FACULTY OF SOCIAL SCIENCES
Department of Economics
University of Copenhagen

PhD Thesis

Fiseha Haile Gebregziabher

Essays in Empirical Development Economics

Insights from Time Series and Panel Data Analysis

Submitted to the Department of Economics, University of Copenhagen, in partial fulfillment of the requirements for the PhD degree in Economics.

Academic Advisor: Katarina Juselius

Submitted: May 2014

Contents

Acknowledgments

This dissertation was written during my three years of PhD studies (2011 − 2014) at the Department of Economics, University of Copenhagen. First and foremost, I am grateful to God that this project has finally come to fruition. Many people deserve my sincere thanks for their encouragements and support throughout this process. I would first like to grab this opportunity to extend my deepest gratitude to my supervisor, Katarina Juselius, for her excellent guidance and unwavering support along the way. Her insightful comments and suggestions have strongly influenced the entire thesis.

In autumn 2012, I visited the University of Nottingham and the World Institute for Development Economics Research (UNU-WIDER). I gratefully acknowledge the financial support provided by the UNU-WIDER to cover transportation costs to and from Helsinki, and living expenses in Helsinki during my visit. I really enjoyed my stay abroad and I would like to thank everyone in these institutions for their outstanding hospitality. In particular, Oliver Morrissey and Finn Tarp deserve credit for being instrumental in organizing the visits to the University of Nottingham and UNU-WIDER respectively as well as for their academic and personal assistance during the stay. Oliver's comments on earlier versions of the first two chapters in this dissertation are gratefully appreciated. This thesis also benefited a lot from a number of contructive comments from seminar and conference participants, both in and outside Denmark.

Special thanks are due to my parents, Haile Gebregziabher and Nigisti Gebregziabher, for making education a top priority in our home and for their unflinching support and love. I also wish to thank my brothers and sisters for their unfailing love and encouragement. In particular, special appreciation goes to Dawit for making the past two and a half years of my stay in Copenhagen more enjoyable and for help in proofreading and editing each manuscript. Who would have thought just a few years ago that both of us would end up here in Copenhagen?

Finally, I owe thanks to friends and colleagues at the Department of Economics for making my time as a Ph.D. student so enjoyable. I am also grateful to my best friends in Copenhagen who have been there for me in all dimensions of my life.

Fiseha Haile Gebregziabher
Copenhagen, May 2014

Summary

This thesis, entitled *Essays in Empirical Development Economics: Insights from time series and panel data analysis*, contains three articles focusing on empirical questions in development economics. Each chapter is self-contained and can be read independently. The first chapter investigates the long-run effects of foreign aid flows in Ethiopia. The second chapter examines the link between IMF-World Bank adjustment programs and long-run economic performance in 18 Sub-Saharan African (SSA) countries. The common thread between these two papers is the use of the cointegrated vector autoregressive (CVAR) model as a statistical benchmark. The third and final chapter assesses the impact of social spending on aggregate human welfare. In terms of methodology, this paper lies squarely within the standard cross-country regression of the effectiveness of social spending. The remainder of this section provides brief summaries of the manuscripts.

Chapter 1, **The Long-Run Macroeconomic Effects of Aid and Disaggregated Aid in Ethiopia**, investigates the long-run impact of foreign aid flows in Ethiopia for the period $1960 - 2009$. Notwithstanding the voluminous and multifaceted literature on aid effectiveness, the question of whether foreign aid works is still mired in controversy. We attempt to advance this longstanding debate by analyzing whether aid from different sources (multilateral and bilateral aid) and different aid modalities (grant and loan) exert different long-run effects on several macroeconomic variables. Using a well-specified CVAR model as a statistical framework, we show that aid affects Gross Domestic Product (GDP), investment, and imports positively, whereas it is negatively associated with government consumption spending. Disaggregating aid by source and type, we find results that stand in stark contrast to the findings reported in most previous studies. Bilateral aid increases investment and GDP, and is negatively linked with government consumption, whereas multilateral aid is only positively associated with imports. Grants contribute to GDP, investment, and imports, whereas loans affect none of the variables.

Chapter 2, **Adjustment and Long-Run Economic Performance in 18 African Countries**, examines the nexus between IMF-World Bank adjustment programs and long-run economic performance in 18 SSA countries for the period $1960 - 2009$. During the 1980s and 1990s, IMF and World Bank-sponsored adjustment programs were ubiquitous in SSA. Nearly-all SSA countries adopted a wide range of policy reforms to arrest severe macroeconomic imbalances and foster long-term economic growth. Existing studies have almost universally used cross-country regressions to analyze the impact of adjustment reforms. This paper contributes to the literature by investigating how the introduction of adjustment policy reforms affected the long-run growth trajectories in a sample of 18 SSA countries on a country-by-country basis. Applying

our multivariate cointegration model to each of these countries, we find that only a handful of countries have shown positive and sustained results. The traditional (*first-generation*) Fund-Bank adjustment package has been associated with resurgence of long-run growth in GDP, export, and investment only in two countries (Ghana and Uganda). Most African economies remained on their *pre*-reform growth paths whereas some others saw their long-run growth rates plummet, despite having successfully implemented more-than-a-decade-long adjustment reforms. Taken as a whole, countries in the CFA franc currency zone fared much worse than their non-CFA counterparts.

Chapter 3, **Social Spending and Aggregate Welfare in Developing and Transition Economies** (joint with Miguel Niño-Zarazúa), analyzes the impact of government spending on social sectors (health, education, and social protection) on two major indicators of aggregate welfare (the *Inequality-adjusted* Human Development Index (IHDI) and child mortality), using a panel dataset comprising 55 developing and transition countries from 1990 to 2009. Notwithstanding the unprecedented attention devoted to fostering human development via scaling up social sector spending, the evidence on the relationship between social spending and welfare outcomes remains inconclusive. This paper reconsiders the empirical evidence on the nexus between social spending and aggregate human welfare. The data on social spending come from the *Government Finance Statistics* database and are given in local currency units (LCU). Unlike previous studies, we transform the data in LCU into purchasing power parity dollars. We show that social spending has a significantly positive causal effect on the IHDI, while government expenditure on health has a significant negative impact on child mortality rate. The preferred (System GMM) specification indicates that a 1 percent increase in social spending, in percent of GDP, increases the IHDI by 0.004 points, which appears modest, albeit not negligible. The long-run effect of a similar increase in social spending is an increase in the IHDI of about 0.057 points.

Chapter 1

The Long-Run Macroeconomic Effects of Aid and Disaggregated Aid in Ethiopia

Journal of International Development, 26(4): 520 – 540, May 2014

Fiseha Haile Gebregziabher
Department of Economics, University of Copenhagen
fiseha.haile.gebregziabher@econ.ku.dk

The Long-Run Macroeconomic Effects of Aid and Disaggregated Aid in Ethiopia*

Fiseha Haile Gebregziabher†

Abstract

This paper investigates the long-run macroeconomic effects of foreign aid and disaggregated aid flows in Ethiopia for the period $1960-2009$. We contribute to the literature by examining whether aid from different sources (multilateral and bilateral aid) and different aid modalities (grant and loan) exert different long-run effects on several macroeconomic variables (Gross Domestic Product (GDP), investment, import, export, and government consumption spending). The results show that aid affects GDP, investment, and imports positively, whereas it is negatively associated with government consumption. Our findings concerning the effects of disaggregated aid flows stand in stark contrast to the findings reported in most previous studies. Bilateral aid increases investment and GDP, and is negatively linked with government consumption, whereas multilateral aid is only positively associated with imports. Grants contribute to GDP, investment, and imports, whereas loans affect none of the variables. Finally, there is evidence to suggest that multilateral aid and loans have been disbursed in a procyclical fashion.

Keywords: Foreign aid; disaggregated aid; macrovariables; cointegrated VAR model; Ethiopia.

JEL classification: C32; F35; O11.

---

*Discussions with Katarina Juselius and Oliver Morrissey are gratefully acknowledged. I also wish to thank participants at the PhD seminar series at the Department of Economics, University of Copenhagen; CREDIT Development seminar at the Department of Economics, University of Nottingham; and the 2013 Nordic Conference in Development Economics (NCDE); and an anonymous referee for helpful comments and suggestions. All errors and omissions are my own.

†Fiseha Haile Gebregziabher is affiliated with the Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, 1353 Copenhagen, Denmark. E-mail: *fiseha.haile.gebregziabher@econ.ku.dk*.

1.1 Introduction

Recent years have witnessed a spate of studies on the contentious issue of aid effectiveness. Extant studies provide statistical findings that are scattered all over the spectrum of possible conclusions, from 'aid works' (Hansen and Tarp, 2001; Dalgaard and Hansen, 2001; Gomanee *et al.*, 2005b; Arndt *et al.*, 2010; Juselius *et al.*, 2014) through 'aid works but only in certain environments' (Burnside and Dollar, 2000; Collier and Dehn, 2001; Dalgaard *et al.*, 2004[1]) to that 'aid is growth-neutral or even pernicious' (Boone, 1996; Easterly *et al.*, 2004; Rajan and Subramanian, 2008).

Burnside and Dollar (2000) (BD) find that aid spurs growth only in countries with 'good' policies and has been extraordinarily influential in policy circles. BD has spawned a new wave of studies, most of which show that its far-reaching policy implications rest on shaky empirical ground. Using the same data as BD, but employing different specifications and estimators, Hansen and Tarp (2001), Dalgaard and Hansen (2001), and Dalgaard *et al.* (2004) show that aid does stimulate growth irrespective of the policy environment on the receiving end. In a similar vein, Easterly *et al.* (2004) re-estimate the BD model using an expanded dataset and refute the BD result. Recently, Rajan and Subramanian (2008) (RS) find that aid has no detectable effect on growth. In contrast, Arndt *et al.* (2010) reexamine RS using a different methodological approach and conclude that aid has a positive effect on growth over the long-run. More recently, Juselius *et al.* (2014) perform a country-by-country analysis of 36 African economies and show that aid contributes to investment and income in the long-run.

This article attempts to advance the debate on aid effectiveness by investigating whether aid from different sources (multilateral and bilateral aid) and different aid modalities (grant and loan) exert different long-run effects on several macroeconomic variables (Gross Domestic Product (GDP), investment, import, export, and government consumption spending) in Ethiopia.

Much of the empirical literature on aid effectiveness is underpinned by cross-country regressions. Notwithstanding that cross-country studies may throw light on general trends, results from such analysis are only statistically valid under a number of fairly stringent assumptions. The implicit assumption of parametric invariance across countries is highly restrictive as it disregards the myriad of heterogenous characteristics of the countries included in the sample. Attempting to prove the effectiveness of aid (or lack thereof) using an aggregate data is tantamount to claiming that aid's benign or malign effects are automatic, detached from the context in which it is given and the conditions attached to its provision (Riddell, 2007). Hence, analytical country-studies provide a more reliable backdrop for addressing the question of whether aid works

[1] Dalgaard *et al.* (2004) show that aid spurs growth outside the tropics but is less potent in them.

(*Ibid.*).

Another critical flaw in most previous studies is their neglect of aid disaggregation. The existing literature focuses on the impact of aggregate aid flows on such aggregate measures of economic performance as GDP growth. However, since a single figure for aid does not capture the heterogeneity of aid, empirical estimates are likely to be plagued with aggregation bias (Clemens *et al.*, 2012). Donor motives and the conditions attached to different kinds of aid may be different and these could render the use of a somewhat artificial aggregate figure implausible. Moreover, the proportions of different types of aid in total aid have seen marked fluctuations over time.

It is against this backdrop that the paper applies the cointegrated VAR model to examine the long-run macroeconomic effects of aid and disaggregated aid flows in Ethiopia for the period $1960-2009$. Ethiopia, harboring one of the most impoverished economies in the world, has to date seen massive influx of official development assistance (ODA) and is currently the largest recipient in the globe (at \$3.8 billion in 2009− about 10 percent of total ODA *net*-disbursements). Thus, the country provides a prime test-case for investigating the impact of foreign aid on recipient economies.

Our approach is similar to that of Juselius, Møller, and Tarp (2014) (JMT14). However, this paper differs from JMT14 in several respects. To begin with, JMT14 paint a broad picture of the impact of foreign aid in 36 African countries and do not study the dynamics of the transmission mechanisms through which aid impacts on the macroeconomy. Moreover, the analysis in JMT14 is confined to closed economy effects, thereby relying on the *ceteris paribus* assumption that 'open economy effects' remain constant. We relax this assumption by including variables proxying for openness to international trade. Further, although JMT14 control for extraordinary events using dummies, they do not model shifts in the long-run trends underlying the macrovariables. Finally, and most importantly, this paper, unlike JMT14, assesses the long-run effects of disaggregated aid flows.

The rest of the paper is organized as follows: Section 1.2 provides a brief theoretical overview of the aid-growth nexus. Section 1.3 discusses aid and macroeconomic trends in Ethiopia. Section 1.4 presents the data, while Section 1.5 is devoted to model specification. Section 1.6 discusses the identified structures of long-run relations. Section 1.7 discusses the long-run impact of shocks to aid and disaggregated aid flows. Finally, Section 1.8 concludes.

1.2 Aid and growth: Transmission mechanisms

The prototype model underlying the nexus between aid and economic growth was the Harrod-Domar model and the two-gap model by Chenery and Strout (1966). The idea behind the two-gap model is that low-income countries lack sufficient domestic resources to finance the investment required to attain a target rate of growth and foreign exchange to import capital goods and technology. Thus, aid can be instrumental in bridging the investment-saving and the foreign exchange gaps. In addition, aid flowing through government channels could potentially relax the fiscal constraints on government spending by supplementing insufficient domestic tax revenue, i.e. the 'fiscal-gap' (Bacha, 1990). However, the gap models, and the Harrod-Domar model at their base, have been severely criticized on several grounds. The back-of-the-envelope calculation of aid requirements provided by the gap models has led to drastic underestimation of aid requirements (Dalgaard and Eriksson, 2009) and fueled overly optimistic expectations about the impact of foreign aid on recipient economies (Easterly, 1999).

Aid is likely to impact on growth through a host of channels. Investment, import, export, and government consumption are the potential transmission mechanisms considered in this paper. Investment in physical capital is often deemed the most important transmission route. Aid may spur economic growth by inducing investment. However, if aid allocated for investment leaks into government consumption, this may reduce its efficacy or even be harmful. Note, however, that some aid is directly intended for consumption and can be beneficial if directed to the social sectors (Gomanee *et al.*, 2005c; Mosley *et al.*, 2004; Mosley and Suleiman, 2007). Low income countries need to import capital goods and technology, but export earnings are often inadequate and volatile. Hence, aid could be instrumental in financing import bills. Aid can also foster exports by, *inter alia*, enhancing infrastructure, 'encouraging' export-oriented policy reforms, and directly supporting the export sector. However, massive influx of aid could cause exchange rate appreciation and thus have detrimental effects on exports by rendering the traded goods sector uncompetitive, i.e. a "Dutch disease" effect generating an anti-export bias.

All these pathways need to be accounted for to make a proper assessment of the scope for aid in fostering (or impairing) growth. The literature identifies many other channels, but the paucity of data circumscribes their use in our multi-equation time-series analysis, which relies on access to long time-series data. Moreover, some of the transmission mechanisms are not measured annually while others change slowly over time and thus do not lend themselves to time-series analysis. Although important variables are omitted from our analysis, this does not, in general, invalidate the long-run estimates. The reason is that cointegration property is invariant to changes in the information set, i.e. a long-run relation detected within a given set of variables will also be found in an enlarged variable set (Johansen, 2000).

As regards the disaggregated measures of aid, the conventional wisdom posits that bilateral aid is dictated by political and strategic considerations, whereas multilateral aid is development-oriented and hence is bound to be more effective. However, the one-size-fits-all blanket policy prescriptions attached to multilateral aid may undermine its efficacy or even be inimical to the economic development of aid receiving countries. In contrast, bilateral aid comes with less policy strings attached to it. Besides, bilateral donors have historical ties with and knowledge about particular recipient countries and superior experience in specific fields of development (Cassen, 1994, *pp.* 245). Concerning grants vis-à-vis loans, the case for grants is that they entail no future repayment burden for recipients, whereas loans could engender debt overhang and stifle growth. However, the repayment obligation associated with loans may render them more effective than grants. Moreover, some scholars purport that grants may discourage domestic resource mobilization and are likely to end up lining the pockets of corrupt leaders.

1.3 Aid and macroeconomic trends in Ethiopia

The past half century has witnessed three remarkably different political regimes in Ethiopia: the Imperial era (*pre*-1974), the socialist (*Derg*) regime (1974 – 1991), and the Ethiopian People's Revolutionary Democratic Front (EPRDF) (*post*-1991). In terms of macro policy stance, these regimes entailed swings from a fairly market-oriented system to a command economy and then back to a market economy. Hence, aid and macroeconomic trends in Ethiopia largely mirror the policies pursued by the three successive political and economic regimes.

In the *pre*-1974 period, an upward trend can be detected in all variables (see Figure 1.1). GDP has been steadily increasing in the 1960s, while aid inflows increased slowly. Investment was also on its rise and appears to have moved in tandem with aid, which is to be expected given that aid is a major source of funds for investment. Of the total aid during this period, 83 percent was from bilateral donors and around 77 percent was given in the form of outright grants (Table A.2 in Appendix A). Exports and imports have also been increasing in this period.

However, these promising trends were nipped in the bud with the coming to power of the socialist regime in 1974. GDP grew slowly and experienced large drops associated with the war with Somalia (1977 – 1978) and the famine in 1984/85. The latter caused aid to escalate dramatically. The *Derg* period witnessed a drastic decline in aid flows (excluding emergency relief) because the country was then considered as being on the 'wrong' camp during the Cold war (Furtado and Smith, 2009).[2] The bulk of the reduction in aid was because of a decline in bilateral aid, whereas multilateral aid continued to rise slowly. About 50 percent of the aid in this period was from multilateral donors, while 75 percent of the aid inflows was composed of grants. Although

[2] A good chunk of the aid inflows in this period was channeled to emergency relief efforts.

investment shows no discernible trend[3], government consumption expenditure surged for most of the *Derg* period. Military expenditure accounted for about half of the government consumption spending in this period (Geda, 2001). Imports and exports declined considerably because of the highly restrictive trade policies put in place by the *Derg* regime. The cumulative effect of the abysmal policy record of the *Derg*, coupled with the political turmoil and the ensuing drying up of external finance, placed the economy on the brink of collapse in the early 1990s.

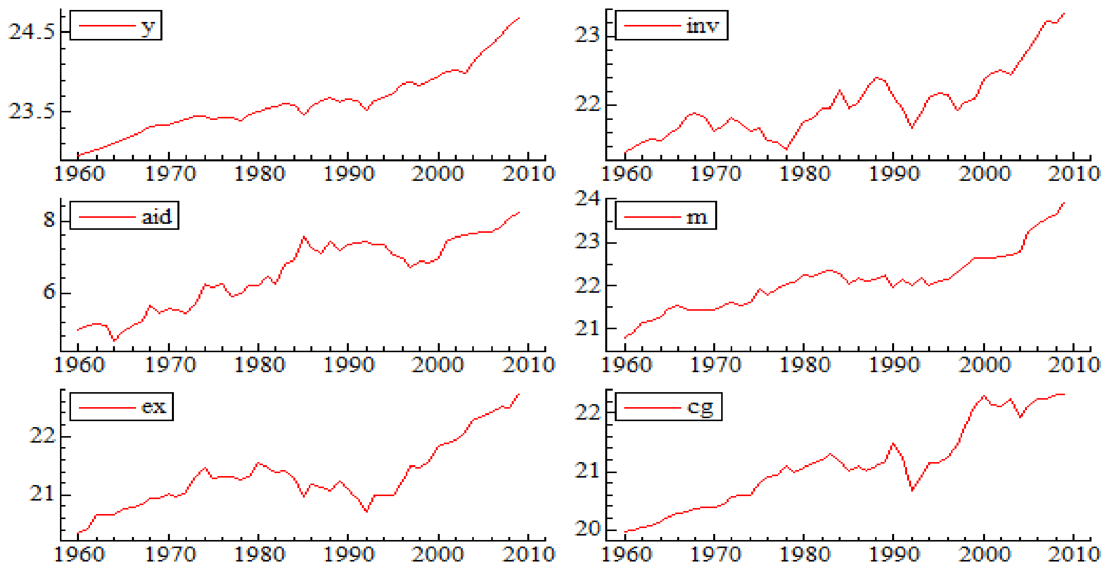

Figure 1.1: The variables of the model in levels

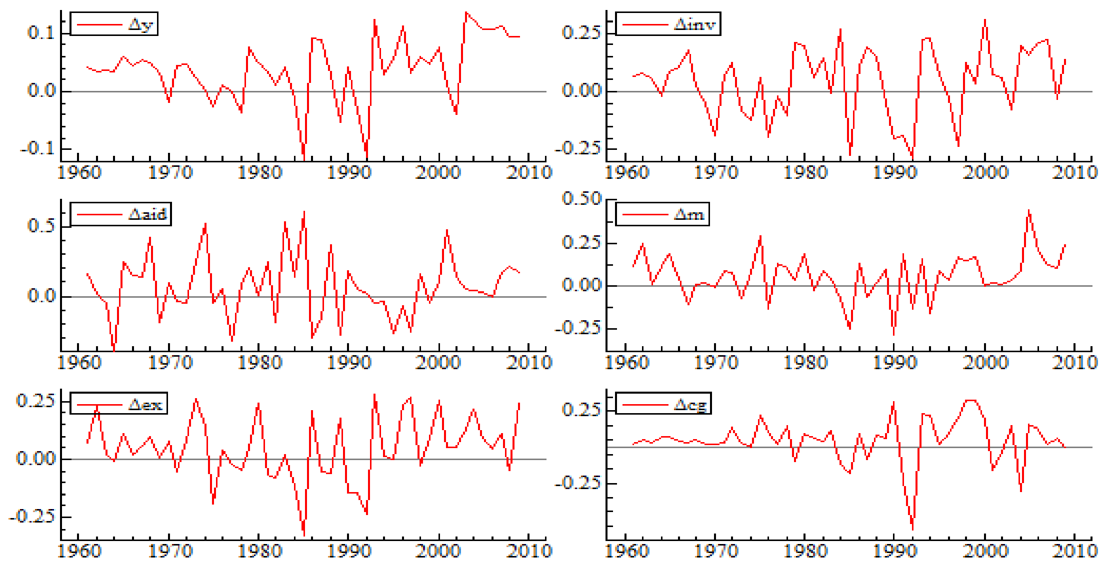

Figure 1.2: The variables of the model in first differences

[3]The rise in investment in the 1980s was largely due to the huge government expenditure on 'agriculture and land settlement' and 'manufacturing', and possibly invigorated by the economic 'Zemecha' (mass mobilization) campaigns (Abegaz, 1999).

