, there are also bureaucratic costs of tax and aid. The former relates to the costs of tax administration, while the latter refer to the costs of organizing and attending meetings with different donor agencies (Morrissey and Torrance 2015). Morrissey (2015b) states that, in the last decade, most developing countries have implemented fiscal reforms and improved tax administration, reducing the bureaucratic cost of taxation. In contrast, the bureaucratic costs of aid are more of a function of the number of donors than the amount of aid. Donor proliferation, disbursement heterogeneity (i.e., aid disbursed in different ways for different purposes) and the changing requirements on monitoring aid result in high bureaucratic costs of aid. Even if donors coordinate and form donor groups, the bureaucratic costs of aid will still be high as recipients will have to attend meetings with the multiple donor groups. This means the bureaucratic costs of taxation are declining while those of aid are still high.

In general, the political calculus between taxes and aid is heterogeneous across countries. Our results suggest costs of taxation are likely to be lower than those of aid, which is one of the main reasons why taxes are the main driver of expenditures in recipient countries. At disaggregated levels, however, the political costs of tax may be higher than the political costs of aid, especially in SSA countries.

5.3 Exploratory analysis

We undertake a series of robustness tests to explore further the relationship between aid, taxes and government expenditures.

Disaggregated expenditures

As government spending is a composite of capital (development) spending and recurrent (consumption) spending, it would be insightful to estimate the disaggregated spending impact of net aid so one can get an idea of what aid is funding. Thus, we re-estimate Equation (14) with capital expenditures and recurrent expenditures as dependent variables, respectively. Results for the disaggregated spending impact are reported in Table 4.

Table 4: Heterogeneity in expenditures

	Capital expenditure	**Recurrent expenditure**
Long-run		
Tax	0.718*** [0.100]	0.820*** [0.105]
Aid	0.040** [0.018]	0.012 [0.021]
Short-run		
Tax	0.692*** [0.069]	0.492*** [0.045]
Aid	0.015 [0.010]	0.009 [0.006]
EC coefficient		
y_{it-1}	-0.783*** [0.038]	-0.604*** [0.041]
t-statistic	-20.36	-14.88
Diagnostics		
RMSE	0.108	0.074
CD test (*p*-value)	0.54 (0.592)	-0.20 (0.839)
Observations	1954	1962

Note: 'Capital expenditure' and 'Recurrent expenditure' refer to error correction models for all 69 countries with the first difference of log capital expenditure and log recurrent expenditure as dependent variables, respectively. The CCEMG with two lags of cross-section averages is used. For all other details, see Table 3.

Source: author's calculations.

Investigation of the lagged dependent variable indicates that aid, capital expenditures and taxes form an equilibrium (cointegrating) relationship. This implies that physical capital projects in recipient countries are sustained by aid and taxes. On average, aid has a positive long-run impact on capital expenditures. A one-percentage-point increase in aid is associated with a 0.040% increase in capital expenditures. As a large share of aid is intended to boost recipients' development prospects (i.e., much aid is intended for investment), we would expect aid to have a positive long-run impact on capital spending, the magnitude of which will be influenced by the productivity of investment in the recipient country. The productivity of investment itself depends on the institutional set-up of the country.

Taxes have a larger coefficient than aid which, again, may suggest that public-sector decision-makers in recipient countries rely more on their domestic sources of revenue for financing their development projects. A one-percentage-point increase in taxes is associated with a 0.718%

increase in capital expenditures. The absence of evidence of a short-run impact of aid on capital spending is intuitive. As capital expenditures consist mainly of capital projects that mature in the medium to long term, the observable impact of aid is restricted to the long run only. Short-term capital expenditures are financed mainly via tax revenue (as shown by the magnitude of taxes) and possibly domestic borrowing.

Aid does not have any significant impact on recurrent spending, while taxes impact on recurrent spending in both the long run and short run. Martins (2010) also finds no long-run relationship between aid and recurrent spending. We do not overstate this finding as the proportion of aid intended for different spending headings is unknown.[15] Increasing social sector spending (on health, education and sanitation) is one of the intentions of donors, and maintenance costs for these social sector investments (hospitals, schools, water plants) are usually classified as recurrent expenditures. Hence, this result should not be treated as evidence of no aid fungibility, but as absence of evidence of aid fungibility. The results differ from Osei et al. (2005), who find that aid appeared to be associated with higher recurrent spending but not capital spending in Ghana. This was not due to a financing impact of aid *per se*, as policy conditions attached to aid were associated with higher tax revenue (which implied higher recurrent spending) and lower domestic borrowing (which implied lower capital spending).

Level of development

Here the sample is split into two income groups—least developed countries (LDCs) and other low-income countries (LICs)—based on the World Bank's classification. We thus re-estimate Equation (14) for each income group, results of which are presented in Table 5. After re-estimating Equation (14) for each income group, there is still considerable cross-section dependence in the residuals. It is possible that interdependencies between countries create common factors that have not been captured. Chudik and Pesaran (2015) suggest that in addition to cross-section averages and lags of cross-section averages in Equation (14), cross-section averages of one or more covariates (other than aid and taxes) be included. These cross-section averages may help identify the multiple unobserved common factors not fully captured by the original set of cross-section averages (Eberhardt and Presbitero 2015). Hence, in this disaggregated model we include:

$$\sum_{l=0}^{p} \pi_{10i}^{CA} \overline{\Delta \Upsilon}_{t-p}$$

for each covariate Υ and equally for further covariates. The lags of cross-section averages of these further covariates are also determined based on the rule of thumb $p = T^{\frac{1}{3}}$ (Chudik and Pesaran 2015). The country series for these additional covariates *do not* enter the model as regressors; just their cross-section averages and lags of cross-section averages enter the model. The objective here is to help identify the unobserved common factors Γ_t, so including variables that may be directly linked to expenditures is reasonable. Therefore, we include exports (in logs) as an additional covariate, as it potentially influences recipients' expenditures through its impact on taxes, and the largest exporters in developing countries are government owned. Furthermore, Remmer (2004) argues that international exposure generates pressures for recipient governments to increase spending, making exports a good candidate for inclusion here.