The incumbent EPRDF regime assumed power in 1991 and launched IMF/World Bank-sponsored stabilization and structural adjustment programs in 1992 (IMF, 1996). GDP has been rising, notwithstanding some pronounced fluctuations related to, *inter alia*, the war with Eritrea ($1998-2000$) and the drought in 2002/03. The former resulted in a surge in military spending, reflected in the sharp rise in government consumption spending. Investment experienced some pronounced swings around a steadily increasing trend. Investment started to rebound from the downturn in the early 1990s following the adoption of policy reforms and the resurge in external resources. Despite these developments, investment saw a big drop in mid-1990s because of the suspension of lending by the Bretton Woods institutions (BWIs) and in the late 1990s due to the ravages of the war with Eritrea. This period also saw imports and exports rising significantly.[4]

In the past couple of decades, aid flows, particularly the multilateral component, were volatile and fluctuated considerably. The 1990s saw a steady decline in multilateral aid, which was mainly associated with donor policy conditionality (Wade, 2001; Borchgrevink, 2008).[5] Since the early 1990s, the BWIs have been imposing policy conditionalities and cutting back on their aid funding when the government failed to comply (Stiglitz, 2003; Wade, 2001). However, in the past few years, the country has become one of the highest aid recipients in the globe. Of the aggregate aid inflows in the *post*-reform period, 48 percent came from multilateral donors and about 80 percent constituted grants. In sum, a cursory view suggests that aid strongly comoves with investment and GDP.

1.4 Data

The data are annual observations spanning the period $1960-2009$ and comprise the variables: Real Gross Domestic Product (GDP) (y); gross investment (inv) (comprising both private and public outlays); Official Development Assistance (ODA) *net*-disbursements (aid); imports (m); exports (ex); government consumption spending (cg); multilateral aid (ma); bilateral aid (ba); loans (l)[6]; grants (g), and, finally, the inflation rate (Δp). The aid data come from the OECD online database[7] and the inflation rate from the World Bank (2011)'s *World Development Indicators* (WDI), while

[4] Note, however, that the dramatic recovery in exports in the early 1990s was not a reflection of the country's terms-of-trade movements. The steady decline in exports in the *pre*-reform period was caused by the trade-retarding policies instituted by the socialist regime. In addition, the fast growth in exports was due to the return to official channels (rather than exports via the black market) (Dercon, 2000).

[5] Over the period $1992-1997$ aid to Ethiopia was cut by 50 percent, 87 percent of which was accounted for by multilateral donors.

[6] A constant figure was added to loans to be able to take logarithms.

[7] Available at http://stats.oecd.org/Index.aspx.

the rest are from the Penn World Tables (PWT)− version 7.0− database of Heston, Summers, and Aten (2011). The data from the PWT are by and large consistent with data from alternative sources such as the WDI and UN *National Accounts* (NA) database. We opt for the PWT data because it spans a longer period. All variables are at constant market prices and given that all are in logs, their differences represent growth rates. The graphs of all time-series in levels and first differences are shown in Figures 1.1 and 1.2, and Appendix Figure 1.3. The raw ODA data and the percentage shares of the disaggregated aid flows in total ODA are presented in Table A.2 in Appendix A. Table A.1 in Appendix A provides variable definitions.

ODA includes grants and all loans with a grant element of more than 25 percent as well as technical cooperation and assistance, but excludes aid for military purposes. Net ODA may be a reasonable measure of the actual transfer to liquidity-constrained governments (Easterly, 2003).[8] Unfortunately, this conventional measure overstates the amount of aid that is actually spent in the recipient country. Alternatives are the Effective Development Assistance (EDA) indicator[9] (see Chang *et al.*, 1998) and the Ethiopian government's own data. However, these data series are not long and thus do not serve our purpose. Hence, we rely on the OECD data in line with most aid-effectiveness studies. It is also important to note in passing that government consumption spending excludes expenditures associated with the provision of services, primarily education and health, which are recorded under actual household consumption in PWT version 7.0.

Because identifying long-run relations in a system of seven variables is quite demanding, we will pursue a specific-to-general approach in the choice of variables. We first analyze a system containing six variables (y_t, inv_t, aid_t, m_t, ex_t, cg_t) [Model 1]. Model 2 adds the inflation rate to the data vector: (y_t, inv_t, aid_t, m_t, ex_t, cg_t, Δp_t). The advantage of gradually expanding the information set is twofold. First, it greatly facilitates the identification of long-run relations. Second, it enables an analysis of the sensitivity of the results associated with the *ceteris paribus* assumption ingrained in the smaller model. Moreover, extending the variable set with inflation rate allows us to explore how the system of real variables is affected by the inclusion of a nominal variable, which is tantamount to analyzing whether nominal shocks have real effects in

[8] Net ODA is reported net of repayment of past loans and treats forgiveness of past loans as current aid.

[9] The main distinction between ODA and EDA is that the former includes both grants and all concessional loans, whereas the latter is the sum of grants and the grant equivalents of all official loans, and disregards grants that are tied to technical assistance. Although EDA appears starkly different from the conventional aid measure, Dalgaard and Hansen (2001) show that the Pearson correlation between nominal ODA and nominal EDA (both in percent of GDP) is 0.98; between nominal ODA and real EDA it is 0.89. In low-income countries, like Ethiopia, where pure grants account for far greater portion of total aid inflows, the distinction between the two is likely to be immaterial.

the long-run. Model 3 builds upon Model 1 by disaggregating aid into its multilateral and bilateral components: (y_t, inv_t, ma_t, ba_t, m_t, ex_t, cg_t). Model 4 decomposes aid into grants and loans: (y_t, inv_t, g_t, l_t, m_t, ex_t, cg_t).

1.5 Model specification[10]

The analysis in this paper uses the Cointegrated VAR (CVAR) model, which is structured around r cointegration (long-run) relations (the endogenous or *pulling forces*) and $p-r$ stochastic trends (the exogenous or *pushing forces*). The CVAR approach starts from a statistically well-specified VAR model and simplifies the general model by imposing testable and data consistent restrictions. Unlike other approaches in which data are constrained in prespecified directions and are assigned an auxiliary role of 'quantifying' the parameters of an *ad-hoc* theoretical model, the CVAR approach uses strict statistical principles to extract out meaningful relations from the data (Spanos, 2009). Thus, the CVAR can be considered as delineating confidence intervals within which empirically sound models should fall (Hoover *et al.*, 2008).

Our baseline VAR model is specified with two lags and a linear trend restricted to the long-run relations:

$$\Delta x_t = \Gamma_1 \Delta x_{t-1} + \alpha(\beta' x_{t-1} + \beta_1' t) + \mu_0 + \varepsilon_t \tag{1}$$

where x_t is a $p \times 1$ vector of variables defined in the previous section, p is the number of variables; α is a $p \times r$ matrix of adjustment coefficients (denoting the speed of adjustment to equilibrium); $\beta' x_{t-1}$ are r long-run relations; Γ_1 is a $p \times p$ matrix capturing short-run effects; t is a linear trend, β_1 is a r-dimensional vector of trend coefficients of the long-run relations; μ_0 is a vector of constant terms; $\varepsilon_t \sim N.i.i.d(0, \Omega)$ is a $p \times 1$ vector of error terms; and Δ is the first difference operator.[11] If the lag length is 1, then $\Gamma_1 = 0$, which implies that the system, once pushed away from equilibrium by exogenous shocks, adjusts back to equilibrium exclusively through α. This is not, however, the case with higher lag lengths, say $k = 2$, where the adjustment dynamics is also influenced by Γ_1, i.e. the system also adjusts to the lagged short-run changes, Δx_{t-1}. See Johansen (1996) and Juselius (2006) for more technical details and applications of the cointegrated VAR model respectively.

In all models, three observations (1985, 1992, and 2003) were classified as 'too large'. The years 1985 and 2003 coincide with periods of severe drought and famine,

[10] The software packages CATS in RATS (Dennis *et al.*, 2006) and OxMetrics (Doornik and Hendry, 2001) were used to carry out all computations and produce the graphs respectively.

[11] This formulation assumes that x_t is at most integrated of order 1, $I(1)$. To ensure that x_t is not $I(2)$ it is further required that $\alpha_\perp' \Gamma \beta_\perp$ has a full rank, where $\Gamma = I - \sum_{i=1}^{k-1} \Gamma_i$, and $\alpha_\perp$ and $\beta_\perp$ are the $p \times p - r$ orthogonal complements of α and β (Johansen, 1996).

while 1992 marks the advent of IMF/World-sponsored market-oriented reforms.[12] To ensure valid statistical inference we control for these outliers by the impulse dummies $Dp85_t$, $Dp92_t$, and $Dp03_t$.[13] These dummies are defined as $(...0, 0, 0, 1, 0, 0, 0, ...)$ and account for an unanticipated one-period shock effects in $19yy$. The impulse dummies exclusively control for the exceptionally large shocks at the time of occurrence but preserve the information of the observations through their lagged impact.[14] Thus, unlike the case with static regressions, the dummies do not eliminate the corresponding observations.

Of the three outlying observations, the 1992 outlier was found to have significantly caused a mean-shift in the growth rates. The sharp turn from a 'state-led' economy to a 'market-oriented' one appears to have placed some of the macrovariables on a higher growth trajectory (Figures 1.1 and 1.2). Modeling this event boils down to making a choice between allowing for a broken linear trend in the data, x_t, but not in the equilibrium relations, $\beta' x_t$, or also allowing for changing trend slope in $\beta' x_t$.[15] Because the broken trend in x_t may (or may not) imply changes in trend slopes in the equilibrium relations, $\beta' x_t$, the model should be specified to allow for a (testable) broken linear trend in $\beta' x_t$ (see Johansen *et al.*, 2000).

Drawing on *a priori* knowledge of historical facts and the graphical inspection of the data, we control for this event using a piecewise linear trend $(\alpha\beta_1 t + \alpha\beta_{11} t_{92})$ restricted to be in the long-run relations, $\beta' x_t$,[16] and a step dummy $(Ds92_t)$ in the equations, Δx_t. t_{92} is a broken linear trend defined as $(...0, 0, 0, 1, 2, 3, ...)$, which starts in 1992 and accounts for the change in trend slope in the long-run relations. $Ds92_t$ is a step dummy defined as $(...0, 0, 0, 1, 1, 1, ...)$ and controls for the change in growth rates as well as the shift in means of long-run relations. In the preferred model specification, we tested the relevance of the break in the cointegration relationships. The hypothesis

[12] Ethiopia launched an economic reform and structural adjustment program (ERSAP) in 1992 (IMF, 1996). The program encompassed wide-ranging macroeconomic stabilization and structural reform measures: liberalization of foreign exchange and trade systems, exchange rate devaluation, fiscal and monetary reforms, deregulation of prices, private sector development and initial divestiture of state owned enterprises, and sustainable development measures. Although the program was developed in collaboration with the BWIs, the Government steered the pace and sequencing of the liberalization reforms. In fact, the country was one of the very few African countries that pursued 'unorthodox' adjustment path which involved, *inter alia*, active government intervention. Unlike many African countries that were coerced to liberalize at breakneck speed, Ethiopia adopted a conservative stance on liberalization, particularly financial liberalization.

[13] A dummy is used when a residual is larger than $|3.0\sigma_\varepsilon|$.

[14] The dummies account for the unanticipated shocks and given that the latter are no longer unanticipated in the next period, their lagged effects on the system are accounted for by the dynamics of the model.

[15] The VAR model contains both differences and levels of the variables, which makes dummy specification fairly complicated. For instance, a shift in the mean of the equations, $E(\Delta x_t)$, corresponds to a broken linear trend in x_t, whereas a broken trend in $E(\Delta x_t)$ cumulates to a quadratic trend in the data.

[16] Restricting the piecewise linear trend to the long-run relations avoids quadratic trend in the data.

that the broken trend is not needed in the long-run relations has been strongly rejected (p-value: 0.00).

After controlling for extraordinary events, various diagnostic tests indicated that a VAR lag length of 1 satisfactorily and parsimoniously describes the variation in the data (see Section 1.5.1). Incorporating these changes, the final model is given by:

$$\Delta x_t = \alpha(\widetilde{\beta}'\widetilde{x}_{t-1}) + \mu_0 + \Phi_s Ds92_t + \varphi_1 Dp85_t + \varphi_2 Dp92_t + \varphi_3 Dp03_t + \varepsilon_t \qquad (2)$$

where $\widetilde{x}_{t-1} = (x_{t-1}, t, t_{92})$; $\widetilde{\beta}' = (\beta', \beta_1, \beta_{11})$; β_{11} measures the change in the linear trend coefficients of the long-run relations after the advent of policy reforms in 1992; Φ_s is a $p \times 1$ vector of coefficients to the step dummy; and $\varphi_{i,i=1,2,3}$ are vectors of coefficients to the impulse dummies.

On top of the aforementioned events (innovational outliers), economic time series may be affected by additive outliers.[17] Whereas non-modeled innovational outliers may be innocuous (Doornik *et al.*, 1998), additive outliers can distort inference on cointegration rank in small samples and tend to generate a bias towards cointegration or even stationarity (Franses and Haldrup, 1994; Nielsen, 2004). It is not uncommon in applications of the cointegrated VAR model to spot outlying observations and model them with innovational dummies. However, inference based on incorrect configuration of dummies can be seriously misleading (see Nielsen, 2004). A probing look at the plots of the series combined with *a priori* knowledge on the timing of special events pointed to an aberrant observation in the GDP series. GDP increased by about 28 percent in 1987 and dropped by 14 percent in 1988 (see Appendix Figure 1.4).[18] This corresponds to inexplicable measurement error and thus we removed it prior to the econometric analysis.[19] Though an alternative is to account for these outliers using additive dummy variables, this could induce spuriously delayed effects which potentially biases the model estimates (Juselius, 2006, *pp.* 108).

1.5.1 Specification tests

First step in the empirical analysis is to determine the lag length, k, of the VAR model. The choice of an appropriate lag length is based on the Schwarz (SC) and Hannan-Quinn (H-Q) information criteria in tandem with the LM tests for left-over residual autocorrelation. These criteria are only valid under the assumption of a well

[17] An additive outlier is an aberrant realization (i.e. unrelated to the data-generating process) and often occurs due to typing mistakes or gross measurement errors.

[18] This outlier appears to be a transitory shock without any sign of delayed effects in the data and is not observed in the rest of the series.

[19] To substantiate the additive nature of these extreme observations, we thoroughly checked our data against data from alternative sources, such as the *WDI* and the UN *NA* database, and consulted a number of studies (see, among others, Abegaz (1999), Geda (2001) and Geda and Degefe (2005)).

Table 1.1: Determination of the cointegration rank

	τ			$\widehat{\rho}$			$\widehat{\alpha}$		
r	r^*-1	r^*	r^*+1	r^*-1	r^*	r^*+1	r^*-1	r^*	r^*+1
Model 1 ($r^*=4$)	0.06	**0.30**	0.35	0.59	**0.73**	0.79	4.7	**3.8**	2.7
Model 2 ($r^*=5$)	0.14	**0.21**	0.18	0.50	**0.80**	0.82	4.9	**3.2**	2.9
Model 3 ($r^*=5$)	0.10	**0.47**	0.46	0.69	**0.65**	0.78	4.7	**2.9**	2.9
Model 4 ($r^*=5$)	0.19	**0.22**	0.24	0.76	**0.76**	0.82	4.5	**4.1**	2.5

Notes: τ is p-value of the the trace test; $\widehat{\rho}$ is the largest unrestricted characteristic root; $\widehat{\alpha}$ is the largest t-value of the α coefficients; and r^* is the *first-best* choice of rank.
Source: Author's analysis based on the data described in Section 1.4.

Table 1.2: Tests of weak exogeneity (*we*) and pure adjustment (*pa*)

Model 1			Model 2			Model 3			Model 4		
Var.	*we*	*pa*	Var.	*we*	*pa*	Var.	*we*	*pa*	Var.	*we*	*pa*
y	0.00	**0.11**	*y*	0.00	**0.16**	*y*	0.00	**0.16**	*y*	0.00	**0.41**
inv	0.04	0.01	*inv*	0.00	0.02	*inv*	**0.15**	0.01	*inv*	**0.08**	**0.08**
aid	**0.05**	**0.17**	*aid*	0.03	**0.13**	*ma*	0.01	**0.53**	*g*	0.03	**0.23**
m	0.00	**0.75**	*m*	0.00	**0.80**	*ba*	0.01	0.03	*l*	**0.18**	**0.50**
ex	**0.68**	0.01	*ex*	**0.08**	**0.10**	*m*	0.00	**0.37**	*m*	0.00	**0.83**
cg	0.00	**0.80**	*cg*	0.00	**0.20**	*ex*	**0.81**	0.00	*ex*	**0.80**	**0.05**
			Δp	0.00	**0.10**	*cg*	0.00	**0.25**	*cg*	0.00	**0.65**

Notes: reported figures are p-values; significant p-values are given in bold face.
Source: Author's analysis based on the data described in Section 1.4.

specified model; hence, we go back and forth between specification testing and lag length selection. Table 1.5 in Appendix B reports the results.

For Model 1, the SC criteria suggests a lag length of $k = 4$, whereas the H-Q points to $k = 1$. For Models $2 - 4$, the SC points towards $k = 1$, whereas the H-Q suggests three lags. The LM tests indicate that there is no evidence of residual autocorrelation in the VAR(1) for all models. Accordingly, the lag length was truncated to 1. However, the results obtained for $k = 2$ are by and large similar with the ones presented below. Provided that there are no signs of autocorrelation in the residuals, and given the large-dimensional systems and the small sample we have at our disposal, the VAR(1) model is a satisfactory and parsimonious representation of the variation in the data.

Table 1.6 in Appendix B reports the univariate and multivariate misspecification test results. With the deterministic specifications and the dummies included, the four models pass most of the specification tests and describe the data reasonably well. No serious deviation from the underlying assumptions of residual independence and normality has been detected. Although there are some indications of moderate ARCH effects and excess kurtosis, cointegrated VAR results are reasonably robust to such effects (Gonzalo, 1994; Rahbek *et al.*, 2002).

In all four models, the assumption of multivariate normality is by and large satisfied. The univariate test statistics indicate that some big outliers resulted in skewed residuals in Model 2.[20] Nonetheless, given the small number of observations, large dimension of the vector process, and non-rejection of multivariate normality, we continue with this specification. In Model 4, the null of normal errors is not rejected with a very small p-value. However, the univariate test results show that normality is not rejected in all equations. It is possible that the multivariate test result is a finite sample phenomenon. In all models, the multivariate tests for no residual autocorrelation up to order 2 are not rejected at the conventional 5 percent level. Although there is some autocorrelation left in some of the models after having controlled for the largest outliers, our main findings are sufficiently robust to steps that might circumvent the problem, such as increasing the lag length.

However, with only 50 observations, no serious check for parameter constancy can be carried out. Hence, considering the volatile history of Ethiopia, some of the estimated coefficients may represent *average historical effects* over the sample period.

1.5.2 Determination of the cointegration rank

After having established an adequate statistical description of the data, the next step is to determine the cointegration rank. The cointegration rank classifies the data into r long-run relations and $p-r$ common stochastic trends. The choice of rank is of particular interest because it affects the entire model setup and inference procedures in the subsequent stages.

The test for r cointegrating vectors is based on the maximum likelihood test procedure, known as the trace test (see Johansen (1988, 1996)). The trace test is based on a sequence of tests of the null of $p-r$ unit roots for $r = 0, 1, 2, ..., p-1$ and relies on the premise of '*no prior economic knowledge*' about the rank r. Note, however, that some of these statistical null hypotheses may not coincide with plausible economic null hypotheses. This is usually the case for large values of $p-r$ (many stochastic trends) and small values of r (few equilibrium relations), consistent with the presumption in economic theory that macroeconomic variables comove in the long-run. Hence, it is crucial to specify beforehand an economic prior for the number of autonomous shocks, $p-r^*$, where r^* corresponds to the number of long-run relations consistent with this prior. This helps avoid the risk of not rejecting an implausible economic null just because it constitutes a conveniently testable statistical null.

The system variables are likely to be affected by long-run trends associated with cumulative productivity shocks and trends in population, which are proxied by the

[20] The univariate test results for Model 2 are not reported but can be made available upon request.

deterministic trends. Moreover, given that all variables are in real terms, we expect at least two stochastic trends: one emanating from external shocks (such as terms-of-trade fluctuations), captured by shocks to exports, and the other corresponding to persistent medium long-run business cycle movements in the data. However, if aid is found to be exogenous in the system, as it is often assumed, it could constitute a third driving trend. Thus, our preferred choice of rank is $r = 4$ for Model 1 and $r = 5$ for Models $2 - 4$. The statistical evidence shall be assessed against these numbers.

The choice of rank is conventionally made based on the trace test. However, the trace test suffers from substantial power problem when the size of the sample is small (Johansen, 2002). Hence, we base the choice of rank on the largest unrestricted root of the characteristic polynomial for a given r, the t-ratios of the α coefficients for the r^{th} cointegration vector, and the graphs of the r^{th} cointegration relation (see Juselius, 2006: Chapter 8.5).