[15] It is not until the early 1990s that donors started providing data on aid for specific purposes. Historical aid data are void of that level of disaggregation.

Table 5: Heterogeneity in levels of development

	LDCs	**LICs**
Long-run		
Tax	0.675*** [0.064]	0.831*** [0.150]
Aid	0.199*** [0.145]	0.004 [0.010]
Short-run		
Tax	0.603*** [0.079]	0.584*** [0.074]
Aid	0.092** [0.035]	-0.004 [0.004]
EC coefficient		
y_{it-1}	-0.872*** [0.078]	-0.769*** [0.061]
t-statistic	-11.15	-12.52
Diagnostics		
RMSE	0.086	0.048
CD test (p-value)	0.11 (0.914)	4.04 (0.000)
Observations	764	1225
Countries	26	43

Note: LDCs refers to low-income countries and LICs to other low-income countries (lower middle-income and upper middle-income countries). Error correction models are estimated for each income group, with the first difference of log expenditure as dependent variable. For all other details, see Table 3.

Source: author's calculations.

Results indicate that aid has a positive long-run (and short-run) impact in least developed countries, with no discernible impact in other low-income countries. Of the 26 LDCs in the sample, 23 are sub-Saharan African, the highest recipients of aid. Hence finding a positive impact of aid on spending in LDCs is consistent with a positive relationship between aid and spending in sub–Saharan Africa. As mentioned earlier, these LDCs are aid dependent and as such their political costs of aid are lower than their political costs of taxation.

Furthermore, LDCs lack the necessary human capital development, financial market development and infrastructure to attract significant amounts of FDI and obtain foreign debt portfolios. The ability to attract FDI will also depend on their institutional framework (rule of law, level of corruption, government effectiveness, and risk of doing business among others), which is much weaker in LDCs than other low-income countries. Moreover, some of these LDCs have exchange rates pegged to more developed countries, so they cannot print more money to finance their expenditures. All these factors contrive to make them aid dependent. Of course, tax revenues are still the most important source of financing expenditures in developing countries. Nevertheless, low- and middle-income countries have the capability to attract more *suitable* complementary funds (FDI, foreign borrowing) than LDCs, reason why the latter depend so much on aid.

Domestic revenue

Royalties and other revenue from natural resources are important in countries endowed with natural resources. Thus, we re-estimate Equation (14) with two new measures of revenue: total tax revenue and total government revenue. The former comprises revenue from non-resource and resource taxes (both direct and indirect tax components), while the latter is a composite of tax and non-tax revenues. Results are reported in Table 6. Regarding total tax revenue (Column 2), the

results are broadly similar to those with non-resource taxes as the measure of domestic revenue (see Table 3). There is evidence of an equilibrium (cointegrating) relationship between aid, total tax revenue and spending. Aid has a positive long-run and short-run impact on spending, with the coefficients on total taxes higher than those on aid. When total government revenue (Column 3) is used, the results are similar, except that now, there is no significant long-run impact of aid on spending. Additionally, the coefficients on all variables are now smaller. There are many competing uses of total domestic revenue (debt servicing, loan repayment, accumulation of reserves), of which government spending is just one. This may explain the smaller average on spending when total domestic revenue is used.

Table 6: Heterogeneity in domestic revenue

	Total tax revenue	**Total domestic revenue**
Long-run		
Tax	0.805*** [0.088]	0.390*** [0.087]
Aid	0.032** [0.014]	0.019 [0.014]
Short-run		
Tax	0.538*** [0.053]	0.300*** [0.072]
Aid	0.021*** [0.007]	0.023*** [0.009]
EC coefficient		
y_{it-1}	-0.789*** [0.039]	-0.635*** [0.050]
t-statistic	-20.14	-12.59
Diagnostics		
RMSE	0.078	0.092
CD test (p-value)	-0.22 (0.826)	1.63 (0.104)
Observations	1961	1940

Note: 'Total tax revenue' and 'Total domestic revenue' refer to error correction models for all 69 countries, first with total tax revenue (including natural resource taxes) as the measure of domestic revenue (Column 2), then total government revenue as the measure of domestic revenue (Column 3). The CCEMG with two lags of cross-section averages is used. For all other details, see Table 3.

Source: author's calculations.

5.4 Weak exogeneity testing

As with all empirical studies, there are concerns about endogeneity. So far, we have discussed one type of endogeneity, whereby the unobserved common factors drive both the dependent variable and the independent variables ($\gamma_i \neq 0, \Phi_i \neq 0$). Another source of endogeneity may be reverse causality. In the aid literature, this reverse causality arises because of the non-random allocation of aid (Temple and Van de Sijpe 2017). We are interested in investigating if donors respond to recipients' fiscal imbalances when disbursing aid, or if disbursement is independent of the fiscal situation in recipient countries. In a simplified version we can express Equations (9) and (10) as follows:

$$y_{it} = \beta_i x_{it} + u_{it} \qquad u_{it} = \alpha_i + \gamma_i \Gamma_t + \varepsilon_{it} \tag{15}$$

$$x_{it} = \pi_i + \lambda_i g_t + \rho_i \Gamma_t + \phi_i \varepsilon_{it} + \nu_{it} \tag{16}$$

for any independent variable x and a single factor Γ driving both y and x in Equations (15) and (16), respectively. Baseline estimates from Table 3 can be interpreted as covariate(s) x having an 'impact' on covariate y. However, nothing is known about the direction of causality in the model yet. As $\phi_i \varepsilon_{it}$ is also present in Equation (16), we are uncertain whether x 'causes' y or y 'causes' x or both.