Table 1.1 reports the p-values of the *trace test* (τ)[21], the largest unrestricted characteristic root $(\widehat{\rho})$, and the largest t-value of the α coefficients $(\widehat{\alpha})$.[22] The test results indicate that $r = 4$ is the statistically most credible (*first-best*) choice of rank for Model 1. For the choice of $r = 5$, the largest unrestricted root seems a bit far from the unit circle; however, the t-values of α indicate that there is no significant adjustment to the last two cointegration vectors. The latter is confirmed by the presence of four mean reverting cointegration vectors.[23] Hence, we consider $r = 4$ to be empirically optimal.

Extending the variable set with inflation rate and disaggregating aid flows imply two possible outcomes for the cointegration rank:

- $r = 4$, the rank is unchanged but the stochastic trends have increased to $p-r = 3$, which would imply that inflation is cointegrated with neither of the variables. In this case, the new common trend would describe the effect of cumulated nominal shocks. This implies the long-run excludability of inflation as it would not add any significant long-run information compared to Model 1. In Models 3 and 4, this case corresponds to an outcome where the components of aid are driven by different stochastic trends and the aid disaggregation introduces one additional stochastic trend into the system.

- $r = 5$ and $p - r = 2$, this corresponds to the case where the number of common stochastic trends is unaltered but inflation is cointegrated with one or more of

[21] These are simulated with a broken linear trend in the cointegration relations.

[22] If all of the t-values of the α coefficients are small, say less than 2.6, then one would not gain a lot by including the $(r+1)^{th}$ cointegration vector in the stationary part of the model (Juselius, 2006, *pp.* 142).

[23] The graphs of the recursively calculated trace tests (not shown here) exhibit pronounced linear growth in the first four cointegration relations, but not in the last two.

the variables in Model 1. In Models 3 and 4, this refers to the possibility where the disaggregated components of aid are cointegrated either between themselves or with the other variables.

In Model 2, the largest unrestricted root is 0.80 for $r = 5$ and 0.82 for $r = 6$. By choosing $r = 5$, we shall include a cointegration relation with a fairly slow, but nevertheless significant adjustment, as evidenced by the largest t value of the α coefficients of 3.2. A glance at the graphs of the cointegration relations reveals that the last two cointegration vectors exhibit distinct non-stationary behavior, pointing toward $r = 5$. Thus, we continue with our preferred and reasonable choice of $r = 5$.

In Model 3, a t-value of 2.9 in α for $\beta_5' x_t$ might suggest that $r = 5$. Although the largest root for $r = 6$ seems relatively far from unity, the strong persistence in $\beta_6' x_t$ might be used as a safeguard against including it in the stationary part of the model. Thus, we settled for $r = 5$. Note, however, that the main conclusions from such analysis remain broadly unchanged with $r = 4$.

In Model 4, the largest unrestricted roots of the companion matrix suggest that $r = 5$ is a potential candidate, which is reinforced by the insignificance of the adjustment coefficients to $\beta_6' x_t$. The graphs of the cointegration vectors also point to $r = 5$. Nonetheless, the cointegration results with $r = 4$ are by and large similar with the ones presented below.

A sensitivity analysis suggested that $r = 3$ is the *second-best* choice for Model 1 and $r = 4$ for Models $2 - 4$. Due to the importance of the choice of rank for the subsequent steps, we perform a sensitivity analysis to check if the empirical estimates based on the statistically most credible value of rank, r^*, are fairly robust to altering the rank to the *second-best* choice, either $r^* - 1$ or $r^* + 1$.

1.5.3 Weak exogeneity and pure adjustment

Although all variables are modelled in the VAR without any prior restrictions, some may be purely adjusting while others weakly exogenous, which are testable hypotheses. To analyze the composition of the pulling and pushing forces in our VAR models, we report the test results of weak exogeneity and pure adjustment of the variables. A variable is said to be weakly (long-run) exogenous if it affects the long-run stochastic path of the variables in the system, while at the same time is not affected by them. Cumulated shocks to a weakly exogenous variable constitute a common stochastic trend. The test of unit-vector in α (pure adjustment) is a test of whether a variable purely adjusts to the long-run relations. If a variable is purely adjusting, then shocks to that variable have no permanent effects on any of the variables in the system.

The test results of weak exogeneity and pure adjustment (endogeneity) of the variables are reported in Table 1.2. We discuss only the results for our preferred choices

Table 1.3: Identified long-run structure (Model 1)

	Eigenvectors $\widehat{\beta}$				*Weights* $\widehat{\alpha}$			
Var.	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$
y_t	1.00	–	–	–	−0.45 (−5.28)	−0.10 (−3.91)	0.11 (3.80)	−0.05 (−1.91)
inv_t	−0.33 (−9.74)	1.00	–	–	−0.65 (−2.07)	−0.27 (−2.97)	0.31 (3.05)	*
aid_t	−0.06 (−2.77)	−0.52 (−8.40)	−0.26 (−4.21)	1.00 (8.48)	*	0.32 (2.13)	*	−0.32 (−2.17)
m_t	–	–	1.00	–	0.58 (2.39)	−0.24 (−3.43)	−0.33 (−4.19)	−0.13 (−1.91)
ex_t	−0.37 (−10.33)	–	−0.76 (−5.89)	–	*	*	*	*
cg_t	–	0.58 (6.04)	–	1.00	*	−0.26 (−3.22)	0.16 (1.82)	−0.33 (−4.24)
t_{92}	–	−0.13 (−11.83)	−0.06 (−3.97)	–				
t	–	–	–	−0.13 (−12.86)				

Notes: t-values in parentheses; *Denotes insignificant coefficients.
Source: Author's analysis based on the data described in Section 1.4.

of rank.[24] Export is exogenous in all models, which reflects the fact that the long-run path of exports of Ethiopia has been driven mainly by external events such as changes in world market prices. This implies that cumulated shocks to exports will act as an exogenous driving trend in all systems. The hypothesis that loan is weakly exogenous in Model 4 cannot be rejected, albeit with a small p-value. GDP, aid (multilateral aid in Model 3 and, both grant and loan in Model 4), import and government consumption are purely adjusting in all models. In addition, inflation is purely adjusting in Model 2. The results for Model 2 show that the exogeneity/endogeneity status of the variables in Model 1 is robust to the *ceteris paribus* clause. The fact that multilateral aid is strongly adjusting in Model 3 implies that shocks to that variable will have no permanent effect on any of the variables in the system.

1.6 Identified long-run relations

This section discusses the identified structures of four cointegration (long-run) relations for Model 1 and five long-run relations for Models $2-4$. To economize on space the discussion for Models $2-4$ will be brief. When interpreting the results in this section it is important to bear in mind that a cointegration relation only measures the association between the variables over the long-run and as such does not say anything about causality. To say something about causality we need to combine the cointegration coefficients, β, with the adjustment coefficients, α. For example, the hypothetical

[24] Note that these test results are not invariant to changes in the cointegration rank. Thus, the results should be interpreted keeping this caveat in mind.

cointegration relation $(x_{1,t} - \beta_1 x_{2,t}) \sim I(0)$[25] describes a positive comovement between $x_{1,t}$ and $x_{2,t}$. If the adjustment coefficient α_1, of $\Delta x_{1,t}$, is negative and significant but the adjustment coefficient corresponding to $x_{2,t}$ is insignificant, i.e. $\alpha_2 = 0$, we can say that the direction of causality runs from $x_{2,t}$ to $x_{1,t}$, i.e. $x_{1,t} = \beta_1 x_{2,t} + u_t$. However, the interpretation becomes less straightforward in terms of sign effects as the number of variables in a long-run relation increases.

Table 1.3 and Tables 1.7, 1.8, and 1.9 in Appendix B report the identified long-run structures for Models 1, 2, 3, and 4 respectively. The estimated structures are generically and empirically identified as defined in Johansen and Juselius (1994). The graphs of the cointegration relations for Models 1, 2, 3, and 4 are shown in Appendix Figures 1.5, 1.6, 1.7, and 1.8 respectively.

The identified long-run structure for Model 1 is accepted based on a p-value of 0.48. The first long-run relation is between GDP, aid, investment, and export. Only GDP is equilibrium correcting to the relation, signifying its importance as a GDP relationship:

$$y_t = 0.06aid_t + 0.33inv_t + 0.37ex_t$$

The results suggest that *ceteris paribus* exports, investment, and aid all make positive contribution to long-run movements in GDP. The coefficients indicate that the long-run fluctuations in income have mostly been associated with exports and investment. This result, coupled with our finding in the next section that the long-run impact of a shock to exports on income is positive and highly significant, seems to lend credence to the export-led development strategy the country has been pursuing in the last couple of decades. The fact that the country experienced an unprecedented spell of economic growth during this period seems to reinforce this claim. The significant contribution from investment is consistent with Geda (2005), who shows that physical capital explains a good portion of Ethiopia's long-run GDP growth.

The finding that income is positively associated with aid is consistent with previous studies (see, among others, Hansen and Tarp, 2001; Dalgaard and Hansen, 2001; Dalgaard *et al.*, 2004; Gomanee *et al.*, 2005b; Arndt *et al.*, 2010; and Juselius *et al.*, 2014). The coefficient to aid is small suggesting that its contribution to income is relatively small. However, it is important to note that the portion of aid that can be expected to directly affect income, once we accounted for the main transmission channel through which aid operates, is relatively small. This is partly because some aid flows are not intended to enhance productive capacity. For instance, humanitarian aid accounts for about 20 percent of the total aid flows to Ethiopia in the past decade.

It is worth noting that the inclusion of both aid and investment in the GDP rela-

[25] A series is said to be integrated of order 0, $I(0)$, if it is stationary without a trend.

tionship might imply the 'double counting' of aid (as some aid is used for investment). However, collinearity between the variables does not result in imprecise estimates of the cointegration relation (Juselius, 2006; Juselius *et al.*, 2013). The reason for this is that, unlike the case with a regression analysis in levels, the cointegrated VAR formulation more or less circumvents the multicollinearity problem by transforming trending variables, x_t, into stationary differences, Δx_t, and stationary cointegration relations, $\beta' x_t$ (*Ibid.*). However, it should be kept in mind when interpreting the results that a portion of investment has been financed by aid inflows.

The α coefficient to GDP reveals that any deviation from the long-run relation is corrected back to equilibrium on average within two years. The relatively modest speed of adjustment seems to reflect the structural rigidity of the economy, typical of developing countries. The adjustment coefficients also show that imports increase when GDP rises above its steady-state level. Moreover, investment reacts to deviations from the relation, but not in an equilibrium correcting manner.

The second long-run relation describes a strong association between investment, aid and government consumption, and normalization is made on investment. All three variables equilibrium correct to the relation, implying the simultaneous feedback between the variables. Hence, one could also interpret this relation as an aid or government consumption relation:

$$inv_t = 0.52aid_t - 0.58cg_t + 0.13t_{92}$$

The positive coefficient to aid is consistent with our finding in the next section that the final impact of a shock to aid on investment is positive and significant. This is to be expected given that foreign aid inflows have been the major source of financing for capital expenditure in Ethiopia (Geda, 2001; Alemu, 2010). For instance, the share of external assistance and loans in financing capital expenditure stood at 53 percent in mid-1980s, dropped to 24 percent in mid-1990s, and increased to 48 percent in the early 2000s (Geda, 2001). The strong positive relationship between aid and investment, combined with the first long-run relation, seems to suggest that investment is the most important transmission mechanism through which aid impacts on income in the long-run.

The negative coefficient to government consumption suggests a substitution effect between investment and government consumption spending, which is somewhat puzzling. This possibly reflects the fact that a good chunk of the government recurrent expenditure over the period under consideration was unproductive and diverted resources away from investment projects. For instance, defense budget accounted for nearly half the total recurrent expenditure during the *Derg* period (Geda, 2001). Military expenditure also skyrocketed in the 1990s, which culminated in the war with Eritrea $(1998 - 2000)$. Such unproductive expenditures account for a significant portion of government consumption spending, particularly considering that the latter does not include education and health expenditures. The broken linear trend, t_{92}, captures the positive shift in trend slope in investment (alternatively a positive shift in government consumption or a leveling-off in aid).

The third relation comprises import, aid, and export, and the adjustment coefficients show that it is only import that is error correcting to this relation:

$$m_t = 0.26aid_t + 0.76ex_t + 0.06t_{92}$$

We find that aid permits higher level of imports, which is line with the theoretical expectation and suggests that aid helps stabilize the supply of foreign exchange needed for imports. This result basically reflects the fact that foreign aid inflows have been increasingly important source of foreign exchange for financing imports in Ethiopia since the 1960s. Moreover, the time path of foreign exchange earnings from exports has been a crucial determinant of imports. Export possesses a large coefficient, suggesting that the country's import bills were mainly financed by the "free" foreign exchange available from export earnings. This is consistent with our expectations as export is by far the largest foreign exchange earner in Ethiopia, which accounts on average for more than 85 percent of total foreign exchange earnings. The coefficient on the broken linear trend, t_{92}, shows the positive change in trend slope in import. The α coefficients show that GDP and investment increase when import is above its steady-state level.

The fourth long-run relation is a function of aid and government consumption, and normalization is made on the latter. However, our normalization is somewhat *ad hoc* because both variables significantly equilibrium correct to the relation. Thus, the relation could equally well be interpreted to be consistent with a relation for aid being negatively affected by government consumption:

$$cg_t = -aid_t + 0.13t$$

The relation essentially describes that government consumption tends to fall when aid increases and vice versa. This perhaps suggests that the curtailment of the public sector was a *quid pro quo* for an increase in aid. It is also possible that donors give aid on the condition that the country contributes certain resources to development projects for which aid is earmarked or that aid goes into projects in which there is a critical minimum level of investment, which 'compels' Ethiopia to pour resources into the projects once that level is met. Considering that a significant portion of the government recurrent expenditure in the past was more or less unproductive and that government consumption expenditure excludes spending associated with health and education services, the fourth cointegration relation might suggest that aid flows have had positive contribution to development projects through public resource reallocation. This finding is essentially at odds with the claim by fervent aid critics that most of foreign aid finances government consumption rather than investment.

If normalization is made on aid, the relation would reflect the *ex post* reaction of donors that they cutback aid when government recurrent spending surges.

Model 2 adds the inflation rate to the system. The identified structure in Table 1.7 (Appendix B) is accepted based on a p-value of 0.49. The first cointegration relation describes a similar GDP relationship as in Model 1. $\beta_2' x_t, \beta_3' x_t$, and $\beta_4' x_t$ are also maintained in the extended system. The fifth relation contains GDP, investment, and inflation, of which the last two are equilibrium correcting. The estimates show that inflation comoves with income, whereas it is negatively associated with investment. The results may suggest that demand pressures stemming from an increase in income lift prices up. The concurrence of high inflation rate with high economic growth is a typical feature of the last decade in Ethiopia as compared to the period before when inflation was largely associated with supply shocks, such as conflicts or droughts. The general trend until the turn of the 21st century was that inflation picks during periods of crop failures (or political turmoil) and declines during bumper years. Put another way, demand shocks predominated in the last decade, whereas supply shocks in the period before. Thus, a possible interpretation is that the recently rising demand pressures have dominated the past influences of supply shocks.

The negative association between investment and inflation is plausible as an increase in investment spending, while perhaps inflationary in the short-run, contributes to price stability over the long-run via the removal of structural bottlenecks and the ensuing increase in supply of goods and services. However, investment is also error-correcting and the relation might as well suggest that higher inflation constricts investment.

Model 3 decomposes aid into its multilateral and bilateral components. The identified structure in Table 1.8 (Appendix B) is accepted based on a quite high p-value of 0.94. The first long-run relation is between GDP, bilateral aid, multilateral aid, investment, and export. The second cointegration relation in Model 1 remains valid in this system, except that only the bilateral component of aid enters the relation with a positive coefficient. From the third long-run relation we see that multilateral aid is positively associated with imports. The estimates of the fourth long-run relation show that bilateral aid has a strong negative association with government consumption. This may suggest that bilateral aid is closely monitored by donors and that such aid goes into development projects and not into government consumption spending. This stands in stark contrast to the widely held assertion that there is considerable leakage out of bilateral aid to government consumption.

The fifth cointegration relation shows that multilateral aid is positively associated with exports. This suggests that multilateral donors have been making aid disbursements in a procyclical fashion: aid tends to be disbursed in periods when exports are high and declines in the wake of negative export shocks. Put in other terms, multilateral donors have not been cushioning the adverse impact of export shocks by acting as a buffer, i.e. engineering aid flows toward stabilizing the fluctuations in foreign currency inflows. Needless to say, this weakens the effect of aid because procyclical aid reinforces macroeconomic instability, thereby detracting from other beneficial effects of multilateral aid. A possible explanation is that the procyclicality of aid is linked to conditionality by the BWIs. This could be the case when donors have limited information regarding the country's reform effort and thus base aid allotment on weak observable macroeconomic indicators. It is also possible, albeit less reasonable, that aid flows are intertwined with the business cycles of donors which in turn might comove with that of recipients.

Model 4 breaks down aid into grants and loans. The identified structure in Table 1.9 (Appendix B) is accepted based on a fairly high p-value of 0.72. The first long-run relation shows that the positive relationship between aid and GDP we found in Model 1 is due to grants. The second long-run relation indicates that only the grant component of aid is positively related to investment. The third relation reveals that grants contribute to imports, whereas this is not the case for loans. The fourth cointegration relation shows that grants are negatively related to government consumption. This possibly reflects that the lack of repayment obligation associated with grants offers donors greater political scope to impose conditions, such as the curtailment of government recurrent expenditure. Finally, the fifth long-run relation unveils that loan is positively associated with exports. This suggests that loan disbursements have been procyclical, which is in concordance with the finding for multilateral aid.

In sum, the results for grants and loans are not surprising given the findings for bilateral and multilateral aid because most bilateral assistance has been in the form of outright grants and that almost all loans (close to 90 percent in recent years) were issued by multilateral donors.

1.7 The long-run impact of shocks to aid flows

This section discusses the long-run effects of shocks to aid and disaggregated aid flows using the moving average (MA) representation of the cointegrated VAR model.

1.7.1 The MA representation of the CVAR model

The common trends representation of the cointegrated VAR model is given by:

$$x_t = C \sum_{i=1}^{t} (\varepsilon_i + \Phi D_i) + C^*(L)(\varepsilon_t + \Phi D_t) + z_0 \tag{3}$$

where $C = \beta_\perp (\alpha'_\perp (I - \Gamma_1)\beta_\perp)^{-1} \alpha'_\perp$ is a matrix of rank $p - r$; $C^*(L)$ is a stationary lag polynomial; and z_0 depends on the initial values of the process (Johansen, 1996). The $p - r$ common stochastic trends are measured by $\alpha'_\perp \sum_{i=1}^{t} \varepsilon_i$. The impulse responses of a shock to a variable can be calculated from the lag polynomial $C^*(L)\varepsilon_t$. The final impact matrix C is of particular interest because the overriding objective is to analyze the final effects of permanent shocks to the variables on the system. It is important to note that the C matrix gives the end points of the impulse response functions. It is evident that identifying restrictions on the long-run relations, $\beta' x_t$, can possibly be done in multiple ways and thus several long-run structures can reproduce the reduced form Π $(= \alpha\beta')$ matrix. For this reason, we calculate the impulse response functions, for a unitary change of $\widehat{\varepsilon}_{it}$, without imposing any restriction on α and β. Put another way, our discussion concentrates on the unrestricted impulse response matrix because it is estimated uniquely.

To facilitate understanding of the individual elements of the C matrix, we rewrite equation (3) for the variables in Model 1:

$$\begin{bmatrix} y_t \\ inv_t \\ aid_t \\ m_t \\ ex_t \\ cg_t \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{41} & c_{42} & c_{43} & c_{44} & c_{45} & c_{46} \\ c_{51} & c_{52} & c_{53} & c_{54} & c_{55} & c_{56} \\ c_{61} & c_{62} & c_{63} & c_{64} & c_{65} & c_{66} \end{bmatrix} \begin{bmatrix} \sum_{i=1}^{t} \varepsilon_{1i} \\ \sum_{i=1}^{t} \varepsilon_{2i} \\ \sum_{i=1}^{t} \varepsilon_{3i} \\ \sum_{i=1}^{t} \varepsilon_{4i} \\ \sum_{i=1}^{t} \varepsilon_{5i} \\ \sum_{i=1}^{t} \varepsilon_{6i} \end{bmatrix} + C \sum_{i=1}^{t} \Phi D_i + C^*(L)(\varepsilon_t + \Phi D_t) + z_0$$

The elements of a column in the C matrix describe the long-run impacts of a shock to variable x_{ij} on the system, whereas a row is interpreted as the cumulated effect on the long-run movements in x_{ij} of permanent shocks to the variables in the system. Of particular interest here is the long-run effect of a shock to aid on the macrovariables (described by the column c_{i3}) and the effect of the macrovariables on aid (captured by the row c_{3i}). If $c_{13} = c_{23} = c_{43} = c_{53} = c_{63} = 0$, then a shock to aid has no long-run effect on any of the macrovariables. If $c_{31} = c_{32} = c_{34} = c_{35} = c_{36} = 0$ and $c_{33} \neq 0$, then aid is exogenous. The remaining coefficients in the C matrix describe the long-run impulse-responses between the macrovariables alone. Table 1.4, and Tables 1.10, 1.11, and 1.12 in Appendix B report the long-run impact matrix C for Models

Table 1.4: The common trends and impact matrix C (Model 1)

The common stochastic trends $\alpha'_\perp \sum \varepsilon_i$							
	y_t	inv_t	aid_t	m_t	ex_t	cg_t	
$\alpha'_{\perp,1}$	0.00	0.09 (0.26)	0.23 (1.22)	−0.03 (−0.12)	**1.00**	−0.16 (−0.73)	
$\alpha'_{\perp,2}$	**1.00**	**−0.53** (**−2.58**)	−0.15 (−1.36)	−0.11 (−0.79)	0.00	0.05 (0.37)	
The long-run impact matrix C							
	$\widehat{\varepsilon}_y$	$\widehat{\varepsilon}_{inv}$	$\widehat{\varepsilon}_{aid}$	$\widehat{\varepsilon}_m$	$\widehat{\varepsilon}_{ex}$	$\widehat{\varepsilon}_{cg}$	$\widehat{\varepsilon}^*_{aid}$
y	−0.24 (−0.90)	**0.15** (**2.00**)	**0.09** (**2.53**)	0.02 (0.29)	**0.23** (**3.66**)	−0.05 (−0.86)	**0.12** (**3.03**)
inv	−1.80 (−1.76)	**0.97** (**3.41**)	**0.31** (**2.29**)	0.20 (0.75)	0.15 (0.63)	−0.11 (−0.51)	**0.33** (**2.59**)
aid	−1.75 (−1.44)	**0.99** (**2.93**)	**0.42** (**2.59**)	0.18 (0.57)	**0.65** (**2.29**)	−0.19 (−0.74)	**0.50** (**3.63**)
m	0.52 (1.31)	**−0.24** (**−2.23**)	0.00 (0.00)	−0.07 (−0.67)	**0.34** (**3.72**)	−0.03 (−0.36)	0.07 (0.62)
ex	0.84 (1.01)	−0.37 (−1.62)	0.06 (0.55)	−0.12 (−0.56)	**0.81** (**4.22**)	−0.09 (−0.52)	0.24 (0.45)
cg	1.55 (1.71)	**−0.84** (**−3.35**)	**−0.29** (**−2.38**)	−0.17 (−0.72)	−0.21 (−0.99)	0.11 (0.57)	0.37 (0.82)

Notes: Results are based on the *second-best* choice of rank; t-values in parentheses, with $t \geq 1.80$ given in bold face; $\widehat{\varepsilon}^*_{aid}$ is based on the *second-best* choice of cointegration rank.