A standard instrumental variable (IV) approach can address the endogeneity problem, whereby variable x is instrumented for by a set of variables z. Crucially, the instrument(s) must be *valid* (that is they should be uncorrelated with the error term) and *informative* (they should have explanatory power over the endogenous variable(s)). Given the unavailability of adequate instruments (Temple 1999), it is common in the aid effectiveness literature to use lagged aid as an instrument for contemporaneous aid. Nonetheless, in cases where past disbursement patterns do not adequately predict current patterns, there is a weak instruments problem, with severe implications for consistency of the IV estimator. Hence, Temple and Van de Sijpe (2017) fill the gap on 'unavailable and/or inadequate instruments' using a supply-push instrument for aid. To the best of our knowledge, they are the first study to use that instrument in a heterogeneous panel regression framework. They then apply the common correlated effects IV (CCEIV) estimator to address the two types of endogeneity.

Using aid data from the OECD DAC, the supply-push instrument is constructed as follows. For each donor, they calculate the average of the annual shares of any given recipient country in a donor's aid for an initial period (from 1960 to 1970) and multiply that average by the donor's current budget (the sum of the donor's aid over all recipients in period *t*). They then sum this composite across different donors, yielding a time-varying measure of aid used to instrument for net aid in the regressions. This means that, for each recipient country, the instrument approximates the aid that the recipient would have received at each date if its shares in the various donors' budgets had remained constant. This instrument is valid [$E(z, \varepsilon) = 0$], as the total aid budget of most donors is orthogonal to individual, time-varying characteristics of recipient countries, and informative [$E(x, z) \neq 0$], as there are multiple donors considered in the analysis, with different budget shares.

In this study, we are agnostic about the exogeneity and endogeneity status of aid and other fiscal variables. Provided the variables are nonstationary and cointegrated, we can test for weak exogeneity, which entails testing for the direction of causality (Canning and Pedroni 2008; Eberhardt and Presbitero 2015). The Granger representation theorem (Engle and Granger 1987) states that at least one variable must adjust to maintain an equilibrium relation, making it intuitive to know variables that adjust to maintain equilibrium and those that are exogenous for the equilibrium.

If there exists a cointegrating relationship between variables, the Granger representation theorem (Engle and Granger 1987) states that these variable series can be represented in the form of a dynamic ECM. For cointegrated variables, we can then test for weak exogeneity in the following models:

$$\Delta y_{it} = \rho_{1i} + \theta_{1i}\hat{e}_{it-1} + \sum_{j=1}^{K} \lambda_{11ij} \Delta y_{it-j} + \sum_{j=1}^{K} \lambda_{12ij} \Delta x_{it-j} + \epsilon_{1it} \quad (17)$$

$$\Delta x_{it} = \rho_{2i} + \theta_{2i}\hat{e}_{it-1} + \sum_{j=1}^{K} \lambda_{21ij} \Delta y_{it-j} + \sum_{j=1}^{K} \lambda_{22ij} \Delta x_{it-j} + \epsilon_{2it} \quad (18)$$

Where ρ_is are constant terms and $\hat{e}_{it-1}$ is the disequilibrium term $\hat{e} = y - \hat{\beta}_i x - \hat{d}$ constructed using the cointegrating relationship between the variables (*d* represents deterministic terms like a constant and a country-specific trend). The disequilibrium term represents how far the variables

are from the equilibrium relationship, with the error correction mechanism then indicating the speed of adjustment following a deviation from the long-run equilibrium (Canning and Pedroni 2008). Each variable may react to its lagged differences, as well as lagged differences of other variables in the cointegrating relationship. The Granger representation theorem implies that at least one of the adjustment coefficients θ_{1i}, θ_{2i} must be non-zero if a cointegrating (long-run) relationship between the variables is to hold (Canning and Pedroni 2008: 512). If $\theta_{1i} \neq 0$, then x has a causal impact on y (which in our case means expenditures adjust to maintain fiscal equilibrium following an imbalance), and if $\theta_{2i} \neq 0$, then y has a causal impact on x (in our case this will mean donors disburse aid in response to budgetary disequilibria in recipient countries, and that taxes adjust to such fiscal disequilibria as well). If both θ_{1i} and θ_{2i} are non-zero, then x and y determine each other jointly.

One of the advantages of using the disequilibrium term from a cointegrating relationship is that all the variables in Equations (17) and (18) are stationary. This means once ECMs are estimated for each country, estimates for θ_i can be investigated using standard t-ratios (Canning and Pedroni 2008). Following Canning and Pedroni (2008), we use two separate statistics to test for weak exogeneity. The first is the group-mean statistic (*GM* hereafter) which averages the θ_i from individual country estimations of Equations (17) and (18), and the *GM* test for the null of 'no long-run causal impact' is computed from the averaged *t*-ratio from country regressions ($\bar{t}_{\theta_2} = N^{-1}\sum_{i=1}^{N} t_{\theta_2}$). The *GM* statistic follows a standard normal distribution under the null hypothesis of 'no causal impact'. The second statistic is a Fisher-type (*Lambda*-Pearson) statistic which is constructed from the *p*-values of the *t*-tests from the country regressions to get the overall marginal significance associated with those *p*-values. The Fisher statistic follows a χ^2 distribution with *2N* degrees of freedom under the null hypothesis of 'no causal impact'.

The null and alternative hypotheses for both tests are the same when the θ_i coefficients are the same for all members of the panel. This translates into a null that $\theta_i = 0$ for all members in the panel against an alternative $\theta_i \neq 0$ for some non-negligible members of the panel (Canning and Pedroni 2008). The interpretation of the tests, however, differs when θ_i differs across countries. The *GM* test is a two-tailed test so can take on positive or negative values under the null and alternative hypotheses depending on whether $\hat{\theta}_i$ is positive or negative, whereas the Fisher statistic is a one-tailed test that only takes positive values in both the null and alternative hypotheses. If these two tests fail to agree on the direction of causality between variables, this can be interpreted as θ_i being on average zero (allowing for large negative and positive values to cancel each other), but not pervasively zero in the panel (Canning and Pedroni 2008; Eberhardt and Presbitero 2015). If that is the case, it provides evidence of the heterogeneity of θ_i across countries. We also report the robust $\hat{\theta}_i$ estimate, and its associated t-statistic. In the last column of Table 7 we report the share of countries in the sample (N_i/N) that fail to reject the null of 'no causal' impact.

Table 7: Weak exogeneity tests

	GM	*p*	Fisher	*p*	Mean $\hat{\theta}_i$	*t*-stat	N_i/N
Expenditure Equation 1	-1.459	0.14	361.20	0.00	-0.782	-9.95	58%
Expenditure Equation 2	-1.464	0.14	396.61	0.00	-0.755	-8.50	62%
Expenditure Equation 3	-1.132	0.26	282.22	0.00	-0.624	-8.97	64%
Tax Equation 1	0.270	0.79	171.34	0.03	0.156	1.92	83%
Tax Equation 2	0.098	0.92	160.36	0.09	0.078	0.96	83%
Tax Equation 3	0.205	0.84	151.66	0.12	0.200	2.63	73%
Aid Equation	-0.053	0.96	146.23	0.30	-0.193	-0.75	88%
Grants Equation	-0.154	0.88	145.41	0.32	-0.226	-0.99	54%
TC Equation	0.190	0.85	137.80	0.49	0.057	0.69	80%
Loans Equation	0.034	0.97	160.88	0.04	0.069	0.17	65%
Consumption Equation	-1.776	0.08	449.37	0.00	-0.710	-13.11	55%
Capital Equation	-1.703	0.09	428.57	0.00	-0.841	-10.88	54%

Note: we report both statistics developed by Canning and Pedroni (2008). *GM* presents the group-mean statistic which is the average of country-specific *t*-ratios on the disequilibrium term which is distributed *N(0,1)*. Fisher is $-2\sum_{i=1}^{N} ln\Pi$, where Π is the *p*-value of the country-specific t-value on the disequilibrium term. The Fisher statistic is distributed $\chi^2(2N)$. Both test statistics are for the null of 'no causal impact' which in our case can be interpreted as the variable not adjusting to maintain fiscal equilibrium. We also report the robust $\hat{\theta}_i$ estimate, and its associated t-statistic. In the last column we report the percentage of countries in the sample that fail to reject the null of 'no causal' impact.

Source: author's calculations.

We present results for weak exogeneity tests in Table 7, using specifications of Equations (17) and (18) with two lags. The results are based on the dynamic CCEMG model augmented with two lags of cross-section averages (this is the long-run relationship from which the disequilibrium term is constructed). 'Expenditure Equations' (1), (2) and (3) refer to the ECM regression with government expenditure as dependent variable and net ODA, grants and loans as aid variables, respectively. 'Tax Equations' are estimated analogously. The 'Aid Equation' is estimated with net aid as dependent variable and taxes and expenditures as independent variables. Grants, Loans and TC Equations are estimated analogously. 'Consumption' and 'Capital Equations' are estimated with net aid and taxes as independent variables and recurrent spending and capital spending as dependent variables, respectively. We then report the two statistics developed by Canning and Pedroni (2008), the *GM* and Fisher statistics, as well as their respective *p*-values. We also report the robust mean $\hat{\theta}_i$ with its *t*-statistic: we would expect a typically high *t*-statistic on the average $\hat{\theta}_i$ coefficients in the expenditure equation (which can be interpreted as evidence of a long-run causal relationship from aid/taxes to spending) and a low *t*-statistic (below 1.96) in the other equations (Eberhardt and Presbitero 2015).

For the expenditure equations, the *GM* test fails to reject the null of 'no causal impact' from aid (grants and loans) and taxes, whereas the Fisher statistic rejects the null of 'no causal impact'. This offers insights into the cross-country heterogeneity of θ_i. Moreover, the *t*-statistic on the robust $\hat{\theta}_i$ is high for the three expenditure equations, indicating that there is potential long-run causality from aid and taxes to spending. However, approximately 58% of countries in the sample fail to reject the null of 'no causal impact' from x to y when net aid is the main regressor of interest; 62% of countries fail to reject the null in the case of grants; and 64% fail to reject the null in the case of loans. As these test results are not overwhelmingly in favour of (or against) weak exogeneity, we discuss both possibilities. On the one hand, spending policies are often prepared for the medium to long term, meaning spending (especially statutory expenditures and public payroll) is not easily reversed once implemented (Bwire 2012). In the fiscal response context, this implies that expenditures do not adjust to maintain the budgetary equilibrium. On the other hand, spending

may be endogenous to the long-run equilibrium in situations where government spending is planned based on the expected revenue packet but spending allocation is affected when the revenue outcome is realized (Bwire 2012). This latter pattern of results (endogeneity for the long-run equilibrium) is largely replicated when expenditures are disaggregated into capital and consumption expenditures.

In the first two tax equations, the *GM* test fails to reject the null of 'no causal impact' from expenditures and aid to taxes, while the Fisher statistic rejects the null. Following Canning and Pedroni (2008), this can be interpreted as taxes being weakly exogenous on average, but not *pervasively* weakly exogenous in the sample. When loans are the aid variable of interest, the tax equation shows that both the *GM* and Fisher statistics fail to reject the null of 'no causal impact'. When net aid is used as measure for aid, 83% of the countries in the sample cannot reject the null of 'no causal impact'. When grants and loans are used, respectively, 83% and 73% of the countries in the sample fail to reject the null of 'no causal impact' from expenditures and grants/loans to taxes. Additionally, the *t*-statistic on the robust mean $\hat{\theta}_i$ coefficients is less than 1.96 (evidence of no causal impact). This provides tentative evidence that tax revenues are weakly exogenous. Tax systems are also statutory and are not easily changed once implemented. Faced with a deficit, recipients cannot just increase tax rates as such bills usually need to be agreed upon by congress, and that is a time-consuming process. Aid may also impact tax administration such that tax rates do not increase but the collection efficiency of taxes increases. Nonetheless, improvements in efficiency also take time, meaning taxes would not immediately adjust to fiscal imbalance.