Source: Author's analysis based on data described in Section 1.4.

1, 2, 3, and 4 respectively. In the last columns of the C matrices we report the effects of shocks to aid (and its disaggregated components in Models 3 and 4) based on the *second-best* choices of cointegration rank. To facilitate readability we have indicated significant coefficients with bold face.

The common stochastic trends, $\alpha'_\perp \sum \varepsilon_i$, for Model 1 are reported in the upper part of Table 1.4. Consistent with the test results of weak exogeneity, the first common stochastic trend, $\alpha'_{\perp,1} \sum \varepsilon_i$, is measured by cumulated shocks to exports. This is a reflection of the fact that over the long-run exports of Ethiopia have been driven mainly by external factors such as fluctuations in terms-of-trade. The second stochastic trend, $\alpha'_{\perp,2} \sum \varepsilon_i$, is measured by a weighted average of the empirical shocks to real income and investment, and seems to capture medium long-run movements in the real economy. These two common trends are invariant to expansion of the variable set with inflation (Model 2) and the disaggregation of aid flows (Models 3 and 4).[26] This is consistent with the asymptotic invariance of common trends to extensions of the information set demonstrated in Johansen and Juselius (2014).

[26] The common stochastic trends for Models $2-4$ are not reported but can be made available upon request from the author.

1.7.2 The final impact of shocks to aid flows

In Model 1, a shock to aid has significant effect on GDP (positive), investment (positive), and government consumption (negative). The positive effect of aid on income is in contrast to Juselius *et al.* (2014), who found a significant negative effect. However, we believe our result is plausible for the reasons discussed in Section 1.1. The effects of aid on investment and government consumption are consistent with Juselius *et al.* (2014). The positive effect of exports on aid seems at variance with the conventional wisdom that donors inject more aid to countries in which export has declined due to, among others, natural disasters and terms-of-trade shocks. In other words, foreign aid has had no countervailing relationship with fluctuations in export earnings. Shocks to investment affect aid positively, which reflects that higher level of investment attracts more aid perhaps because donors believe that the country uses aid effectively.

These results are maintained in Model 2. In addition, shocks to aid have insignificant effect on inflation and exports. Hence, there is no evidence to suggest that positive shocks to aid inflows induce inflationary pressures and curtail exports by causing exchange rate appreciation.

In Model 3, a shock to bilateral aid has significant effect on GDP (positive), investment (positive), and government consumption (negative), whereas a shock to multilateral aid affects none of the variables. The latter is not surprising given our previous finding that multilateral aid is strongly adjusting (endogenous). Decomposing aid into its bilateral and multilateral components uncovers the interesting result that a shock to exports has positive and significant effect on multilateral aid, whereas it has no significant impact on bilateral aid. This suggests that multilateral aid disbursements have been made in a procyclical manner. Put differently, aid from multilateral donors was not tailored to help the country mitigate fluctuations in export earnings, rather it has been making good years even better and falling when it is needed the most, further exacerbating economic doldrums. Moreover, the final impact of a shock to government consumption on bilateral aid is negative, revealing that a surge in government recurrent spending triggers a fall in aid from bilateral donors. Further, shocks to investment exert positive effect on bilateral aid. This indicates that bilateral donors respond to surges in investment by providing more aid possibly because this reflects that the country uses aid for intended purposes.

The long-run impact matrix for Model 4 shows that a shock to grants has significant effect on GDP (positive), investment (positive), and government consumption (negative), whereas a shock to loans has no effect on the system variables. A shock to investment affects grants positively but has no effect on loans. The final impact of a shock to loan is similar to that of multilateral aid, which comes as no big surprise considering that most of loans to Ethiopia have been issued by multilateral donors. In addition, bilateral aid and grants have similar effects on the macrovariables, reflecting the fact that most bilateral aid has been composed of outright grants.

Finally, the last columns of the C matrices show that the long-run effects of aid and disaggregated aid flows are fairly robust to changes in the cointegration rank.

1.8 Conclusions

In this paper, we investigated the long-run macroeconomic effects of aid and disaggregated aid flows in Ethiopia for the period $1960 - 2009$, using the cointegrated VAR model as a statistical benchmark. We found that foreign aid contributes to income, investment, and imports, whereas it is negatively associated with government consumption spending. The latter stands in stark contrast to the claim often put forward by staunch aid critics that most of aid leaks to government consumption.

Disaggregating aid into its multilateral and bilateral components unveils that the latter increases investment and income, and is negatively correlated with government consumption, whereas multilateral aid is only positively associated with imports. The negative effect on government consumption spending is in contradiction with the rhetoric that there is considerable leakage out of bilateral aid to government consumption. With regard to grants vis-à-vis loans, we found that grants contribute to investment, import and income, and are negatively related to government consumption, whereas loans affect none of the variables. Hence, if the purpose of aid is to increase imports, investment, and income, it may be better to provide outright grants than loans.

Further, we found evidence to suggest that multilateral aid and loans have been disbursed in a procyclical fashion. This sounds awkward as such pattern of aid flows reinforces macroeconomic instability. As Ethiopia has a less-diversified economy and restricted access to international credit markets, it would need external finance in times of economic downturn to recover back to its long-term trend. The results suggest that multilateral donors should take steps to change the hitherto pattern of aid giving into a smoother path of aid disbursements.

Finally, we acknowledge upfront that we only took the initial step toward understanding the differential long-run effects of different forms of aid in Ethiopia. Hence, further work is needed to pin down the mechanisms that could explain why the different types of aid operate differently. In other words, further disaggregation of aid flows would be needed to have a clearer picture of what kind of aid works, what does not, and why.

Appendices

Appendix A

Table A.1. Data description

Var.	Description
Y	Real Gross Domestic Product (GDP).
Inv	Investment. Includes outlays on additions to the fixed assets of the economy, net changes in the level of inventories, and net acquisitions of valuables.
Aid	Net official development assistance (ODA).
M	Import. The value of all goods and services received from the rest of the world.
Ex	Export. The value of all goods and services provided to the rest of the world.
Cg	Government consumption spending. It includes all government current expenditures for purchases of goods and services, compensation of employees, and most expenditures on national defense and security.
Δp	The annual change in the log of consumer price index.
Ba	Bilateral aid. Aid flows from all bilateral donors.
Ma	Multilateral aid. Aid flows from all multilateral donors.
L	Net loans. ODA loans with maturities of over one year. The net data are reported after deduction of amortization receipts and the effect of other measures reducing aid (e.g. forgiveness).
G	Grants. *Aid* minus *net loans.* Debt cancellations do not constitute real resource transfers and thus, the forgiveness of past loans associated with the Highly Ind ebted Poor Countries (HIPC) initiative in 2006 (amounting to $3,616 million) were extracted from the data prior to the analysis.

Table A.2. Raw ODA data and disaggregated aid (% of ODA)*

Year	ODA	Ma	Loan	Year	ODA	Ma	Loan	Year	ODA	Ma	Loan
1960	142.7	3.5	0.0	1977	363.1	50.0	35.6	1994	1521.9	46.8	27.7
1961	167.6	3.6	-0.3	1978	394.1	61.0	33.9	1995	1159.8	38.4	29.4
1962	170.7	5.8	3.5	1979	486.2	56.2	42.1	1996	1074.5	44.8	27.1
1963	161.6	10.3	7.9	1980	491.0	59.8	20.3	1997	826.7	34.0	19.9
1964	108.3	22.9	13.9	1981	630.5	62.6	17.7	1998	972.6	42.5	16.6
1965	138.3	23.3	17.7	1982	519.4	62.2	20.5	1999	929.7	47.2	27.1
1966	161.2	13.4	19.5	1983	892.1	53.0	35.3	2000	1037.3	43.1	19.9
1967	185.9	11.0	21.1	1984	1020.2	48.1	22.2	2001	1685.9	64.7	47.2
1968	284.2	21.7	46.7	1985	1892.0	42.5	13.3	2002	1938.5	61.5	44.8
1969	233.6	23.4	31.7	1986	1395.2	37.1	10.7	2003	2038.7	35.8	10.1
1970	256.7	16.6	39.5	1987	1192.4	49.4	21.1	2004	2110.3	42.6	9.2
1971	245.8	21.8	46.6	1988	1721.3	41.4	20.1	2005	2176.1	37.7	12.9
1972	233.6	21.1	37.3	1989	1291.2	47.3	18.3	2006	2174.0	47.9	23.3
1973	299.7	26.4	34.6	1990	1553.7	43.4	16.1	2007	2589.3	51.8	16.0
1974	505.3	32.8	30.4	1991	1640.7	58.0	10.0	2008	3229.7	43.8	8.8
1975	480.0	45.3	47.1	1992	1662.7	61.3	9.2	2009	3820.0	51.9	24.3
1976	504.8	51.8	42.0	1993	1575.9	61.8	32.9				

*ODA in USD millions. Bilateral aid and grant constitute the remaining portions.
Source: http://stats.oecd.org/Index.aspx.

Appendix B

Table 1.5: Lag length determination

	Model 1				Model 3			
	SC	H-Q	LM(1)	LM(k)	SC	H-Q	LM(1)	LM(k)
VAR[1]	**−23.90**	−25.84	**0.22**	**0.22**	**−25.20**	−27.61	**0.41**	**0.41**
VAR[2]	−22.59	−25.43	0.03	0.18	−23.63	−27.24	0.11	0.11
VAR[3]	−22.26	−25.99	0.59	0.33	−23.07	**−27.88**	0.05	0.68
VAR[4]	−23.68	**−28.30**	0.39	0.33				
	Model 2				Model 4			
	SC	H-Q	LM(1)	LM(k)	SC	H-Q	LM(1)	LM(k)
VAR[1]	**−27.29**	−29.74	**0.10**	**0.10**	**−24.69**	−27.09	**0.17**	**0.17**
VAR[2]	−25.82	−29.60	0.01	0.11	−23.46	−27.07	0.00	0.11
VAR[3]	−27.26	**−32.36**	0.01	0.03	−23.80	**−28.62**	0.48	0.55

Source: Author's analysis based on the data described in Section 1.4.

Table 1.6: Misspecification tests

Panel A. Univariate tests

	Model 1		Model 3		Model 4	
Var.	Normality	ARCH	Normality	ARCH	Normality	ARCH
Δy_t	2.36 [0.31]	1.04 [0.31]	1.71 [0.43]	1.81 [0.28]	3.50 [0.17]	1.32 [0.25]
Δinv_t	2.97 [0.23]	0.38 [0.54]	3.64 [0.16]	1.46 [0.23]	2.59 [0.27]	0.46 [0.50]
Δaid_t	1.15 [0.56]	1.05 [0.31]	–	–	–	–
Δm_t	1.73 [0.42]	0.26 [0.61]	2.20 [0.33]	0.31 [0.58]	1.46 [0.48]	0.003 [0.95]
Δex_t	1.96 [0.38]	0.00 [0.99]	2.46 [0.29]	0.002 [0.96]	3.61 [0.17]	0.06 [0.80]
Δcg_t	5.10 [0.08]	0.67 [0.41]	4.42 [0.11]	1.31 [0.25]	5.09 [0.08]	0.95 [0.33]
Δba_t			2.21 [0.33]	0.13 [0.72]	–	–
Δma_t			1.02 [0.60]	0.02 [0.89]	–	–
Δg_t					2.97 [0.23]	1.24 [0.27]
Δl_t					1.48 [0.49]	0.03 [0.85]

Panel B. Multivariate tests

Test	LM(k)	Model 1	Model 3
Autocorrelation	LM(1)	$\chi^2(36) = 47.95$ [0.09]	$\chi^2(49) = 50.08$ [0.43]
	LM(2)	$\chi^2(36) = 39.34$ [0.32]	$\chi^2(49) = 64.78$ [0.07]
ARCH	LM(1)	$\chi^2(441) = 493.24$ [0.04]	$\chi^2(784) = 815.24$ [0.21]
	LM(2)	$\chi^2(882) = 938.08$ [0.09]	$\chi^2(1568) = 1372$ [1.00]
Normality		$\chi^2(12) = 16.07$ [0.19]	$\chi^2(14) = 18.09$ [0.20]
Test	LM(k)	Model 2	Model 4
Autocorrelation	LM(1)	$\chi^2(49) = 65.31$ [0.06]	$\chi^2(49) = 57.83$ [0.18]
	LM(2)	$\chi^2(49) = 52.76$ [0.33]	$\chi^2(49) = 58.93$ [0.16]
ARCH	LM(1)	$\chi^2(784) = 837.9$ [0.09]	$\chi^2(784) = 820.63$ [0.18]
	LM(2)	$\chi^2(1568) = 1176$ [1.00]	$\chi^2(1568) = 1372$ [1.00]
Normality		$\chi^2(14) = 17.52$ [0.23]	$\chi^2(14) = 23.62$ [0.05]

Notes: p-values in parentheses.

Source: Author's analysis based on the data described in Section 1.4.

Table 1.7: Identified long-run structure (Model 2)

	Eigenvectors $\widehat{\beta}$					*Weights* $\widehat{\alpha}$				
Var.	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\beta}_5$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
y_t	1.00	–	–	–	−0.49 (−5.13)	−0.73 (−8.40)	−0.07 (−3.26)	0.17 (6.32)	−0.11 (−4.94)	*
inv_t	−0.16 (−5.56)	1.00	–	–	0.19 (3.61)	−0.89 (−2.51)	−0.36 (−3.83)	0.35 (3.22)	*	−0.69 (−2.97)
aid_t	−0.13 (−3.99)	−0.52 (−8.62)	−0.36 (−3.61)	0.94 (10.62)	–	*	0.55 (3.71)	*	−0.36 (−2.49)	*
m_t	–	–	1.00	–	–	0.63 (2.25)	−0.27 (−3.72)	−0.29 (−3.50)	*	*
ex_t	−0.31 (−6.53)	–	−1.41 (−11.28)	–	–	*	−0.18 (−1.94)	0.24 (2.23)	−0.21 (−2.23)	*
cg_t	–	0.56 (6.24)	–	1.00	–	*	−0.35 (−3.83)	*	−0.48 (−5.40)	*
Δp_t	–	–	–	–	1.00	*	*	*	*	−0.85 (−4.97)
t_{92}	−0.03 (−4.97)	−0.13 (−12.81)	–	–	–					
t	–	–	–	−0.13 (−16.95)	–					

Notes: t-values in parentheses. *Denotes insignificant adjustment coefficients.
Source: Author's analysis based on the data described in Section 1.4.

Table 1.8: Identified long-run structure (Model 3)

	Eigenvectors $\widehat{\beta}$					*Weights* $\widehat{\alpha}$				
Var.	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\beta}_5$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
y_t	1.00	–	–	–	–	−0.48 (−5.97)	*	0.10 (2.90)	−0.08 (−2.47)	*
inv_t	−0.21 (−5.36)	1.00	–	–	–	−0.75 (−2.58)	−0.32 (−3.40)	0.53 (4.28)	*	*
ba_t	−0.06 (−6.51)	−0.46 (−9.10)	–	0.52 (9.60)	–	*	0.54 (2.92)	*	−0.61 (−2.75)	−0.61 (−2.23)
ma_t	−0.06 (−6.51)	–	−0.25 (−13.05)	–	1.00	*	*	*	*	−0.48 (−5.44)
m_t	–	–	1.00	–	–	0.50 (2.12)	−0.29 (−3.81)	−0.39 (−3.85)	−0.21 (−2.26)	*
ex_t	−0.14 (−3.68)	–	–	–	−1.58 (−4.34)	*	*	*	*	*
cg_t	–	0.24 (3.21)	–	1.00	–	*	−0.20 (−2.36)	*	−0.46 (−4.56)	*
t_{92}	−0.03 (−4.75)	−0.10 (−8.68)	−0.13 (−14.77)	−0.03 (−3.25)	0.19 (4.60)					
t	–	–	–	−0.08 (−17.81)	−0.11 (−8.38)					

Notes: t-values in parentheses. *Denotes insignificant adjustment coefficients.
Source: Author's analysis based on the data described in Section 1.4.

Table 1.9: Identified long-run structure (Model 4)

	Eigenvectors $\widehat{\beta}$					*Weights* $\widehat{\alpha}$				
Var.	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\beta}_5$	$\widehat{\alpha}_1$	$\widehat{\alpha}_2$	$\widehat{\alpha}_3$	$\widehat{\alpha}_4$	$\widehat{\alpha}_5$
y_t	1.00	–	–	–	–	−0.54 (−6.03)	*	0.12 (4.39)	−0.09 (−2.70)	*
inv_t	−0.14 (−4.81)	1.00	–	–	–	−0.69 (−2.05)	−0.18 (−2.19)	0.35 (3.39)	*	*
g_t	−0.09 (−4.67)	−0.40 (−7.33)	−0.23 (−4.15)	0.49 (8.20)	–	*	0.52 (3.86)	*	*	−0.13 (−2.91)
l_t	–	–	–	–	1.00	*	*	*	*	−0.37 (−3.38)
m_t	–	–	1.00	–	–	0.85 (3.28)	−0.29 (−4.80)	−0.33 (−4.24)	*	*
ex_t	−0.29 (−5.36)	–	−0.76 (−4.63)	–	−2.44 (−7.98)	*	*	*	*	*
cg_t	–	0.46 (4.57)	–	1.00	–	*	−0.19 (−2.77)	0.16 (1.85)	−0.49 (−4.73)	−0.05 (−2.06)
t_{92}	−0.04 (−5.72)	−0.12 (−10.16)	−0.08 (−3.96)	–	0.20 (−5.02)					
t	–	–	–	−0.10 (−16.23)	−0.11 (−8.12)					

Notes: t-values in parentheses. *Denotes insignificant adjustment coefficients.
Source: Author's analysis based on the data described in Section 1.4.

Table 1.10: The long-run impact matrix C (Model 2)

	$\widehat{\varepsilon}_y$	$\widehat{\varepsilon}_{inv}$	$\widehat{\varepsilon}_{aid}$	$\widehat{\varepsilon}_m$	$\widehat{\varepsilon}_{ex}$	$\widehat{\varepsilon}_{cg}$	$\widehat{\varepsilon}_{\Delta p}$	$\widehat{\varepsilon}^*_{aid}$
y	−0.35 (−1.08)	0.11 (1.59)	**0.10 (2.08)**	0.06 (0.86)	**0.27 (2.00)**	**−0.12 (−2.62)**	0.01 (0.11)	0.03 (1.21)
inv	−2.00 (−1.49)	**0.92 (3.13)**	**0.40 (2.10)**	0.20 (0.70)	0.29 (0.52)	−0.08 (−0.41)	−0.62 (−1.58)	**0.29 (2.62)**
aid	−2.10 (−1.37)	**0.83 (2.47)**	**0.50 (2.28)**	0.28 (0.86)	0.93 (1.47)	**−0.40 (−1.81)**	−0.31 (−0.68)	**0.39 (3.26)**
m	0.21 (0.22)	−0.28 (−1.34)	0.06 (0.44)	0.08 (0.37)	**0.82 (2.07)**	**−0.42 (−3.02)**	**0.54 (1.89)**	−0.04 (−0.36)
ex	0.61 (0.87)	**−0.39 (−2.52)**	−0.06 (−0.63)	−0.00 (−0.03)	0.40 (1.40)	**−0.22 (−2.21)**	**0.46 (2.24)**	0.04 (0.38)
cg	1.81 (1.44)	**−0.75 (−2.70)**	**−0.41 (−2.28)**	−0.23 (−0.84)	−0.65 (−1.25)	0.27 (1.47)	0.35 (0.94)	**−0.26 (−3.05)**
Δp	0.17 (0.54)	**−0.13 (−1.95)**	−0.00 (−0.07)	0.01 (0.18)	**0.23 (1.80)**	**−0.12 (−2.71)**	**0.19 (2.11)**	−0.05 (−1.03)

Notes: Results based on *second-best* choice of rank; t-values in parentheses, with $t \geq 1.80$ given in bold face; $\widehat{\varepsilon}^*_{aid}$ is based on the *second-best* choice of cointegration rank.
Source: Author's analysis based on the data described in Section 1.4.

Table 1.11: The long-run impact matrix C (Model 3)

	$\widehat{\varepsilon}_y$	$\widehat{\varepsilon}_{inv}$	$\widehat{\varepsilon}_{ba}$	$\widehat{\varepsilon}_{ma}$	$\widehat{\varepsilon}_m$	$\widehat{\varepsilon}_{ex}$	$\widehat{\varepsilon}_{cg}$	$\widehat{\varepsilon}^*_{ba}$	$\widehat{\varepsilon}^*_{ma}$
y	−0.14 (−0.63)	0.08 (1.27)	**0.06 (2.65)**	0.00 (0.13)	0.01 (0.23)	**0.23 (3.94)**	**−0.08 (−1.83)**	**0.07 (2.73)**	−0.01 (−0.86)
inv	−1.06 (−1.62)	**0.51 (2.77)**	**0.26 (3.58)**	−0.06 (−0.92)	0.21 (1.13)	0.21 (1.19)	**−0.31 (−2.37)**	**0.25 (2.65)**	0.01 (0.16)
ba	−2.03 (−1.56)	**0.98 (2.70)**	**0.51 (3.58)**	−0.10 (−0.85)	0.40 (1.08)	0.50 (1.46)	**−0.62 (−2.38)**	**0.52 (4.18)**	**−0.22 (−2.86)**
ma	0.71 (0.58)	−0.24 (−0.71)	0.07 (0.51)	0.11 (1.02)	−0.24 (−0.71)	**1.44 (4.53)**	−0.11 (−0.45)	0.04 (0.18)	**0.44 (3.51)**
m	0.40 (1.42)	**−0.17 (−2.17)**	−0.05 (−1.51)	0.04 (1.42)	−0.10 (−1.26)	**0.25 (3.40)**	0.05 (0.91)	−0.05 (−1.73)	**0.08 (4.41)**
ex	0.56 (0.80)	−0.21 (−1.09)	0.00 (0.04)	0.07 (1.15)	−0.17 (−0.86)	**0.81 (4.39)**	−0.02 (−0.13)	0.00 (0.03)	**0.08 (1.81)**
cg	0.98 (1.55)	**−0.48 (−2.68)**	**−0.25 (−3.58)**	0.05 (0.83)	−0.19 (−1.06)	−0.25 (−1.52)	**0.30 (2.38)**	**−0.25 (−4.09)**	**0.09 (2.35)**

Notes: Results based on *second-best* choice of rank; t-values in parentheses, with $t \geq 1.80$ given in bold face; $\widehat{\varepsilon}^*_{ba}$ and $\widehat{\varepsilon}^*_{ma}$ are based on the *second-best* choice of cointegration rank.

Source: Author's analysis based on the data described in Section 1.4.

Table 1.12: The long-run impact matrix C (Model 4)

	$\widehat{\varepsilon}_y$	$\widehat{\varepsilon}_{inv}$	$\widehat{\varepsilon}_g$	$\widehat{\varepsilon}_l$	$\widehat{\varepsilon}_m$	$\widehat{\varepsilon}_{ex}$	$\widehat{\varepsilon}_{cg}$	$\widehat{\varepsilon}^*_g$	$\widehat{\varepsilon}^*_l$
y	−0.17 (−0.63)	0.11 (1.52)	**0.07 (2.08)**	0.00 (0.26)	0.01 (0.18)	**0.27 (3.42)**	−0.07 (−1.22)	**0.07 (2.68)**	0.01 (0.52)
inv	−1.20 (−1.34)	**0.82 (3.25)**	**0.30 (2.56)**	−0.06 (−0.92)	0.13 (0.57)	0.42 (1.56)	−0.15 (−0.75)	**0.20 (2.19)**	0.04 (0.92)
g	−1.58 (−1.40)	**1.08 (3.39)**	**0.38 (2.53)**	−0.08 (−1.08)	0.18 (0.61)	0.40 (1.19)	−0.16 (−0.64)	**0.47 (3.73)**	**−0.17 (−2.56)**
l	0.23 (0.11)	−0.16 (−0.29)	0.28 (1.07)	0.15 (1.10)	−0.09 (−0.18)	**2.25 (3.81)**	−0.53 (−1.21)	−0.24 −0.92	**0.64 (4.62)**
m	0.25 (0.81)	**−0.17 (−1.97)**	−0.01 (−0.18)	0.04 (1.64)	−0.04 (−0.49)	**0.30 (3.24)**	−0.06 (−0.90)	**0.16 (2.10)**	**−0.13 (−3.06)**
ex	0.46 (0.66)	−0.32 (−1.62)	0.01 (0.11)	0.07 (1.55)	−0.08 (−0.43)	**0.72 (3.45)**	−0.15 (−1.00)	0.06 (0.83)	0.03 (0.64)
cg	1.02 (1.29)	**−0.70 (−3.13)**	**−0.27 (−2.57)**	0.04 (0.81)	−0.11 (−0.54)	−0.43 (−1.81)	0.14 (0.82)	−0.15 (−1.73)	−0.07 (−1.67)

Notes: Results based on *second-best* choice of rank; t-values in parentheses, with $t \geq 1.80$ given in bold face; $\widehat{\varepsilon}^*_g$ and $\widehat{\varepsilon}^*_l$ are based on the *second-best* choice of cointegration rank.

Source: Author's analysis based on the data described in Section 1.4.

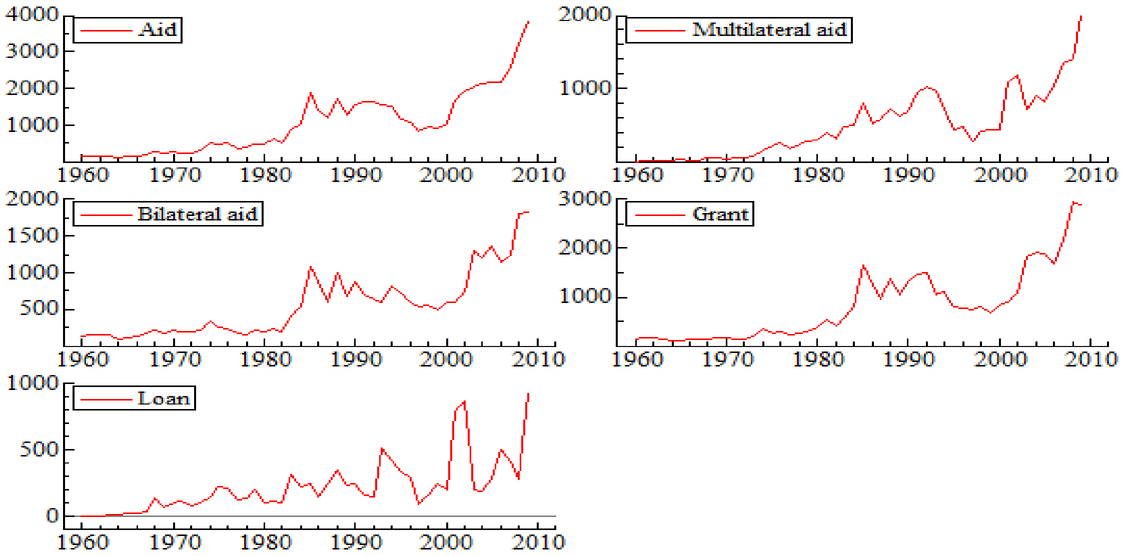

Figure 1.3: Aid and disaggregated aid flows (prior to log transformation)

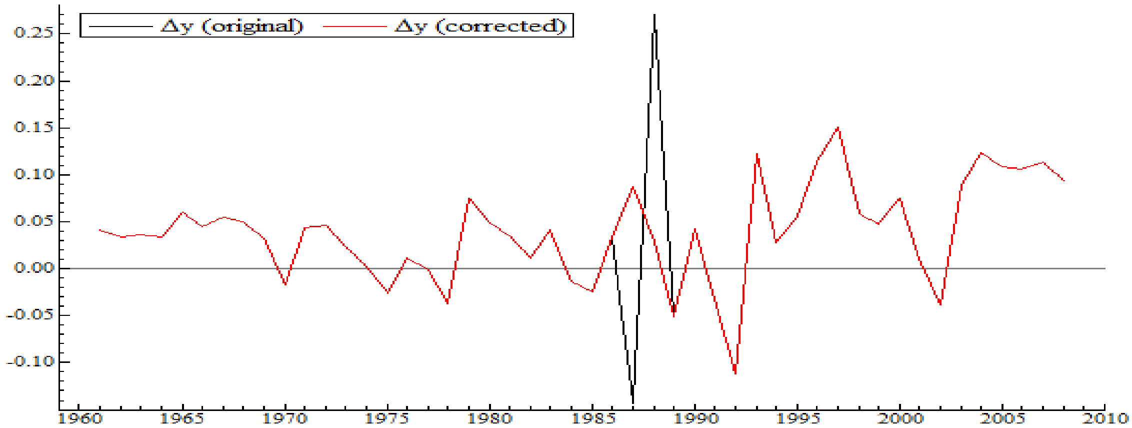

Figure 1.4: GDP: Original and corrected series

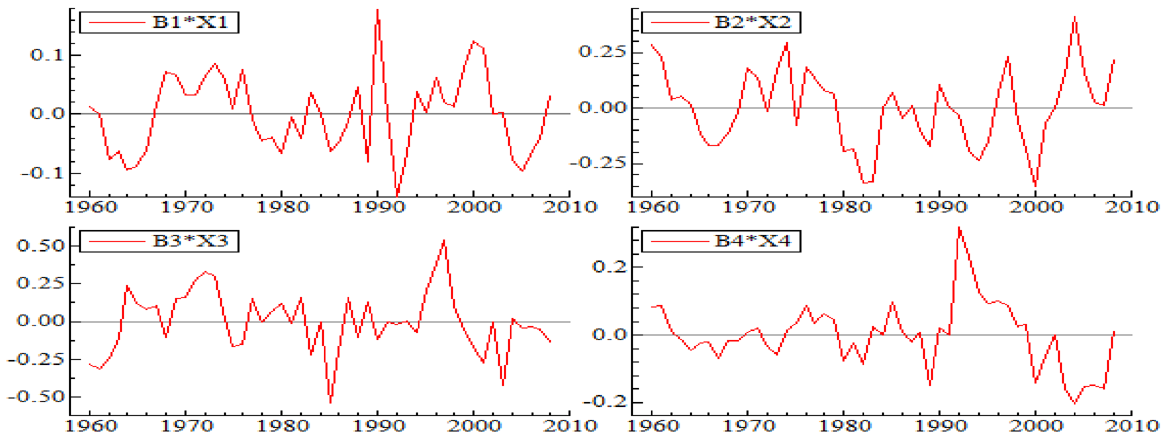

Figure 1.5: Cointegration relations (Model 1)

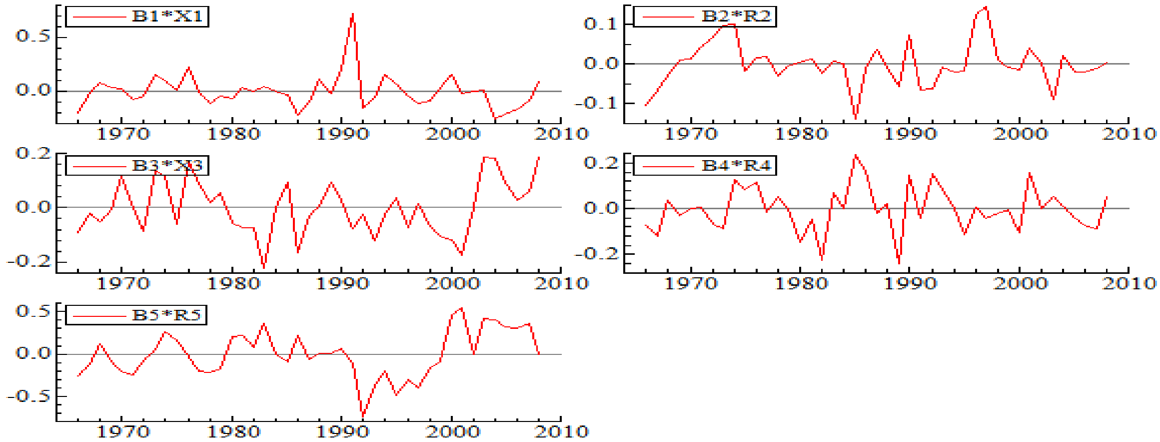

Figure 1.6: Cointegration relations (Model 2)

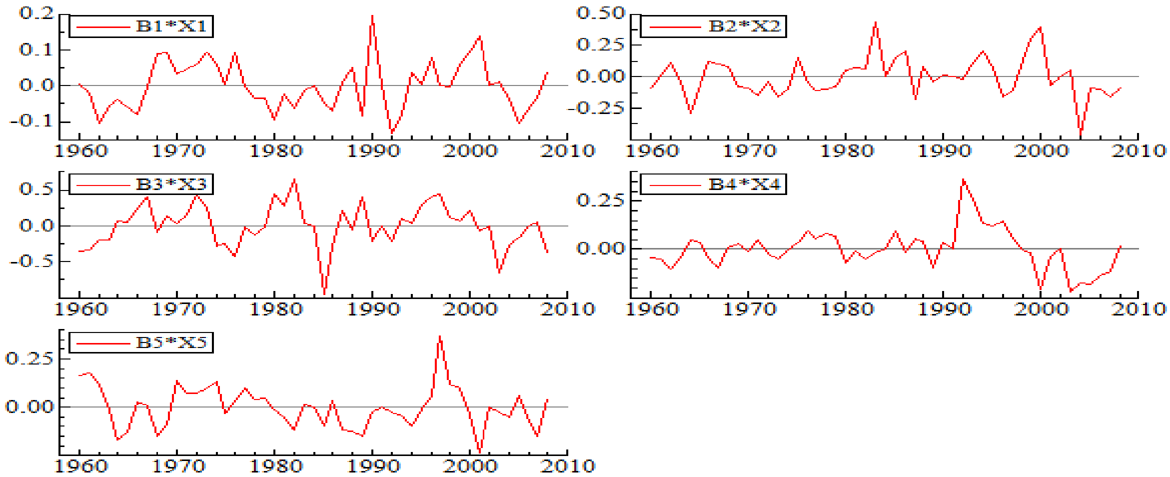

Figure 1.7: Cointegration relations (Model 3)

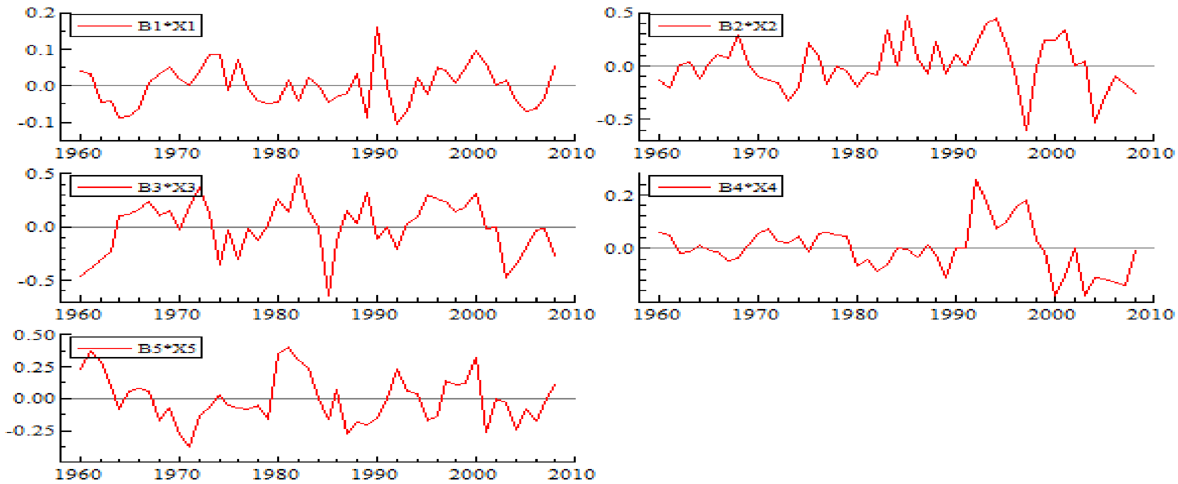

Figure 1.8: Cointegration relations (Model 4)

Chapter 2

Adjustment and Long-Run Economic Performance in 18 African Countries

Accepted for publication in the *Journal of International Development*

Fiseha Haile Gebregziabher
Department of Economics, University of Copenhagen
fiseha.haile.gebregziabher@econ.ku.dk

Adjustment and Long-Run Economic Performance in 18 African Countries*

Fiseha Haile Gebregziabher†

Abstract

This paper investigates the nexus between IMF-World Bank stabilization-cum-structural adjustment programs (SAPs) and long-run economic performance in 18 African countries on a country-specific basis for the period 1960-2009. We employ a structural break approach to study the impact on long-run growth trajectories of the introduction of SAPs. The analysis reveals that only few countries have shown positive and sustained results. The traditional (*first-generation*) Fund-Bank adjustment package is linked with sustained increase in Gross Domestic Product (GDP), export, and investment growth rates only in two countries (Ghana and Uganda). Many African economies remained on their *pre*-reform growth paths whereas some others experienced growth deceleration, despite more-than-a-decade-long adjustment. Taken as a whole, countries in the CFA franc currency zone fared much worse than their non-CFA counterparts due to the different adjustment strategies pursued.

Keywords: Adjustment programs; Sub-Saharan Africa; CFA franc countries; IMF; World Bank; multivariate cointegration.

JEL classification: O1; O4; C32; E6.

2.1 Introduction

During the 1980s and 1990s, IMF and World Bank-sponsored stabilization-cum-structural adjustment programs (SAPs) dominated policy-making in Sub-Saharan Africa (hence-

*I wish to thank two anonymous referees; participants at the PhD seminar at the Department of Economics, University of Copenhagen; and seminar participants at the World Institute for Development Economics Research (UNU-WIDER) for useful comments and suggestions. Discussions with Katarina Juselius and Oliver Morrissey are also gratefully acknowledged. Errors and omissions are my own.

†Fiseha Haile Gebregziabher is affiliated with the Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, 1353 Copenhagen, Denmark. E-mail: *fiseha.haile.gebregziabher@econ.ku.dk*.

forth SSA). Many SSA countries adopted a wide array of policy reforms to redress severe macroeconomic imbalances (burgeoning fiscal and current account deficits, mounting external debt, and misalignment in relative prices) and reinvigorate long-term economic growth.

The question of whether these programs were effective in fostering economic growth has been a subject of widespread contention. Most previous studies find that adjustment lending has insignificant or negative impact on economic growth (e.g. Przeworski and Vreeland, 2000; Easterly, 2005; Barro and Lee, 2005; Dreher, 2006). In contrast, some internal World Bank and IMF studies found positive effects of their programs on growth (e.g. World Bank, 1994, 2002) and contended that Africa's disappointing economic performance reflects a failure to adjust rather than a failure of adjustment. Note, however, that some recent World Bank reports explicitly note that the policy reforms in the 1980s and 1990s have yielded results far below expectations (e.g. World Bank, 2005, *pp.* 9; 2008b, *pp.* 29).

Existing studies have almost universally used cross-country regressions. Although such analysis may throw light on 'what holds on average', it is fraught with numerous limitations. The implicit assumption that parameters are invariant across countries is at best highly restrictive and at worst untenable. No lesson is perhaps more evident than that of the diversity of SSA's adjustment experience in terms of the timing, degree, as well as the nature of policy reforms. In addition, the political, social, and economic circumstances at the country level vary greatly. Moreover, standard cross-country analysis masks distinct periods of sustained changes in growth performance within countries (Jones and Olken, 2008; Hausmann *et al.*, 2005). Thus, lumping such diverse countries together could conceal important information and lead to unfounded conclusions. For this reason, analytical country-studies seem to provide a more reliable backdrop for addressing the question of whether adjustment works. Furthermore, existing country-specific studies tend to be *ad hoc* and are not based on common econometric framework, which impedes comparative analysis.

This paper is an attempt to fill these gaps by investigating how the introduction of adjustment programs affected long-run growth in, *inter alia*, GDP, investment, and export in a sample of 18 SSA countries spanning the period 1960-2009. This is of particular interest because the *raison d'etre* of SAPs was to achieve 'adjustment with growth', i.e. to restore macroeconomic balances while at the same time attaining higher growth rates. Various World Bank/IMF studies indicate that boosting long-term growth was the focal concern of SAPs (see, among others, Corbo *et al.* (1987); World Bank (1994, 2002, 2005)). Against this backdrop, we delve deeper to explore if adjustment policies have resulted in higher and sustained growth rates in SSA on a country-by-country basis. Our country-specific approach allows us to study the similarities and differences of the adjustment experience of SSA countries, and to draw common lessons and garner a broader understanding of how adjustment operates under distinct economic environments.

To set the tone for our analysis: Section 2.2 provides a brief overview of stabilization and structural adjustment programs. Section 2.3 describes the econometric approach, while Section 2.4 discusses the data. Section 2.5 deals with empirical model specification. Section 2.6 presents the results. Finally, Section 2.7 provides concluding remarks.

2.2 Adjustment: A bird's eye view

In the 1980s and 1990s, endeavors to address SSA's severe macroeconomic imbalances centered on the adjustment programs championed by the IMF and World Bank. Broadly speaking, the adjustment strategy designed to restore equilibria was two-pronged: 'stabilization' and 'structural adjustment'. Although both stabilization and structural adjustment strive to restore macroeconomic balances, the means differ.