When the tests are applied to net aid (as well as grants and technical cooperation), both test statistics, in addition to the robust mean θ_i coefficient, agree that the three aid variables are indeed weakly exogenous. This provides insights into the disbursement behaviour of donors. Aid plays an important role in determining the budget, but its level does not reflect budget imbalance in recipient countries. Nonetheless, as shown in Table 3 aid influences other variables in the system (it is *long-run forcing*). While fiscal planners in recipient countries have a planned target for aid revenue (as portrayed by the long-run relationship), they take the aid as given. Donors do not adjust the level of aid to recipients, but possibly adjust how the aid is delivered (i.e., the modality of aid) according to certain recipients' characteristics (such as rule of law and government effectiveness). In the case of loans, the *GM* test fails to reject the null of no causal impact, while the Fisher statistic rejects the null. However, 65% of countries in the sample fail to reject the null of 'no causal impact', implying we can conclude that loans, like grants and net aid, are weakly exogenous. These results depict pervasive cross-country heterogeneity in the form of disequilibrium such that it could be a budget surplus in some countries while it is a deficit in other countries. Nevertheless, fiscal variables do not adjust to maintain equilibrium in the face of such imbalances.

6 Conclusion

This paper investigated the nature of the relationship between aid, taxes and spending using novel panel time series methods to address the problems of cross-country heterogeneity, cross-section dependence, and variable non-stationarity. Using annual data for 69 developing countries over the period 1980 to 2013, we provide evidence of an average long-run (cointegrating) relationship between aid, taxes and spending, which differs across countries. Estimates show that on average, aid has both a positive long-run and short-run impact on spending. This average, positive long-run relationship is quite small, indicating that increases in aid are positively associated with marginal increases in spending. Aid also has a short-run impact on spending. In addition, we find that the coefficient on taxes is higher than the coefficient on aid. This implies taxes are the main driver of

domestic expenditures, in the long run as well as the short run. We ascertain that aid positively influences spending, but tax revenue is more strongly associated with recipients' expenditures.

Delving further into the relationship between aid, taxes and spending, we find that the long-run and short-run impact of aid on spending is indeed very large in LDCs (especially SSA countries). This is to be expected as LDCs, of which SSA form the bulk, are aid dependent. Thus, political calculus between aid and taxes in those countries favours more aid than increased taxation. Moreover, the fiscal capacity in these LDCs is relatively low, hence the maximum amount of tax they collect may still not be enough for development gains. While improving tax administration and fiscal capacity in LDCs is very important, aid for financing spending is clearly as important for those countries. Using the estimated long-run equilibrium relationship, we do not find evidence of a donor 'disbursement rule', as donor allocation is independent of the fiscal situation of the recipient (irrespective of the measure of aid used). Aid, as well as its components, is weakly exogenous. This implies donors have a fixed amount of aid they intend to disburse, and this amount does not change in light of the fiscal situation of recipients.

References

Bai, J., Kao, C. and Ng, S. (2009). Panel cointegration with global stochastic trends. *Journal of Econometrics*, *149*(1), 82-99. https://doi.org/10.1016/j.jeconom.2008.10.012

Banerjee, A., and Carrion-i-Silvestre, J. L. (2017). Testing for panel cointegration using common correlated effects estimators. *Journal of Time Series Analysis*, *38*(4), 610-636. https://doi.org/10.1111/jtsa.12234

Bwire, T. (2012). 'Aid, fiscal policy and macroeconomy of Uganda: a cointegrated vector autoregressive (CVAR) approach'. PhD thesis. Nottingham: University of Nottingham.

Bwire, T., Lloyd, T., and Morrissey, O. (2017). Fiscal reforms and the fiscal effects of aid in Uganda. *The Journal of Development Studies*, *53*(7), 1019-1036. https://doi.org/10.1080/00220388.2017.1303677

Canning, D. and Pedroni, P. (2008). Infrastructure, long-run economic growth and causality tests for cointegrated panels. *The Manchester School*, *76*(5), 504-527. https://doi.org/10.1111/j.1467-9957.2008.01073.x

Chudik, A., and Pesaran, M. H. (2015). Common correlated effects estimation of heterogeneous dynamic panel data models with weakly exogenous regressors. *Journal of Econometrics*, *188*(2), 393-420. https://doi.org/10.1016/j.jeconom.2015.03.007

Coakley, J., Fuertes, A.M. and Smith, R. (2006). Unobserved heterogeneity in panel time series models. *Computational Statistics & Data Analysis*,*50*(9), 2361-2380. https://doi.org/10.1016/j.csda.2004.12.015

De Hoyos, R.E. and Sarafidis, V. (2006). Testing for cross-sectional dependence in panel-data models. *Stata Journal*, *6*(4), 482. https://doi.org/10.1177/1536867X0600600403

Eberhardt, M. (2012). Estimating panel time-series models with heterogeneous slopes. *Stata Journal*, *12*(1), 61. https://doi.org/10.1177/1536867X1201200105

Eberhardt, M. and Teal, F. (2011). Econometrics for grumblers: a new look at the literature on cross-country growth empirics. *Journal of Economic Surveys*, *25*(1), 109-155. https://doi.org/10.1111/j.1467-6419.2010.00624.x

Eberhardt, M. and Presbitero, A.F. (2015). Public debt and growth: Heterogeneity and non-linearity. *Journal of International Economics*, *97*(1), 45-58. https://doi.org/10.1016/j.jinteco.2015.04.005

Engle, R.F. and Granger, C.W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: Journal of the Econometric Society*, 251-276. https://doi.org/10.2307/1913236

Feeny, S. (2006). Policy preferences in fiscal response studies. *Journal of International Development*, *18*(8), 1167-1175. https://doi.org/10.1002/jid.1281

Franco-Rodriguez, S., Morrissey, O. and McGillivray, M. (1998). Aid and the public sector in Pakistan: evidence with endogenous aid. *World Development*, *26*(7), 1241-1250. https://doi.org/10.1016/S0305-750X(98)00048-5

Granger, C.W. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics*, *2*(2), 111-120. https://doi.org/10.1016/0304-4076(74)90034-7

Guillaumont, P., Jeanneney, S.G. and Brun, J.F. (1999). How instability lowers African growth. *Journal of African Economies*, *8*(1),87-107. https://doi.org/10.1093/jae/8.1.87

Gyimah-Brempong, K. and Traynor, T.L. (1999). Political instability, investment and economic growth in Sub-Saharan Africa. *Journal of African Economies*, *8*(1), pp.52-86. https://doi.org/10.1093/jae/8.1.52

Hamilton, L.C. (1992). How robust is robust regression? *Stata Technical Bulletin*, *1* (2).