Stabilization seeks to achieve short-term objectives (e.g. curbing inflation, establishing a sustainable balance of payment position) through the use of certain adjustment instruments (e.g. exchange rate devaluation, monetary restraint, reduction of fiscal deficit). Hence, stabilization attempts to narrow the gap between aggregate production and demand by a reduction in demand, i.e. via *expenditure-reducing policies*. Structural adjustment, on the other hand, aims to foster long-term growth through unleashing the production potential and enhancing the efficiency of the economy. Structural adjustment thus seeks to bridge the gap between aggregate production and demand by increasing production, i.e. via *expenditure-switching* and *long-term supply policies*. Liberalization, deregulation, and privatization constitute the centerpiece of structural adjustment programs (Summers and Pritchett, 1993). As succinctly put by Fischer (1989): "stabilization means adjusting to living within your means whereas structural adjustment means changing the structure of your economy to enable your means to grow more rapidly". Almost by definition, stabilization has been funded by the IMF whereas structural adjustment by the World Bank.

The conventional wisdom posits that *stabilization* is a prior condition for *structural adjustment*. However, the distinction between the two is often tenuous and the same set of policy instruments are often deployed in stabilizing or restructuring an economy. At a more practical level, measures such as currency devaluation, monetary contraction, budget retrenchment, interest rate reform, and tax adjustment can have stabilizing as well as structural effects (Rodrik, 1990). Stabilization and structural adjustment can be mutually reinforcing. Structural adjustment is more likely to be effective when preceded by sufficient doses of stabilization and many structural reforms foster macroeconomic stability. Notwithstanding this congruence, certain policies under structural adjustment, like trade reforms, may engender short-term instability. Thus, the distinction between stabilization and structural adjustment is not of considerable pertinence to our analysis and hereafter the term 'adjustment' will refer to policy reforms that fall under the domain of structural adjustment and stabilization.

2.3 Econometric approach

This paper uses a structural break approach to study the impact on long-run growth trajectories of the introduction of SAPs. In particular, we employ the multivariate cointegration model in Hungnes (2010), which is an alternative formulation of the conventional cointegration model in Johansen *et al.* (2000) and accommodates different types of structural breaks. More specifically, the model in Hungnes (2010) provides a suitable framework for identifying and testing for structural breaks as well as for conducting hypothesis testing on long-run growth rates and means of long-run relations. Using such a multivariate framework, hypotheses about the effect of reforms can be formulated as shifts in equilibrium means and tested as shifts in cointegration means. Similarly, shifts in growth rates can be formulated and properly tested within such framework. In some respects, our approach is similar to that of Hausmann *et al.* (2005), Wacziarg and Welch (2008), and Jones and Olken (2008), who identify episodes of sustained shifts in growth rates and examine explanations for such transitions.

An alternative would be to use a univariate approach and apply some variant of the method proposed by Perron (1989). However, Bai *et al.* (1998) show that there are substantial gains in precision from using multivariate models in which several variables are modelled as cointegrated system. The use of multiple series sharpens inference about the existence and dates of shifts in the mean levels (Bai *et al.*, 1998, *pp.* 420). In other words, a break in mean growth rates might be more readily detected and estimated in a multivariate setting including variables that are purportedly comoving. Moreover, for our purposes, it would be econometrically more tractable to use a system-of-equations approach.

We identify and test for trend breaks using univariate as well as multivariate statistical procedures. In particular, an algorithm searching for breaks developed by Doornik *et al.* (2013) and the procedure in Hungnes (2010) were used to determine the existence, timing, as well as the significance of breaks in mean macroeconomic growth rates.

Sustained shifts in growth rates are defined following Hausmann *et al.* (2005) and Imam and Salinas (2008): (i) For a shift in mean growth rate to be categorized as a growth turnaround it should be sustained for at least 8 years and the change in growth rate has to be at least 2 percentage points; (ii) Countries can have more than one instance of growth turnaround as long as the dates are more than 5 years apart; (iii) Trend breaks were selected at 1% 'target size' (i.e. $\alpha = 0.01$) in the *Autometrics* options in OxMetrics 7 (see Doornik *et al.* (2013)).[1] Note that we thoroughly tested the sensitivity of our findings to relaxing these criteria. We find that the main results prove robust.

2.3.1 The cointegrated VAR model

The cointegrated VAR model is a dynamic adjustment model in growth rates and equilibrium errors:

$$\Delta x_t = \alpha(\beta' x_{t-1}) + \sum_{i=1}^{k-1} \Gamma_i \Delta x_{t-i} + \Phi D_t + \epsilon_t, \qquad t = 1, ..., T, \tag{1}$$

where x_t is a p-dimensional vector of macroeconomic variables that are integrated of order 1, $I(1)$, at most[2]; p is the number of (endogenous) variables; $\beta' x_{t-1}$ are r cointegration (long-run) relations; α is a $p \times r$ matrix of adjustment coefficients (denoting the speed of adjustment to equilibrium); Γ_i is a $p \times p$ matrix of short-run adjustment coefficients, k is the lag length; D_t is an $n \times 1$ vector of n deterministic terms (such as a constant, trend, and dummy variables), Φ is a $p \times n$ matrix of coefficients; $\epsilon_t \sim N.i.i.d(0, \Omega)$ is a $p \times 1$ vector of error terms; and Δ is the first difference operator. If the lag length is 1, then $\Gamma_i = 0$, which implies that the system, once pushed away from equilibrium by exogenous shocks, adjusts back to equilibrium exclusively through α. This is not, however, the case with higher lag lengths, say $k = 2$, where the adjustment dynamics is also influenced by Γ_1, i.e. the system also adjusts to the lagged short-run changes, Δx_{t-1}. See Johansen (1996) and Juselius (2006) for more technical details and applications of the cointegrated VAR model respectively.

2.3.1.1 Deterministic components in the cointegrated VAR model

Allowing for two lags, a constant term (μ_0), a piecewise linear trend ($\mu_1 t + \mu_{11} t_b$), and a step dummy ($\Phi_s Ds_t$), the model in Equation (1) can be rewritten as:

$$\Delta x_t = \alpha(\beta' x_{t-1}) + \Gamma_1 \Delta x_{t-1} + \mu_0 + \mu_1 t + \mu_{11} t_b + \Phi_s Ds_t + \epsilon_t \tag{2}$$

where μ_0 is a vector of constant terms; t is a linear trend $(1, 2, 3, ...)$, μ_1 is a vector of trend coefficients; t_b is a broken linear trend $(...0, 0, 0, 1, 2, 3, ...)$, μ_{11} is a vector measuring change in trend slope; Ds_t is a $q \times 1$ vector of unrestricted step dummies

[1] The target size determines the significance level below which a break is not kept in the model. Note that for Burkina Faso the trend break was significant only at the 5% level.

[2] The series x_t is said to be integrated of order d, i.e. $x_t \sim I(d)$, if it is stationary when differenced d times, which means x_t contains d unit roots. In other words, x_t is integrated of order d if x_t has the representation $(1-L)^d x_t = C(L)\varepsilon_t$, where $C(1) \neq 0$, and $\varepsilon_t \sim IN(0, \Omega)$.

$(...0, 0, 0, 1, 1, 1, ...)$, and Φ_s is a $p \times q$ matrix of coefficients.

Because the cointegrated VAR model describes p equations, Δx_t, and r long-run relations, $\beta' x_t$, in the same model, the effect of deterministic terms will differ between the long-run relations and the short-run dynamics. Hence, the coefficients (μ_0, μ_1, μ_{11}, Φ_s) have to be decomposed into components that induce growth in the system, $E(\Delta x_t)$, and those that capture the means of long-run relations, i.e. $\mu_0 = \alpha\beta_0 + \theta_0$, $\mu_1 = \alpha\beta_1 + \theta_1$, $\mu_{11} = \alpha\beta_{11} + \theta_{11}$, and $\Phi_s = \alpha\phi_0 + \phi_1$. See Juselius (2006: Chapter 6). The piecewise linear trend has to be restricted to the cointegration relations to avoid quadratic trends in the data and thus we impose the restriction $\theta_1 = \theta_{11} = 0$ (i.e. $\mu_1 = \alpha\beta_1$ and $\mu_{11} = \alpha\beta_{11}$):

$$\Delta x_t = \alpha(\beta' x_{t-1}) + \Gamma_1 \Delta x_{t-1} + (\alpha\beta_0 + \theta_0) + \alpha\beta_1 t + \alpha\beta_{11} t_b + (\alpha\phi_0 + \phi_1) Ds_t + \epsilon_t \quad (3)$$

where $\beta_0 + \phi_0 Ds_t$ is the equilibrium mean, with $\phi_0 Ds_t$ capturing the shift in equilibrium mean. θ_0 is the slope of the linear trend, whereas $\phi_1 Ds_t$ measures the change in the slope of the linear trend. Restricting the piecewise linear trend to the cointegration relations is tantamount to allowing for the possibility that the long-run relations can be stationary around a linear trend, possibly with a changing slope in the year $19yy$.

As alluded to above, the deterministic components in the conventional formulation of the cointegrated VAR model, as in Equations (2) and (3), are not amenable to straightforward interpretation. This can be demonstrated, without loss of generality, using the following simple model with a linear trend restricted to the cointegration relations:

$$\Delta x_t = \alpha(\beta' x_{t-1} + \beta_1' t) + \Gamma_1 \Delta x_{t-1} + \mu_0 + \epsilon_t \quad (4)$$

where all remaining deterministic terms in Equation (3) are excluded for brevity.

Under some regularity conditions, Johansen (1991, 1996) shows that the expression in Equation (4) has the moving-average representation:

$$x_t = C \sum_{i=1}^{t} \epsilon_i + C(L)\epsilon_t + \gamma t + z_0 \quad (5)$$

where $C = \beta_\perp(\alpha_\perp' \Gamma \beta_\perp)^{-1} \alpha_\perp'$ is a $p \times p$ long-run impact matrix with rank $p - r$; $\Gamma = I_p - \Gamma_1$, I_p is a $p \times p$ identity matrix; $\beta_\perp$ and $\alpha_\perp$ are orthogonal complements of β and α respectively (such that $\beta_\perp' \beta = 0$, $\alpha_\perp' \alpha = 0$, and $(\alpha, \alpha_\perp)$ and $(\beta, \beta_\perp)$ have full rank); $C(L)$ is a stationary lag polynomial describing impulse response effects; $\gamma = C\mu_0 + C(L)\alpha\beta_1 = C\mu_0 - (I_p - C\Gamma)\beta(\beta'\beta)^{-1}\beta_1$; and z_0 is a function of the initial values of the process (see Johansen, 1996, *pp.* 82). This shows that the trend coefficients, γ, are complicated nonlinear functions of the coefficients in Equation (4).

2.3.2 Modelling the effect of major adjustment reforms

Our baseline VAR model is specified with two lags and a linear trend restricted to the cointegration relations:

$$\Delta x_t = \alpha(\beta' x_{t-1} + \beta_1' t) + \Gamma_1 \Delta x_{t-1} + \mu_0 + \epsilon_t \tag{6}$$

where $x_t' = [y_t, inv_t, ex_t, m_t, cg_t, aid_t]$ is a 6×1 vector of macroeconomic variables (see Section 2.4) and the rest are as defined previously. The differences Δx_t represent growth rates because the variables are given in logs.

Many SSA countries have experienced a significant shift in the underlying long-run trends (and thus mean-shift in the growth rates) of the macrovariables after the advent of IMF/World Bank adjustment programs (see Section 2.6). The effect of major adjustment policy reforms on the growth rates, Δx_t, can be modelled by allowing for a piecewise linear trend, $\beta_1 t + \beta_{11} t_b$, in the long-run relations. This is tantamount to augmenting the baseline VAR model in Equation (6) with $\alpha\beta_{11}t_{yy}$, $\Phi_s D_s yy_t$, and $\psi D_p yy_t$:

$$\Delta x_t = \alpha(\beta' x_{t-1} + \beta_0 + \beta_1' t + \beta_{11}' t_{yy}) + \Gamma_1 \Delta x_{t-1} + \theta_0 + \Phi_s D_s yy_t + \psi D_p yy_t + \epsilon_t \tag{7}$$

where t_{yy} is a broken linear trend $(...0, 0, 0, 1, 2, 3, ...)$ starting at the date of reform, $19yy$, and restricted to the cointegration relations; β_{11} measures the change in the linear trend coefficient of the cointegration relations after the advent of policy reforms; $D_s yy_t$ is an $m_1 \times 1$ vector of unrestricted step dummies $(...0, 0, 0, 1, 1, 1, ...)$ starting in the year $19yy$ and controls for change in growth rates as well as shift in the means of long-run relations, Φ_s is a $p \times m_1$ matrix of coefficients to the step dummies; $D_p yy_t$ is an $m_2 \times 1$ vector of unrestricted impulse dummies $(..., 0, 0, 0, 1, 0, 0, 0...)$ and accounts for an unanticipated one-period shock effects in $19yy$, ψ is a $p \times m_2$ matrix of coefficients to the impulse dummies; and yy stands for the year $19yy$.[3]

As shown in the previous section, the conventional formulation of the cointegrated VAR model does not allow straightforward interpretation of the deterministic components. This necessitates reformulating the model in Equation (7) such that the deterministic terms are easily identifiable and directly interpretable. Hungnes (2010) suggests an alternative formulation that facilitates the decomposition of all deterministic terms into components that induce growth in the system, Δx_t, and those that

[3] The impulse dummies exclusively control for the exceptionally large shocks at the time of occurrence but preserve the information of the observations through their lagged impact. Put differently, the dummies account for unanticipated shocks and given that the latter are no longer unanticipated in the next period, their lagged effects on the system are accounted for by the dynamics of the model. Thus, unlike the case with static regressions, the dummies do not eliminate the corresponding observations.

capture the means of cointegration relations, $\beta' x_t$:

$$\Delta x_t - \gamma \Delta D_t = \alpha \left[\beta'(x_{t-1} - \gamma D_{t-1}) - \beta_0\right] + \Gamma_1 \left[\Delta x_{t-1} - \gamma \Delta D_{t-1}\right] + \epsilon_t \qquad (8)$$

where D_t is an $m \times 1$ vector containing the deterministic variables in Equation (7) and γ is a $p \times m$ matrix of coefficients.[4] Note that $\delta \equiv \beta' \gamma$ measures the effect of the deterministic variables, D_t, on the long-run relations, $\beta' x_t$.[5]

It can be shown that x_t in Equation (8) has the moving-average representation:

$$x_t = C \sum_{i=1}^{t} \epsilon_i + C(L)\epsilon_t + \gamma D_t + z_0 \qquad (9)$$

Unlike the case with the conventional formulation of the cointegrated VAR model, each coefficient in γ has a straightforward interpretation and describes the effect of the deterministic components, D_t, on the endogenous variables, x_t.

For countries that have had no statistically significant shift in the long-run trends (i.e. mean-shift in the growth rates) of the macrovariables following the introduction of adjustment programs, we include only a linear trend restricted to lie in the cointegration space:

$$\Delta x_t - \gamma_1 = \alpha(\beta' x_{t-1} - \beta_0 - \beta_1(t-1)) + \Gamma_1 \left[\Delta x_{t-1} - \gamma_1\right] + \epsilon_t \qquad (10)$$

where $\beta_1 \equiv \beta' \gamma_1$ is the vector of linear trends in the cointegration relations. $E\left[\Delta x_t\right] = \gamma_1$ picks up the unconditional growth rates of the variables, while $E\left[\beta' x_t\right] = \beta_0 + \beta_1 t$ captures the means of the cointegration relations. Put another way, γ_1 describes the long-run (steady state) growth rates of the macrovariables for countries that have not experienced shift in long-run trends. Note, however, that Equation (10) disregards, for brevity, the step and impulse dummies that we will include to account for extraordinary events (see Section 2.5).

For countries that have experienced a statistically significant change in trend slopes (i.e. when the model also contains a broken linear trend in the long-run relations) following the introduction of adjustment programs, the model to be estimated is given

[4] Note that Equations (7) and (8) are alternative ways of writing the same system only in some exceptional cases (see Hungnes, 2010). Generally, a one-to-one relationship between (7) and (8) does not exist. For example, when D_t includes step or impulse dummies, as in Equation (7), there is no straightforward way to transform Equation (8) into Equation (7). The underlying assumption in Equation (8) is that the deterministic variables in D_t are such that the $m \times T$ matrices $(D_t)_{t=1}^{T}$ and $(\Delta D_t)_{t=1}^{T}$ have full rank (Hungnes, 2010). This implies that we cannot include a constant term in D_t (because ΔD_t becomes 0). Put in other words, we cannot estimate the levels of the variables, x_t, in a cointegrated system of $I(1)$ variables. We can, however, estimate the levels in the direction of the cointegration space, which is captured by β_0.

[5] This can be seen by rewriting the cointegration space as $\beta'(x_{t-1} - \gamma D_{t-1}) - \beta_0 = \beta' x_{t-1} - \delta D_{t-1} - \beta_0$.

by:

$$\Delta x_t - \gamma_1 - \gamma_{11} D_s yy_t = \alpha \left[\beta' x_{t-1} - \beta_0 - \beta_1 (t-1) - \beta_{11}(t_{yy} - 1)\right] + \Gamma_1 \left[\Delta x_{t-1} - \gamma_1 - \gamma_{11} D_s yy_{t-1}\right] + \epsilon_t \tag{11}$$

where $\beta_1 \equiv \beta' \gamma_1$ and $\beta_{11} \equiv \beta' \gamma_{11}$. β_{11} is a vector of broken trends measuring the change in trend slopes in the long-run relations in $19yy$. In this case, there is a need to additionally account for the change in underlying trends (and thus the corresponding shift in long-run growth rates) that ensued adjustment policy reforms, which is captured by $\gamma_{11} D_s yy_t$. Thus, the long-run (steady state) growth rates of the macrovariables are given by $E[\Delta x_t] = \gamma_1 + \gamma_{11} D_s yy_t$ (where $D_s yy_t = \Delta t_{yy}$). In other words, γ_1 is the overall long-run trend underlying the macrovariables prior to the introduction of adjustment programs, while $\gamma_{11} D_s yy_t$ accounts for the change in long-run trends that accompanied adjustment programs. Note that equation (11) will be augmented with dummies to control for special events.

2.4 Data

The data are annual observations for the period 1960-2009 and comprise the variables: GDP (y_t); gross investment (inv_t) (comprising both private and public outlays); exports of goods and services (ex_t); imports of goods and services (im_t); government consumption expenditure (cg_t); and Official Development Assistance (ODA) *net*-disbursements (aid_t).[6] All variables are at constant market prices. The aid data are obtained from the OECD online database[7], while the rest are from the *Penn World Tables* (PWT− Version 7.0) and *World Development Indicators* (WDI).

For most of the countries included in the sample, we opt for the data from the PWT as they span a longer period of time. For Cote d'Ivoire, Zambia, and Sudan, the data are from WDI. PWT data are of low quality for the first two countries and available only from 1970 for Sudan. WDI data are scanty for many SSA countries, the use of which would hamper comparative analysis. To make sure that the results are not artifacts of the PWT data set, we test their robustness using data from the WDI (see Section 2.6). It should be pointed out upfront that we are fully cognizant of the problems with the quality of data for SSA countries. However, these are the very same data that have been extensively used in the literature and thus we rely on the same imperfect figures.

We judge the long-run impact of policy reforms by their effect on some major indicators of economic performance: GDP, investment, export, and import growth

[6] ODA includes grants and all loans with a grant element of more than 25% as well as technical cooperation and assistance, but excludes aid for military purposes.

[7] Available at http://stats.oecd.org/.

rates. GDP growth rate is regarded as the most telling indicator of the success or failure of adjustment efforts over the long-term. Adjustment programs were targeted at, *inter alia*, redressing external imbalances. Hence, the analysis includes export and import as indicators of trade performance. Investment is another leading indicator because adjustment programs were directed toward increasing gross domestic investment. Government spending and aid flows are included to explore whether changes in these variables can help explain growth transitions. The choice of variables is partly justified by the findings of some recent studies (e.g. Hausmann *et al.* (2005); Imam and Salinas (2008)), who find that growth turnarounds are associated with increases in investment and trade, and with real exchange rate depreciations, whereas the link with changes in terms-of-trade and aid flows is weak. Due to the paucity of data, our analysis omits some important variables such as the terms-of-trade and real exchange rate. However, we complement the econometric analysis with a descriptive analysis using the available data.[8]

Finally, it should be pointed out upfront that we are fully cognizant of the problems with the quality of data for SSA countries. However, these are the very same data that have been extensively used in the literature and thus we rely on the same imperfect figures.

2.5 Model specification[9]

2.5.1 Specification tests

SSA countries are quite diverse in terms of the genesis and severity of the economic crisis they encountered in the late 1970s and 1980s, the policy reforms implemented in response, and the consequences of these changes. The main thrust in the selection of countries was thus the desire to allow for as much diversity as possible. Some countries were omitted from the analysis either because they have not adopted SAPs (such as Botswana, South-Africa, Angola, Namibia, Swaziland, and Liberia) or due to a large number of missing observations and poor quality of the data. Many SSA countries became independent in the early 1960s and the initial years of the *post*-colonial era in the newly established fledgling states were often turbulent and gradual. Using a statistical test procedure in Nielsen (2008), Juselius *et al.* (2014) spot the first five

[8]Note, however, that although some variables are omitted from our analysis, this does not, in general, invalidate the long-run estimates. The reason is that cointegration property is invariant to extensions of the information set (Juselius, 2006).

[9]The software packages CATS in RATS (Dennis *et al.*, 2006) and GRaM (*G*rowth *Ra*tes and cointegration *M*ean levels) (Hungnes, 2005) were used to carry out all computations. The latter was used to decompose the deterministic components in the VAR models of interest. The program code for CATS as well as the details of the specification test results for each country can be obtained from the author upon request.

years, $1960-1965$, as excessively influential for many of the countries in our sample. There are substantial gains from leaving out such non-representative years and thus we removed them prior to the subsequent empirical analysis. Table 2.1 reports the choice of sample period for each country.

The VAR model is derived under the assumption of constant parameters. Although parameter stability can be assessed using recursive test procedures, the small number of observations at our disposal circumscribes the power of the available recursive procedures. However, since parameter non-constancy is often engendered by periods of political and economic turbulence, such as war, severe droughts, civil unrest, interventions, and policy reforms, we improve parameter stability by controlling for the most dramatic events using different types of dummy variables. For example, a shift in the equilibrium mean can be captured by a step dummy, $D_s yy_t$, defined as $(0, ..., 0, 0, 0, 1, 1, 1, ..., 1)$, while a one-period shock effect can be accounted for by an impulse dummy, $D_p yy_t$, defined as $(0, ..., 0, 0, 0, 1, 0, 0, 0, ..., 0)$. If the step dummy, $D_s yy_t$, is restricted to lie in the cointegration space and the model has two lags, an impulse dummy, $D_p yy_t = \Delta D_s yy_t$, will automatically enter the model unrestrictedly. A transitory impulse dummy, $D_{tr} yy_t$, defined as $(0, ..., 0, 0, 0, 1, -1, 0, 0, 0, ..., 0)$ also enters the VAR model for some of the sample countries.

Although controlling for the effect of special events improves model specification, another reason for concern is the presence of additive outliers.[10] Such outliers often distort inference in small samples (Franses and Haldrup, 1994; Nielsen, 2004). Although it is not uncommon in time-series analysis to spot outlying observations and model them with dummies, inference based on incorrect configuration of dummies can be seriously misleading (Nielsen, 2004). A probing look at the plots of the series (see Appendix Figure 2.1) coupled with *a priori* knowledge on the timing of special events pointed to three aberrant observations. These correspond to GDP in 1987 in Ethiopia and government consumption spending in 2005 in both Ghana and Tanzania. Because these outliers are likely to be excessively influential and are not part of the VAR dynamics, we removed them prior to the econometric analysis.[11]

With the deterministic specifications and the dummies included, the individual country models pass most of the specification tests and describe the data reasonably well. It is, however, hard to ensure complete parameter constancy and thus it should be borne in mind that some of the estimates may represent *average historical effects* over the sample period.

[10] An additive outlier is an aberrant observation (i.e. unrelated to the data-generating process) and often occurs due to typing mistakes or gross measurement errors.

[11] Though an alternative is to account for these outliers using additive dummy variables, this could induce spuriously delayed effects which potentially biases the model estimates (Juselius, 2006, *pp.* 108).

2.5.2 Determination of the cointegration rank

After having established an adequate statistical description of the data, the next step is to determine the cointegration rank for the individual country models. The cointegration rank classifies the data into r relations towards which the process is adjusting (the *pulling forces*) and $p - r$ relations which are pushing the process (the *pushing forces*). The choice of rank is of paramount importance because it affects the entire analysis in subsequent stages.

The test for r cointegrating vectors is based on the maximum likelihood test procedure, known as the trace test (see Johansen (1988, 1996)). The trace test is based on a sequence of tests of the null of $p - r$ unit roots for $r = 0, 1, 2, ..., p - 1$ and relies on the premise of '*no prior economic knowledge*' about the rank r. Note, however, that some of these statistical null hypotheses may not coincide with plausible economic null hypotheses (Juselius, 2006: Chapter 8). This is usually the case for large values of $p-r$ (many stochastic trends) and small values of r (few equilibrium relations), consistent with the presumption in economic theory that macroeconomic variables co-move in the long-run. Hence, it is crucial to specify beforehand an economic prior for the number of autonomous shocks, $p-r^*$, where r^* corresponds to the number of long-run relations consistent with this prior. This helps avoid the risk of not rejecting an implausible economic null just because it constitutes a conveniently testable statistical null.

The system variables are likely to be affected by long-run trends associated with cumulative productivity shocks and trends in population, which are proxied by the deterministic trends. Moreover, given that all variables are in real terms, we expect at least two stochastic trends: one emanating from external shocks (such as terms-of-trade fluctuations), captured by shocks to exports, and the other corresponding to persistent medium long-run business cycle movements in the data. However, if aid is found to be exogenous in the system, as it is often presumed, it could constitute a third driving trend. Thus, in most cases, the economic prior would correspond to "$r = 4$, $p - r = 2$" or "$r = 3$, $p - r = 3$". Due to the importance of the choice of rank for the subsequent steps, we perform a sensitivity analysis to check if the empirical estimates based on the statistically most credible value of rank, r^*, are fairly robust to altering the rank to the *second-best* choice, either $r^* - 1$ or $r^* + 1$.

The choice of rank is made based on a range of statistical criteria, such as the trace test, the largest unrestricted root of the characteristic polynomial for a given r, the t-ratios of the α coefficients for the r^{th} cointegration vector, and the graphs of the r^{th} cointegration relation (see Juselius, 2006: Chapter 8). Rank determination is conventionally made based on the trace test. However, substantial size and power distortion plague the trace test when the size of the sample is small. A cause for concern in this regard is that the application of a small-sample correction to the trace test results in a correct size, but it does not necessarily solve the power problem (Johansen, 2002). Thus, instead of just using the trace test, we base rank determination on the aforementioned sources of information.

The last two columns of Table 2.1 report the most credible value (*first-best* choice) of cointegration rank as well as the *second-best* alternative for each country. The test results show that $r = 4$ is the *first-best* choice of rank for the vast majority of 11 countries, whereas $r = 3$ is empirically optimal for 5 countries. A sensitivity analysis suggested that $r = 3$ is the *second-best* choice for 12 countries. Due to the importance of the choice of rank for the subsequent analysis we report all estimates both for the *first-* and *second-best* choices of rank.

2.6 Results

In discussing the link between adjustment reforms and long-run economic performance in the 18 SSA countries, we define two broad groups: the CFA franc zone countries[12] and non-CFA countries. Among the sample countries, Burkina Faso, Cameroon, Cote d'Ivoire, Niger, and Senegal are members of the CFA franc currency zone. Up until 1994, these countries maintained a fixed exchange rate with the French franc. In January 1994, the CFA franc was devalued for the first time since 1948, by 50 percent. Hence, throughout the 1980s and early 1990s, the rules of the CFA zone precluded nominal devaluation and the member countries were left only with alternative expenditure-switching policy measures (trade taxes, quotas, export subsidies) to arrest external imbalances. As a result, trade and exchange rate policy reforms in the CFA countries differed markedly from those in their non-CFA counterparts.

In general, the macroeconomic performance of the CFA countries outstripped that of the non-CFA countries before the early 1980s (Devarajan and De Melo, 1987). The superior performance is partly ascribed to the monetary and fiscal discipline imposed by the rules of membership in the CFA zone (*Ibid.*). However, the performance of the franc zone countries deteriorated in the 1980s due to the difficulty in adjusting the real exchange rate in response to terms-of-trade shocks (Devarajan and Rodrik, 1991). This reveals an important tradeoff during the adjustment period. On the one hand, CFA zone members enjoyed spectacularly lower inflation owing to the fixed exchange rate regime. On the other hand, they were virtually unable to use nominal devaluation as an instrument of adjustment. Apart from the fixed exchange rate, non-CFA and CFA zone countries shared many commonalities. They became independent states at similar times and were roughly at comparable stages of development. All relied on

[12] The CFA franc zone consists of 14 countries in SSA which maintain the same currency, the CFA franc.

Table 2.1: Countries, sample period, dummy variables, broken linear trend (t_{yy}), and cointegration rank (first- and second-best choices)

Country	Sample period	Dummy variables‡	t_{yy}	Rank 1^{st}	Rank 2^{nd}
Burkina Faso	1964-2009	D$p94t$, D$p99t$, D$p00t$	$\mathbf{t_{94}}$	3	4
Cameroon	1965-2009	D$s88t$, D$s95t$, D$p76t$, D$p91t$	$\mathbf{t_{88}}$,$\mathbf{t_{95}}$	4	3
Cote d'Ivoire	1960-2009	D$s81t$, D$p80t$, D$p84t$	$\mathbf{t_{81}}$	4	3
Ethiopia	1960-2009	D$s92$t, D$p85t$, D$p03t$	$\mathbf{t_{92}}$	4	3
Gambia, The	1960-2009	D$s68t$, D$p82t$, D$p05t$	None	3	4
Ghana	1966-2009	D$s83t$, D$p72t$, D$p83t$	$\mathbf{t_{83}}$	4	5
Guinea	1963-2009	D$s97t$, D$p67t$, D$p70t$, D$p90t$	None	4	3
Kenya	1965-2009	D$s85t$, D$p68t$, D$p93t$	$\mathbf{t_{85}}$	3	2
Madagascar	1965-2009	D$p86t$, D$p97t$, D$p02t$	None	2	3
Malawi	1965-2009	D$s82t$, D$p94t$	$\mathbf{t_{82}}$	2	3
Mozambique	1960-2009	D$s87t$, D$p83t$, D$tr94t$, D$p92t$	$\mathbf{t_{87}}$	4	3
Niger	1965-2009	D$s73t$, D$s84t$	None	4	3
Rwanda	1960-2009	D$s95t$, D$p63t$, D$p94t$, D$p00t$	$\mathbf{t_{95}}$	3	4
Senegal	1965-2009	D$s94t$, D$p69t$, D$p94t$	$\mathbf{t_{94}}$	4	3
Sudan	1960-2009	D$s92t$, D$p96t$, D$p00t$	$\mathbf{t_{92}}$	4	3
Tanzania	1960-2009	D$s92t$	None	4	3
Uganda	1964-2009	D$s87t$, D$p79t$, D$p94t$, D$p05t$	$\mathbf{t_{87}}$	3	4
Zambia	1963-2009	D$p07t$	None	4	3

Notes: t_{yy} is a broken linear trend, where 19yy is the period at which a change in trend slope was allowed for; ‡$D_{tr}yy_t$ is a transitory impulse dummy; †Data are from *World Development Indicators* (WDI); *CFA franc zone countries.
Source: http://stats.oecd.org/; WDI; Penn World Tables 7.0 (Heston *et al.*, 2012).

primary products and faced similar external shocks.

Of the 18 SSA countries, the analysis finds that 8 non-CFA and 4 CFA franc zone countries experienced significant trend breaks after the advent of adjustment programs. Seven of these countries are characterized by growth acceleration. Table 2.1 reports the dates at which we allowed for a break in the linear deterministic trend for these countries. We identified these dates based on the statistical procedures discussed in Section 2.3. The hypotheses that these countries have had no significant shift in the mean growth rates of the series at the specified dates are strongly rejected (p-value: 0.00). However, we perform a sensitivity analysis to examine if the estimates based on the statistically as well as economically most credible break date, denoted here by $19yy^*$, are fairly robust to alternative candidate break points in the vicinity of $19yy^*$. We find that the main findings of this paper are sufficiently robust to changes in the break dates. In particular, the growth estimates corresponding to the variables of primary interest, namely GDP, investment, and exports, are found to be fairly consistent across alternative break points. To facilitate understanding of the shift in long-run trends that ensued adjustment policy reforms, Appendix Figures 2.2 – 2.10

show the graphs of some of the macrovariables (in levels and first differences) for many of the countries in the sample.

Tables 2.2 and 2.3 report the estimates of long-run growth rates, γ_1, along with their t-values, for those 12 countries and show how they changed after adjustment reforms were ushered in, which is measured by γ_{11}^{19yy}. Tables 2.4 and 2.5 (Appendix) show the estimates of long-run growth rates, γ_1 and γ_{11}^{19yy}, for the *second-best* choices of cointegration rank.

The remaining 5 non-CFA and 1 CFA countries have seen no significant shift in mean growth rates. The estimates of long-run trends for these countries are reported in Table 2.7 (Appendix). We see from Tables 2.4, 2.5, and 2.7 (Appendix) that our main findings are sufficiently robust to changes in the cointegration rank. Table 2.6 (Appendix) reports the results obtained using data from the WDI, albeit only for countries for which WDI data are available over a sufficiently long period of time and for the *first-best* choices of rank. The main findings are fairly robust to changes in the source of data and thus are not artifacts of the PWT data set.

Sections 2.6.1 presents the summary of the results, while Section 2.6.2 discusses the adjustment experience of the non-CFA and CFA countries that have seen sustained shift in long-run trends. Appendix A provides a very brief discussion on the adjustment experience of the remaining countries.

2.6.1 Summary of results

2.6.1.1 Non-CFA countries[13]

Of the 13 non-CFA countries, 5 of them (Ethiopia, Ghana, Mozambique, Rwanda, and Uganda) saw a considerable upturn in long-run GDP growth rate after reforms, whereas Malawi suffered a severe decline. The remaining 7 countries managed only to preserve their *pre*-reform growth rates. It may be surprising that many countries experienced no sustained increase in GDP growth, despite having successfully implemented more-than-a-decade-long adjustment programs. In particular, the case of Malawi represents a notable disaster as it experienced a sharp drop in long-run GDP growth, though it was one of the strong adjusters in SSA (World Bank, 1994).

Five countries (Ethiopia, Ghana, Kenya, Rwanda, and Uganda) saw their long-run average growth rate of investment pick up after adjustment; Malawi had a striking drop; whereas the remaining 6 countries experienced no change.

In 5 of the 13 non-CFA countries (Ethiopia, Ghana, Mozambique, Rwanda, and Uganda), the long-run growth rate of exports edged up significantly. Kenya saw a slight increase. The performance was, however, uneven and export growth remained unal-

[13]Note that in summarizing the results we disregard Sudan since its reform program in the early 1990s was not backed by Fund-Bank adjustment loans.

Table 2.2: Long-run (steady state) growth rates of the macrovariables (countries that experienced shift in long-run trends): *first-best* choice of rank

	Burkina Faso*			**Cameroon***			
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1994}}$	$\gamma_1+\gamma_{11}^{1994}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1988}}$	$\boldsymbol{\gamma_{11}^{1995}}$	$\gamma_1+\gamma_{11}^{88}+\gamma_{11}^{95}$
Δy_t	0.03 (31.96)***	0.02 (8.75)***	0.05	0.07 (12.41)***	−0.10 (8.42)***	0.06 (5.20)***	0.03
Δinv_t	0.08 (7.09)***	−0.01 (2.72)***	0.07	0.10 (12.15)***	−0.14 (6.00)***	0.10 (4.03)***	0.06
Δex_t	0.05 (3.66)***	0.04 (1.48)	0.09	0.05 (2.05)**	0.01 (0.24)	−0.03 (0.57)	0.03
Δim_t	0.04 (2.87)***	0.01 (0.44)	0.05	0.05 (2.36)**	−0.02 (0.51)	0.03 (0.75)	0.06
Δcg_t	0.06 (3.75)***	0.004 (0.14)	0.064	0.08 (6.87)***	−0.10 (4.13)***	0.07 (2.53)**	0.05
Δaid_t	0.06 (3.93)***	0.00 (0.012)	0.06	0.00 (0.00)	0.02 (0.28)	0.03 (0.36)	0.05

	Cote d'Ivoire*			**Ethiopia**		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1981}}$	$\gamma_1+\gamma_{11}^{1981}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1992}}$	$\gamma_1+\gamma_{11}^{1992}$
Δy_t	0.08 (35.91)***	−0.06 (17.30)***	0.02	0.02 (5.43)***	0.04 (6.32)***	0.06
Δinv_t	0.13 (4.32)***	−0.15 (3.23)***	−0.02	0.02 (2.88)***	0.06 (3.92)***	0.08
Δex_t	0.05 (5.29)***	−0.02 (1.56)	0.03	0.02 (1.69)*	0.10 (4.30)***	0.12
Δim_t	0.10 (8.39)***	−0.07 (3.80)***	0.03	0.03 (3.37)***	0.08 (4.04)***	0.11
Δcg_t	0.12 (15.25)***	−0.12 (10.93)***	0.00	0.04 (3.63)***	0.06 (3.03)***	0.10
Δaid_t	0.07 (2.83)***	−0.10 (2.78)***	−0.03	0.09 (4.40)***	−0.03 (0.76)	0.06

	Ghana			**Kenya**		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1983}}$	$\gamma_1+\gamma_{11}^{1983}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1985}}$	$\gamma_1+\gamma_{11}^{1985}$
Δy_t	0.01 (6.67)***	0.04 (24.55)***	0.05	0.04 (6.79)***	0.00 (0.00)	0.04
Δinv_t	−0.04 (3.45)***	0.09 (6.12)***	0.05	−0.02 (1.15)	0.07 (2.67)***	0.05
Δex_t	−0.05 (3.94)***	0.14 (7.59)***	0.09	0.01 (0.95)	0.03 (1.54)	0.04
Δim_t	−0.04 (11.02)***	0.12 (18.41)***	0.08	−0.06 (3.63)***	0.14 (6.27)***	0.08
Δcg_t	0.06 (7.97)***	−0.01 (0.73)	0.05	0.03 (2.23)**	0.01 (0.34)	0.04
Δaid_t	−0.01 (0.16)	0.07 (1.35)	0.06	0.12 (2.91)***	−0.10 (1.93)*	0.02

Significance level: *10%; **5%; ***1%

Notes: t-values in parentheses; γ_1 refers to the long-run growth rates of the macrovariables; γ_{11}^{19yy} measures the shift in long-run growth rates that accompanied the introduction of adjustment programs in 19yy; $\gamma_1+\gamma_{11}$ is the overall long-run growth rate of the macrovariables over the sample period.
*CFA franc zone countries.
Source: Author's analysis based on the data described in Section 2.4.

Table 2.3: Long-run (steady state) growth rates of the macrovariables (countries that experienced shift in long-run trends): *first-best* choice of rank

	Malawi			Mozambique		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1982}}$	$\gamma_1+\gamma_{11}^{1982}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1987}}$	$\gamma_1+\gamma_{11}^{1987}$
Δy_t	0.06 (4.17)***	−0.04 (2.06)**	0.02	−0.01 (0.88)	0.07 (7.40)***	0.06
Δinv_t	0.06 (1.59)	−0.06 (1.28)	0.00	0.05 (4.89)***	0.00 (0.00)	0.05
Δex_t	0.06 (3.76)***	−0.01 (0.38)	0.05	0.00 (0.00)	0.16 (6.90)***	0.16
Δim_t	0.03 (2.17)**	0.02 (1.17)	0.05	−0.02 (1.49)	0.09 (4.07)***	0.07
Δcg_t	0.09 (3.67)***	−0.05 (1.58)	0.04	−0.03 (2.75)***	0.11 (5.46)***	0.08
Δaid_t	0.04 (0.09)	0.02 (0.43)	0.06	0.27 (2.70)***	−0.31 (2.06)**	−0.04
	Rwanda			**Senegal***		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1995}}$	$\gamma_1+\gamma_{11}^{1995}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1994}}$	$\gamma_1+\gamma_{11}^{1994}$
Δy_t	0.03 (3.85)***	0.06 (8.68)***	0.09	0.02 (10.50)***	0.02 (4.77)***	0.04
Δinv_t	0.06 (2.86)***	0.12 (2.47)**	0.18	0.04 (4.14)***	0.05 (2.21)**	0.09
Δex_t	0.04 (2.59)***	0.08 (2.33)**	0.12	0.01 (2.24)**	0.00 (0.00)	0.01
Δim_t	0.06 (3.31)***	0.04 (1.14)	0.10	0.02 (2.79)***	0.03 (2.13)**	0.05
Δcg_t	0.03 (3.33)***	0.04 (2.02)**	0.07	0.02 (2.89)***	0.01 (0.61)	0.03
Δaid_t	0.06 (3.93)***	0.00 (0.00)	0.06	0.04 (2.07)**	−0.02 (0.67)	0.02
	Sudan			**Uganda**		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1992}}$	$\gamma_1+\gamma_{11}^{1992}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1987}}$	$\gamma_1+\gamma_{11}^{1987}$
Δy_t	0.03 (10.39)***	0.04 (5.87)***	0.07	0.00 (0.00)	0.07 (23.10)***	0.07
Δinv_t	0.04 (4.02)***	0.11 (6.26)***	0.15	0.01 (0.30)	0.06 (1.55)	0.07
Δex_t	−0.04 (4.34)***	0.19 (8.16)***	0.15	−0.06 (1.81)*	0.17 (3.40)***	0.11
Δim_t	−0.01 (0.37)	0.11 (2.90)***	0.10	0.00 (0.00)	0.08 (1.51)	0.08
Δcg_t	−0.01 (0.43)	0.09 (2.32)**	0.08	0.01 (1.17)	0.03 (2.14)**	0.04
Δaid_t	0.01 (5.39)***	0.12 (24.20)***	0.13	0.07 (2.23)**	−0.01 (0.23)	0.06

Significance level: *10%; **5%; ***1%

Notes: t-values in parentheses; γ_1 is the long-run growth rate of the variables; γ_{11}^{19yy} measures the shift in long-run growth rates that accompanied the introduction of adjustment programs in 19yy; $\gamma_1+\gamma_{11}^{19yy}$ is the overall long-run growth rate of the variables over the sample period.

*CFA franc zone countries.

Source: Author's analysis based on data described in Section 2.4.

tered for the rest of the countries. The turnaround in export growth was quite dramatic for Ghana, Uganda, and Mozambique, which may reflect the fact that these countries undertook bold trade and exchange rate reforms (World Bank, 1994). Nonetheless, the effects were so large partly because they started from a very low base.

In sum, only 4 countries (Ethiopia, Ghana, Rwanda, and Uganda) enjoyed a sustained increase in GDP, investment, and export growth rates. The traditional (*first-generation*) Fund-Bank adjustment package[14] is associated with growth acceleration only in Ghana and Uganda.[15] This can be deemed disappointing if one takes into account the fact that most non-CFA countries have made significant strides in improving their macroeconomic policies (World Bank, 1994, *pp.* 57). Ghana and Uganda were not only Africa's most aggressive reformers but also tended to be those which had the sharpest economic decline in SSA prior to reforms. This might suggest that the traditional SAPs boosted growth rates only in countries where protracted economic decline preceded reforms.