Heller, P.S. (1975). A model of public fiscal behaviour in developing countries: Aid, investment, and taxation. *The American Economic Review*, *65*(3), 429-445.

Hendry, D.F., (1995). *Dynamic econometrics*. Oxford: Oxford University Press. https://doi.org/10.1093/0198283164.001.0001

Herzer, D. (2017). The Long-run relationship between trade and population health: Evidence from five decades. *The World Economy*, *40*(2), 462-487. https://doi.org/10.1111/twec.12419

Herzer, D. and Morrissey, O. (2013). Foreign aid and domestic output in the long run. *Review of World Economics*, *149*(4), 723-748. https://doi.org/10.1007/s10290-013-0169-y

Jones, S. (2015). Aid supplies over time: Addressing heterogeneity, trends, and dynamics. *World Development*, *69*, 31-43. https://doi.org/10.1016/j.worlddev.2013.12.014

Kapetanios, G., Pesaran, M.H. and Yamagata, T. (2011). Panels with non-stationary multifactor error structures. *Journal of Econometrics*, *160*(2), 326-348. https://doi.org/10.1016/j.jeconom.2010.10.001

Keen, M. and Simone, A. (2004). Tax policy in developing countries: some lessons from the 1990s and some challenges ahead. In S. Gupta, B. Clements, and G. Inchauste (eds), *Helping countries develop: The role of fiscal policy*, pp.302-52. Washington, DC: International Monetary Fund.

Keen, M. and Mansour, M., (2009). *Revenue Mobilization in Sub-Saharan Africa; Challenges from Globalization* (No. 09/157). Washington, DC: International Monetary Fund. https://doi.org/10.5089/9781451873047.001

Lloyd, T., McGillivray, M., Morrissey, O. and Opoku-Afari, M., (2009). The fiscal effects of aid in developing countries: A comparative dynamic analysis. In G. Mavrotas and M. McGillivray (eds) M. *Development Aid*, pp. 158-179. London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230595163_7

Lütkepohl, H. (2007). General-to-specific or specific-to-general modelling? An opinion on current econometric terminology. *Journal of Econometrics*, *136*(1), 319-324. https://doi.org/10.1016/j.jeconom.2005.11.014

McGillivray, M. and Morrissey, O., (2001). *Fiscal effects of aid*. Helsinki: United Nations University, World Institute for Development Economics Research (UNU-WIDER).

Martins, P.M., (2010). *Fiscal dynamics in Ethiopia: the cointegrated VAR model with quarterly data* (No. 10/05). CREDIT Research Paper. Nottingham: Centre for Research in Economic Development and International Trade, University of Nottingham.

Mascagni, G., and Timmis, E. (2017). The fiscal effects of aid in Ethiopia: Evidence from CVAR applications. *The Journal of Development Studies*, *53*(7), 1037-1056. https://doi.org/10.1080/00220388.2017.1303676

Morrissey, O., (2012). *Aid and government fiscal behaviour: What does the evidence say?* (No. 2012/01). WIDER Working Paper. Helsinki: UNU-WIDER.

Morrissey, O., (2015a). Aid and government fiscal behaviour: assessing recent evidence. *World Development*, *69*, 98-105. https://doi.org/10.1016/j.worlddev.2013.12.008

Morrissey, O., (2015b). Aid and domestic resource mobilization with a focus on Sub-Saharan Africa. *Oxford Review of Economic Policy*, *31*(3-4), 447-461. https://doi.org/10.1093/oxrep/grv029

Morrissey, O., Isopi, A. and Clist, P., (2011). Aid and Government Spending. Report A to Agence Francaise de Developpement (Paris).

Morrissey, O., Prichard, W. and Torrance, S., (2014). Aid and Taxation: Exploring the relationship using new data. International Centre for Taxation and Development (ICTD), ICTD Working Paper 21 (www.ictd.ac). https://doi.org/10.2139/ssrn.2496453

Morrissey, O., and Torrance, S. (2015), 'Aid and Taxation', ch. 31 in B. M. Avin and B. Lew (eds), Handbook on the Economics of Foreign Aid. Cheltenham: Edward Elgar.

Mosley, P., Hudson, J. and Horrell, S. (1987). Aid, the public sector and the market in less developed countries. *The Economic Journal*, *97*(387), 616-641. https://doi.org/10.2307/2232927

Murdoch, J.C. and Sandler, T., (2002). Economic growth, civil wars, and spatial spillovers. *Journal of Conflict Resolution*, *46*(1), 91-110. https://doi.org/10.1177/0022002702046001006

Osei, R., Morrissey, O. and Lloyd, T. (2005). The fiscal effects of aid in Ghana. *Journal of International Development*, *17*(8), 1037-1053. https://doi.org/10.1002/jid.1258

Pesaran, M.H. (2006). Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, *74*(4), 967-1012. https://doi.org/10.1111/j.1468-0262.2006.00692.x

Pesaran, M.H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics*, *22*(2), 265-312. https://doi.org/10.1002/jae.951

Pesaran, M. H. (2015). Testing weak cross-sectional dependence in large panels. *Econometric Reviews*, *34*(6-10), 1089-1117. https://doi.org/10.1080/07474938.2014.956623

Pesaran, M. H., and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, *68*(1), 79-113. https://doi.org/10.1016/0304-4076(94)01644-F

Prichard, W., Cobham, A. and Goodall, A., (2014). ICTD Government Revenue Dataset ICTD working paper 19. *Institute of Development Studies, Brighton*. https://doi.org/10.2139/ssrn.2496442

Remmer, K.L. (2004). Does foreign aid promote the expansion of government?. *American Journal of Political Science*, *48*(1), 77-92. https://doi.org/10.1111/j.0092-5853.2004.00057.x

Roger, L., (2015). Foreign Aid, Poor Data, and the Fragility of Macroeconomic Inference. *CREDIT Research Paper*, (15/06). https://doi.org/10.2139/ssrn.2781908

Smith, R.P. and Fuertes, A.-M. (2007) Panel time series. Mimeo, April, Centre for Microdata Methods and Practice (cemmap).

Söderbom, M., Teal, F., Eberhardt, M., Quinn, S. and Zeitlin, A. (2014).*Empirical development economics*. London: Routledge. https://doi.org/10.4324/9780203070925

Temple, J., and Van de Sijpe, N. (2017). Foreign aid and domestic absorption. *Journal of International Economics*, *108*, 431-443. https://doi.org/10.1016/j.jinteco.2017.07.010

Temple, J. (1999). The new growth evidence. *Journal of Economic Literature*,*37*(1), 112-156. https://doi.org/10.1257/jel.37.1.112

Van de Sijpe, N. (2013). Is Foreign Aid Fungible? Evidence from the Education and Health Sectors. *The World Bank Economic Review*, *27*(2), 320-356. https://doi.org/10.1093/wber/lhs023

Williams, G. (1994). Why structural adjustment is necessary and why it doesn't work. *Review of African Political Economy*, 60, 214–25. https://doi.org/10.1080/03056249408704057

Appendix

Table A1: List of countries

Sub-Saharan Africa (SSA)	Middle East and North Africa (MENA)	Asia and the Pacific	Latin America and the Caribbean (LAC)
1. Angola 2. Benin 3. Botswana 4. Burkina Faso 5. Burundi 6. Central African Republic 7. Chad 8. Comoros 9. Congo, Dem Rep 10. Congo Rep 11. Côte d'Ivoire 12. Equatorial Guinea 13. Gabon 14. Gambia, The 15. Ghana 16. Guinea 17. Guinea-Bissau 18. Kenya 19. Lesotho 20. Madagascar 21. Malawi 22. Mauritania 23. Mauritius 24. Mozambique 25. Niger 26. Rwanda 27. Senegal 28. Seychelles 29. Sudan 30. Swaziland 31. Togo 32. Uganda 33. Zimbabwe	1. Algeria 2. Egypt 3. Iran, Islamic Republic 4. Jordan 5. Morocco 6. Turkey	1. Bangladesh 2. China 3. Fiji 4. India 5. Indonesia 6. Nepal 7. Pakistan 8. Philippines 9. Sri Lanka 10. Thailand 11. Vanuatu	1. Argentina 2. Barbados 3. Belize 4. Chile 5. Colombia 6. Costa Rica 7. Dominica 8. Dominican Republic 9. Ecuador 10. El Salvador 11. Guatemala 12. Honduras 13. Jamaica 14. Mexico 15. Nicaragua 16. Panama 17. Peru 18. Uruguay 19. Venezuela

Source: author's calculations.

Table A2: Classification of countries according to level of development

Low-income countries	Non-low-income countries: (GNIpc <= $1,045 in 2013)	Lower middle-income countries (LMICs): (GNI*pc* $1,046 - $4,125 in 2013)	Upper middle-income countries (UMICs): (GNIpc $4,126 - $12,745 in 2013)
1. Angola 2. Bangladesh 3. Benin 4. Burkina Faso 5. Burundi 6. Central African Republic 7. Chad 8. Comoros 9. Congo, Dem Rep 10. Equatorial Guinea 11. Gambia, The 12. Guinea 13. Guinea-Bissau 14. Lesotho 15. Madagascar 16. Malawi 17. Mauritania 18. Mozambique 19. Nepal 20. Niger 21. Rwanda 22. Senegal 23. Sudan 24. Togo 25. Uganda 26. Vanuatu	1. Kenya 2. Zimbabwe	1. Congo Rep 2. Côte d'Ivoire 3. Egypt 4. El Salvador 5. Ghana 6. Guatemala 7. Honduras 8. India 9. Indonesia 10. Morocco 11. Nicaragua 12. Pakistan 13. Philippines 14. Sri Lanka 15. Swaziland	1. Algeria 2. Argentina 3. Barbados 4. Belize 5. Botswana 6. Chile 7. China 8. Colombia 9. Costa Rica 10. Dominica 11. Dominican Republic 12. Ecuador 13. Fiji 14. Gabon 15. Iran 16. Jamaica 17. Jordan 18. Mauritius 19. Mexico 20. Panama 21. Peru 22. Seychelles 23. Thailand 24. Turkey 25. Uruguay 26. Venezuela

Note: GNIpc refers to *per capita* GNI. For estimation purposes, 'Other low-income countries', 'Lower middle-income countries' and 'Upper middle-income countries' are grouped into 'Other low-income countries' (see Table 5 of Subsection 5.3).

Source: author's calculations.

Table A3: Variables, descriptions and data sources

Variable name	Variable description	Data source
Aid	Net aid (net of repayments)	OECD DAC
Grants	Gross ODA grants	OECD DAC
Loans	Gross ODA loans	OECD DAC
Technical cooperation	Technical cooperation	OECD DAC
Tax revenue	Non-resource tax revenue	GRD Database
Total tax revenue	Resource + non-resource tax revenue	GRD Database
Total government revenue	Government revenue excluding grants	GRD Database
Spending	Total government spending	World Development Indicators
Capital spending	Public investment	World Development Indicators
Recurrent (consumption) spending	Government consumption spending	World Development Indicators

Source: author's calculations.

Table A4: Summary statistics

Panel A: Raw (non-logged) variables

Variable	Mean	S.D.	Min.	Max.
Aid	4.69e+08	7.14e+08	-1.40e+09	8.77e+09
Tax revenue	1.25e+10	6.52e+10	5982453	1.69e+12
Spending	4.20e+10	2.34e+11	5.05e+07	5.44e+12
Gross loans	3.18e+08	6.01e+08	0.000	9.44e+09
Gross grants	3.23e+08	5.35e+08	-582187	8.08e+09
TC	1.35e+08	1.63e+08	1357134	1.46e+09
Cap. spending	3.00e+10	1.81e+11	0.000	4.24e+12
Cons. spending	1.21e+10	5.43e+10	0.000	1.21e+12

Panel B: Regression variables (in logs or first differences of logs)

Variable	Mean	S.D.	Min.	Max.
y_{it-1}	22.113	2.013	17.738	29.107
Aid_{it-1}	18.020	4.663	0.000	22.895
Tax_{it-1}	21.034	2.031	15.604	27.952
Cap_{it-1}	21.577	2.137	16.598	28.862
$Cons_{it-1}$	21.150	1.913	16.872	27.577
$Loans_{it-1}$	18.260	1.937	9.321	22.969
$Grants_{it-1}$	18.686	1.471	12.313	22.812
TC_{it-1}	18.172	1.134	14.121	21.105
Δy_{it}	0.037	0.224	-1.421	1.974
ΔAid_{it}	-0.035	3.611	-21.143	21.566
ΔTax_{it}	0.040	0.213	-1.682	2.232
ΔCap_{it}	0.038	0.287	-2.380	2.914
$\Delta Cons_{it}$	0.031	0.209	-1.894	2.314
$\Delta Loans_{it}$	-0.034	0.904	-8.028	5.166
$\Delta Grants_{it}$	0.029	0.548	-3.293	3.181
ΔTC_{it}	-0.014	0.238	-1.281	1.095

Note: descriptive statistics are presented for the full sample of 2,346 observations from N= 69 countries (average T=34). In Panel A the variables are non-logged. In Panel B we report descriptive statistics for the main ECM regression variables (as well as those used for exploratory analysis), namely Δy_{it}—expenditure growth rate; y_{it-1}—lagged level of government expenditure; Aid_{it-1}—lagged level of net aid; Tax_{it-1}—lagged level of taxes; ΔAid_{it}—growth rate of net aid; ΔTax_{it}—growth rate of taxes. All other variables used for exploratory analysis follow the same interpretation.

Source: author's calculations.

Table A5: Cross-section dependence

Panel A	**Variables in levels**			
	$Grants_{it}$	$Loans_{it}$	$Cons_{it}$	Cap_{it}
avg $\hat{\rho}_{ij}$	0.247	0.133	0.446	0.485
avg $\lvert\hat{\rho}_{ij}\rvert$	0.333	0.260	0.542	0.555
CD	66.10	34.39	116.41	126.84
p-value	0.00	0.00	0.00	0.00
Panel B	**Variables in first differences**			
	$\Delta Grants_{it}$	$\Delta Loans_{it}$	$\Delta Cons_{it}$	ΔCap_{it}
avg $\hat{\rho}_{ij}$	0.046	0.015	0.086	0.097
avg $\lvert\hat{\rho}_{ij}\rvert$	0.175	0.176	0.182	0.176
CD	11.94	4.01	22.32	25.25
p-value	0.00	0.00	0.00	0.00

Note: we use the stata routine 'xtcd' developed by Markus Eberhardt. We report the average correlation and average absolute correlation coefficients of the $N(N-1)$ sets of correlations. CD is the Pesaran (2004) test for cross-section dependence distributed $N(0,1)$ under the null of cross-section independence. Panels A and B test for cross-section dependence in the variable series for levels and first differences, respectively. Grants, loans, recurrent expenditures (Cons), capital expenditures (Cap).

Source: author's calculations.

Table A6: Panel unit root tests

Levels: CIPS with intercept only								
Variable	Grants		Loans		Cons		Cap	
Lags	Z*t*bar	*p*	Z*t*bar	*p*	Z*t*bar	*p*	Z*t*bar	*p*
0	-12.46	0.00	-11.43	0.00	0.37	0.64	-3.78	0.00
1	-7.18	0.00	-5.22	0.00	-1.68	0.05	-4.06	0.00
2	-3.60	0.00	-3.65	0.00	-0.81	0.21	-1.72	0.04
3	-2.30	0.01	-3.55	0.00	-0.84	0.20	-1.21	0.11
4	0.94	0.83	-0.60	0.28	-0.42	0.34	0.30	0.62
Levels: CIPS with intercept & trend								
Variable	Grants		Loans		Cons		Cap	
Lags	Z*t*bar	*p*	Z*t*bar	*p*	Z*t*bar	*p*	Z*t*bar	*p*
0	-11.29	0.00	-9.68	0.00	-0.20	0.42	-2.95	0.00
1	-5.58	0.00	-1.96	0.03	-1.67	-1.66	-3.84	0.00
2	-1.32	0.09	-0.87	0.19	-0.04	0.49	-0.88	0.19
3	0.65	0.74	0.48	0.69	0.40	0.66	-0.04	0.48
4	4.47	1.00	3.82	1.00	-0.33	0.37	1.58	0.94
Differences: CIPS test with drift								
Variable	Grants		Loans		Cons		Cap	
Lags	Z*t*bar	*p*	Z*t*bar	*p*	Z*t*bar	*p*	Z*t*bar	*p*
0	-36.83	0.00	-35.86	0.00	-27.10	0.00	-31.96	0.00
1	-27.47	0.00	-23.46	0.00	-16.75	0.00	-21.02	0.00
2	-16.98	0.00	-13.44	0.00	-8.92	0.00	-12.89	0.00
3	-12.93	0.00	-7.63	0.00	-5.73	0.00	-8.04	0.00
4	-8.47	0.00	-4.11	0.00	-4.51	0.00	-4.46	0.00

Note: grants, loans, recurrent spending (Cons) and capital spending (Cap) all in logs. 'Lags' denote the number of lags of the differenced dependent variable included to wipe out serial correlation. H_0 = non-stationarity in all countries' variable series; H_1 = stationarity in some countries' variable series.

Source: author's calculations.