2.6.1.2 CFA franc zone countries

Most of the CFA franc countries under review posted decent GDP growth prior to reforms. Despite such an outstanding performance, two countries (Cameroon and Cote d'Ivoire) saw their long-run GDP growth rate plummet after adjustment, whereas Cameroon, Senegal, and Burkina Faso experienced an increase in the wake of the 1994 devaluation of the CFA franc. The remaining countries witnessed no change. The poor growth performance of these countries is at best in stark contrast to the 'adjustment with growth' rationale for SAPs.

Two countries (Cameroon and Cote d'Ivoire) suffered a severe decline in long-run growth in investment; Niger and Burkina Faso maintained their *pre*-reform performance; whereas Senegal managed an increase, albeit modest and only in the aftermath of the 1994 devaluation. This is consistent with Devarajan and De Melo (1990, *pp.* 25), who show that expenditure-reduction in general and investment-reduction in

[14]The *first-generation* SAPs were ubiquitous in the 1980s and focused on restoring macroeconomic imbalances with a view to promoting economic growth. They mainly consisted of fiscal and monetary objectives, and as such placed emphasis on *expenditure-reducing* as well as *expenditure-switching* policies. This generation of SAPs were subject to scathing criticisms for their excessive emphasis on demand reduction, resulting in unwarranted contraction of output and declining living standards. The *second-generation* of SAPs (from the late 1980s through the end of the 1990s) included more comprehensive policy packages aimed at structural and institutional reforms. Moreover, they expanded to other sectors such as health, education, and agriculture. Additional components comprised, among others: provision of basic infrastructure, mitigation of the adverse effects of adjustment on the poor, better sequencing of reforms, and efforts to engender a deeper ownership of reforms. The *third-generation* SAPs (since the late 1990s) have poverty reduction as a fundamental objective, along with sustainable growth. In addition, attention has been given to governance, institution building, social sector reforms, as well as to the sustainability and ownership of reforms.

[15]Note that Ethiopia and Rwanda introduced major adjustment policy reforms in the 1990s and pursued 'unorthodox' adjustment path.

particular bore the brunt of adjustment in the CFA franc zone.

Of the 5 CFA countries, two of them (Cameroon and Cote d'Ivoire) experienced curtailment of export growth; Burkina Faso saw an increase; whereas the remaining two countries have had no change. The substantial appreciation of the CFA franc in the 1980s and early 1990s exacerbated the poor export performance in the CFA franc zone (World Bank, 1994). Adjustment to external shocks was nearly impossible because a fixed exchange rate precluded a major policy instrument for halting prolonged recession.

To sum up, adjustment has generally failed to put the CFA franc economies back on their original trend path. The contraction in GDP growth in the CFA zone has been principally engendered by the sharp cutback in investment spending that accompanied the introduction of deflation-orchestrated adjustment. The one-size-fits-all tight fiscal and credit policies prescribed by the Bretton Woods institutions (henceforth BWIs), combined with the inability to use exchange rate devaluation and the inefficacy of alternative instruments, appear to have wrought havoc on the CFA franc economies.

A stocktaking of the adjustment experiences of the 18 SSA countries drives home several points. Growth takeoffs are primarily associated with large increases in export and a relatively less pronounced increases in investment, which is consistent with Jones and Olken (2008). Growth collapses appear to be strongly linked with sharp declines in investment and government consumption spending. Most of the countries which have had growth acceleration experienced considerable terms-of-trade declines for most of the decade after they embarked on reforms (e.g. Uganda, Ghana, Mozambique, Senegal, Ethiopia). This seems in line with the finding in Hausmann *et al.* (2005) and Jones and Olken (2008) that the link between terms-of-trade changes and sustained shifts in growth performance is quite weak. In addition, some of the growth accelerations are preceded by political-regime changes and/or the coming to an end of major wars (e.g. Uganda, Ethiopia, Rwanda, Ghana). Thus, the *post*-reform growth spurt could partly be attributed to political changes, although it is unlikely that the latter alone fully explains the growth turnarounds.

Moreover, for some of the countries (e.g. Ghana and Uganda), the *post*-reform growth resurgence partly reflects the protracted economic decline that preceded reforms. Furthermore, most of the countries that saw growth acceleration experienced considerable depreciation in their real effective exchange rates in the first five years of reforms (Ghana (61%), Ethiopia (28%), Uganda (25%), Sudan (34%), Mozambique (12%), Burkina Faso (11%), Senegal (10%)). The remaining countries (except Tanzania (25%) during 1986-1990) have seen only a relatively modest real exchange rate depreciation in at least the first decade after they embarked on adjustment. The sharp *post*-reform increase in export growth in most of the aforementioned countries might reflect the sizable real exchange rate depreciation. The finding that some countries in the CFA zone experienced growth acceleration after the devaluation of the CFA franc in 1994 seems to reinforce this claim.

2.6.2 Country experiences

2.6.2.1 Non-CFA countries

Ethiopia: Ethiopia launched an economic reform and structural adjustment program (ERSAP) in 1992 (IMF, 1996). The program encompassed wide-ranging macroeconomic stabilization and structural reform measures: liberalization of foreign exchange and trade systems, exchange rate devaluation, promotion of price signals and competition, fiscal and monetary reforms, deregulation of prices, private sector development and initial divestiture of state owned enterprises, and sustainable development measures. The program aimed at achieving, among others, an annual GDP growth rate of 6.0 percent in the medium term (1992 – 1995 and beyond) (AfDB, 1997). As part of the program, the *Birr* was devalued by 143 percent in 1992. Although the program was developed in collaboration with the BWIs, the Government steered the pace and sequencing of the liberalization reforms. In fact, the country was one of the very few African countries that pursued 'unorthodox' adjustment path which involved, *inter alia*, active government intervention. Unlike many SSA countries that were coerced to liberalize at breakneck speed, Ethiopia adopted a conservative stance on liberalization, particularly financial liberalization.

Ethiopia experienced brisk recovery after market-oriented policy reforms were ushered in, as evidenced by the considerable increase in the long-run growth rates of GDP, investment, and exports. The significant increase in GDP growth following the introduction of adjustment programs appears to be due to the sharp increase in export and import growth, notwithstanding the less pronounced increase in average annual growth in investment. It seems surprising that mean annual growth in aid inflows declined in the *post*-reform period. This can be mainly explained by the steady decline in multilateral aid in the 1990s (Gebregziabher, 2014), which was mainly associated with donor policy conditionality (Stiglitz, 2003; Wade, 2001). Since the early 1990s, the BWIs have been imposing policy conditionalities and cutting back on their aid funding when the government failed to comply.

Ghana: Ghana launched orthodox SAPs, locally dubbed *Economic Recovery Program* (ERP), in 1983 (IMF, 1998). Ghana has been touted as a success story and has often been put forward by the BWIs as a country to emulate in Africa. The policy reforms, undertaken at a time when the economy was teetering on the verge of total collapse, were extensive and contained several foci: liberalization of the exchange rate and trade systems; price decontrol and reform; market liberalization; monetary and fiscal management; rehabilitation of the economic and social infrastructure; and structural and institutional reforms to enhance economic efficiency and encourage private saving and investment (Hadjimichael *et al.*, 1996). Currency devaluation was the centerpiece of the ERP and the *Cedi* was devalued by 900 percent in 1983, which was followed by eight other devaluations between 1983 and 1986 (Alderman, 1994). It bears noting that the reform effort in Ghana was backed by a drastic expansion in government expenditure (Rodrik, 1998, Alderman, 1994), atypical of most SSA countries.

Ghana's macroeconomic performance after adjustment has been impressive. GDP growth surged from a mere 1 percent annually during 1966 − 1983 to a hefty positive rate of 5 percent in the *post*-1983 period. The growth rate of exports also edged up considerably, marking a break with the secular descent before reforms. Further, aid inflows to Ghana increased considerably. Aid flows to Ghana increased fivefold between 1983 and 1989 (Alderman, 1994). This may suggest that absent the large inflows of aid the success of the program would have been far more modest. It is, finally, noteworthy that Ghana's spectacular performance partly reflects the depth of the economic decline before adjustment.

Kenya: Kenya's first adjustment period (1980-84) was marked by a half-hearted and haphazard implementation of reform. According to Swamy (1994), the lack of compliance can be traced to design and timing problems but also due to patchy and intermittent commitment to the program. Kenya embarked upon a more concerted and sustained effort at adjustment in 1985 (Husain and Faruqee, 1994). However, even the second period of adjustment (1985-91) was characterized by stop-and-go reforms and, according to Mosley (1991, *pp.* 270), "few country lending experiences have given the Bank so much cause for frustration". Trade liberalization (and exchange rate depreciation) and, to a lesser extent, export development constituted the most successful area of reform during 1985-91 (Swamy, 1994). Note in passing that Kenya's adjustment experience has been a byword for the failure of World Bank policy conditionality. Mosley (1991, *pp.* 300) argues that one would be hard-pressed to name a single policy condition attached to the Bank's adjustment lending that was implemented because of its pressure.

Only the reforms launched in 1985 appear to have been associated with shift in long-run growth rates. The economic performance in the *post*-adjustment period was rather disappointing. The long-run growth rate of GDP remained unchanged. Growth in imports crept up significantly. The surge in imports is consistent with the fact that adjustment focused on the external sector and that there was substantial import liberalization (Mwega and Ndung'u, 2008; Swamy, 1994).

Malawi: Malawi implemented a reform program during 1981-87, which comprised liberalization of agricultural prices and markets; price decontrol; subsidy removal; public expenditure reductions; parastatal reform; tax reform; and public sector institutional and management reform (Harrigan, 1991). The program aimed at reducing domestic absorption, through cutbacks in public sector expenditure, and stimulating tradable production by adjusting price incentives– summed up in what Lipton (1987) referred to as "pricism and state minimalism". The reform package focused single-mindedly on 'getting prices right' and lacked non-price policies to address deep-seated structural constraints and elicit supply response (Sahn and Arulpragasam, 1991, 1994; Harrigan, 1997, 2003). However, growth and development require a lot more than just getting prices right. In the early 1990s, a general consensus was reached regarding the failure of past adjustment reforms (Sahn and Arulpragasam, 1991; World Bank, 1995, 1997). An overwhelming number of studies conclude that the dismal performance in the *post*-reform period can be mainly ascribed to the adverse impact of the Bank's inappropriate policies, the poor sequencing of the reform process, and the failure of the program to tackle fundamental structural bottlenecks (see, among others, Cromwell, 1992; Harrigran, 1991, 1997, 2003; Lele, 1990; Sahn and Arulpragasam, 1991, 1994). Note, however, that a number of exogenous shocks reinforced the lack of economic recovery. However, most of the adverse external shocks were similar to those experienced by other African countries (Lele, 1990). Against this backdrop, a new series of SAPs were introduced in late 1980s and 1990s to salvage the deteriorating economy.[16]

Only the reforms introduced in late 1981 were associated with a shift in the long-run path of the macrovariables. Although the stabilization-cum-structural adjustment policy measures improved external and internal balance (Lele, 1990), their outcome in terms of economic growth appears quite disappointing. Long-run GDP growth stood at 6 percent between 1965 and 1981, but plunged to 2 percent in the *post*-reform period. Besides, the mean annual growth rates of investment and government consumption fell dramatically. Note, however, that the trends underlying investment and government consumption are statistically insignificant perhaps due to the substantial fluctuation in the annual growth rates, which appears to have increased the standard errors of these estimates. However, the long-run growth rates of exports and imports changed only slightly. This seems to indicate that the major cause of the decline in GDP growth was the sharp fall in investment growth. This is to be expected because cutbacks in development expenditure were at the heart of the reduction in domestic absorption in conjunction with the IMF's credit ceilings (Harrigan, 1991, *pp.* 261). Faini *et al.* (1991)

[16] Although the Bank endeavored in the early 1990s to move away from its narrow pricist approach and tackle some of the structural rigidities of the economy, it reverted back to the orthodoxy of the 1980s in late 1990s, which engendered serious inconsistencies in policy implementation (Harrigan, 2003, *pp.* 854).

find that Malawi lost close to 4 percent of GDP per year during $1982 - 1986$ because of lower investment levels under Fund-Bank SAPs. In sum, a look at Malawi's *post*-reform performance suggests that adjustment has failed to jump-start the stagnant economy.[17] Sahn and Arulpragasam (1994, *pp.* 231) corroborate this conclusion. This comes as somewhat of a surprise as Malawi had sound macroeconomic management prior to adjustment (Lele, 1990; Sahn and Arulpragasam, 1994) and was hailed as a relatively strong adjuster (World Bank, 1994).

Mozambique: Mozambique introduced SAPs, known as the *Program for Economic Rehabilitation* (PRE), in 1987 (IMF, 1999). The country has been described by the Bank as one of the few consistent reformers in SSA. The PRE constituted typical SAPs and the BWIs played a pivotal role in the formulation and implementation of the program. Fiscal adjustment, monetary restraint, currency devaluation, price and trade liberalization, financial sector reforms, and privatization of public enterprises were key pillars of the reform strategy. The first steps toward adjustment were two large devaluations of 80 percent and 50 percent in 1987 (Sahn *et al.*, 1998).

The implementation of the ERP appears to have been associated with a reversal of the economic decline before 1987. GDP growth rate swung from an average of -1 percent per annum over the period $1960 - 1987$ to about 7 percent after reforms. The growth rate of exports also picked up dramatically. However, long-run investment growth experienced no change. The surge in exports seems to have underpinned the unprecedented rise in GDP growth.

Rwanda: Rwanda implemented sweeping macroeconomic reforms under the aegis of the BWIs in the early 1990s. Despite the successful implementation of the reforms (World Bank, 1997), the economy continued its downward slide. Part of the economic decline could be explained by the depressing effect of stabilization policies and the protracted civil war in the early 1990s. Following the end of the genocide and the ascension to power of a new government in mid-1994, Rwanda embarked on major structural reforms, which included, among others, trade reforms, liberalization of the monetary and financial regimes, and privatization of state-owned enterprises (IMF, 1998).

The adjustment programs before 1995 appear to have had no effect on the long-run movements of the macrovariables and thus we only accounted for the shift in growth rates associated with the reforms in mid-1990s. Rwanda witnessed remarkable economic performance in the *post*-reform era. GDP growth, which averaged 3 percent between 1960 and 1994, hovered around 9 percent in the *post*-reform period. The

[17] This is not, however, to say that adjustment lending has yielded no positive economic benefits at all, which cannot be quantified without undertaking a counterfactual analysis.

post-adjustment period also saw a considerable rise in investment and export growth rates.[18]

Sudan: Sudan adopted orthodox SAPs during 1978-85, which failed to attain its objectives (World Bank, 1990). The Government initiated a second wave of radical macroeconomic reforms in 1992 (Ahmed *et al.*, 2004), which included, *inter alia*, price deregulation, trade liberalization, substantial exchange rate devaluation, and massive privatization and private sector revitalization. Note that these policy reforms were 'homegrown' in the sense that they were pursued without negotiation with and external financing from the BWIs. However, the self-imposed SAPs were, on the surface, nothing but a carbon copy of, if not more stringent than, the orthodox SAPs that were imposed on other African countries.

Sudan's turnaround since 1992 is nothing short of remarkable. An impressive GDP growth rate of 7 percent per annum marked the *post*-reform period. Moreover, the country saw a strong resurgence in export and investment growth. Sudan's adjustment experience might suggest that SAPs are more likely to yield positive and sustainable results when adjusters take the driver's seat than when introduced under 'coercion' or 'incentive' of financial assistance.

Uganda: Uganda embarked on comprehensive SAPs in 1987, which has been heralded as one of the few success stories in Africa (Hadjimichael *et al.*, 1996). Exchange rate reforms, trade liberalization, price deregulation, financial sector reform, public enterprise reform, and civil service reform were key pillars of the reform process. Restoring exports was the overriding objective of the program (Botchwey *et al.*, 1998). The devaluation of the Ugandan Shilling by 77 percent in 1987 was one of the cornerstones of the program. Uganda has been lauded for its high degree of 'ownership' of reforms. A massive increase in government expenditure characterized the reform process (Botchwey *et al.*, 1998).

Uganda staged a relatively strong resurgence in economic performance since 1987. GDP growth averaged only 0 percent per annum in the years through 1987, but it spiked to 7 percent after adjustment programs were instituted. The mean annual growth rate of exports also witnessed a sharp increase. This might suggest that the growth spurt after adjustment is primarily related to the strong surge in the trade sector, notwithstanding the contribution from investment. This is in line with the fact that full-scale reforms in the export sector and rapid import liberalization characterized reform in Uganda (Rodrik, 1998). However, the surge in growth rates partly reflects the sharp economic decline that preceded it.

[18] Note, however, that the extraordinary growth rates in mid-1990s were partly due to the coming to an end of the protracted civil war.

2.6.2.2 CFA zone countries

Burkina Faso: Burkina Faso began implementing broad-ranging SAPs, aimed at putting the economy on a faster growth track, in 1991 (IMF, 1999). Unlike most SSA countries which introduced SAPs in response to emergency situations, Burkina Faso had no severe macroeconomic imbalances. Moreover, the country introduced several financial and structural measures in the wake of the devaluation of the CFA franc in 1994. These programs were successfully implemented (IMF, 1999; World Bank, 1994).

The policy reforms introduced following the 1994 devaluation appear to have been linked with a sustained increase in mean macroeconomic growth rates. Annual average GDP growth rate stood at around 3% prior to the introduction of adjustment programs and increased to about 5% in the *post*-1994 period. The country also saw an increase in export growth. However, long-run investment growth remained almost unchanged. This seems to suggest that the growth acceleration can be mainly attributed to the increase in exports, which is in turn likely to have been due to the real exchange rate depreciation.

Cameroon: Cameroon launched an economic reform program in 1988. Cameroon's membership in the franc zone rendered the use of exchange rate devaluation impossible. Besides, the use of other expenditure-switching instruments was circumscribed by the structure of Cameroon's trade. Thus, adjustment was primarily expenditure-reducing and fiscal policy reforms, which slashed current expenditure by a third and halved public investment, were hallmarks of the process of adjustment (Blandford *et al.*, 1994). The austerity program was a *sine qua non* of the IMF standby credit in 1988 and a World Bank SAP in 1989. Reform in Cameroon was characterized by greater austerity than in other SSA countries (Sahn *et al.*, 1998). Moreover, Cameroon introduced macroeconomic and structural reforms in 1994 to reverse its downward economic slide and capitalize on the competitiveness gains from the devaluation of the CFA franc (IMF, 1997). The program was, however, suspended for a year on account of its below-target performance, although it was resumed in 1995 under a new standby arrangement with the IMF (AfDB, 2002).

The estimates in Table 2.2 indicate that the adjustment programs instituted to relieve the country of its economic plight turned out to be "harsh medicine for one of the best-performing and presumably best-managed economies in SSA" (Blandford *et al.*, 1994, *pp.* 161). The country posted an average GDP growth rate of 7% before 1988, which plunged to −3% between 1988 and 1994. GDP growth bounced back after 1994, albeit only slightly. Average investment growth fell dramatically between 1988 and 1994, whereas it increased in the *post*-1994 period. However, the long-run growth rate of exports remained almost unchanged during 1988-94. The sharp fall in GDP growth during 1988-94 appears to have been primarily induced by the draconian cuts in investment expenditure. This does not come as much of a surprise as the program was geared toward curbing government investment expenditure. Government-financed capital outlays shrank by 60% in 1988 and another 38% in 1989 (Blandford *et al.*, 1994). Moreover, private investment plunged 40% between 1986 and 1988, because of the contraction in aggregate demand induced by the reduction in government spending (*Ibid.*). The contraction in investment was so severe that in the early 1990s it fell to 17% of its 1986 value in nominal terms. The key bottleneck to economic recovery before 1994 was the continued overvaluation of the exchange rate. World Bank (1995, *pp.* vi) notes that "that the design of the structural adjustment lending (SAL) was flawed, with the benefit of hindsight. It aimed at reestablishing the competitiveness of the economy through deflationary, internal policies alone which, without an exchange rate adjustment, proved unrealistic."

Cote d'Ivoire: Cote d'Ivoire's first SAPs dates from 1981 and had a life of three years. This was followed by five standby arrangements and three structural adjustment loans (SALs) with the IMF and World Bank respectively until the early 1990s. The program largely centered on macroeconomic and financial management, administrative reforms, and the restructuring of public enterprises. The program came along with a string of conditions that coerced the government to cut expenditures drastically. In the first three years of stabilization, public investment was cut to 33 percent of its value in 1980 in nominal terms (Kanbur, 1990, *pp.* 33). Cote d'Ivoire may be regarded a model pupil of the IMF in the sense that the full range of measures for compressing demand envisaged by the IMF program were implemented and at times surpassed (Duruflé, 1989). The sharp contraction in aggregate demand was compounded by the tighter monetary policies required to enforce the anti-inflationary policies of the CFA franc zone. The failure of earlier adjustment efforts instigated a second wave of reforms in the late 1980s, which included, among others, measures that reduced capital and current expenditures by 15 percent and 30 percent respectively (IMF, 1998). Furthermore, in the wake of the devaluation of the CFA franc in 1994, the government successfully implemented a comprehensive adjustment program (IMF, 1998).

While economic performance in the *pre*-1981 period was nothing short of remarkable, often referred to as the 'Ivorian miracle' (Kanbur, 1990), the *post*-adjustment period recorded discouraging indices of performance. GDP growth rate averaged 8 percent before adjustment, but collapsed to 2 percent in the *post*-adjustment period. This may be quite surprising because one would expect even incomplete adjustment to lead to macroeconomic improvement as argued in World Bank (1994). Moreover, the mean growth rates of investment and government consumption experienced a severe decline. This seems to indicate that the strong decline in GDP growth was mainly caused by the drastic decline in investment and government consumption spending given that the fall in export growth was small.[19]

Senegal: Senegal adopted a number of adjustment programs in the first half of the 1980s (Hadjimichael *et al.*, 1996). The country strengthened adjustment efforts during 1985-88 and undertook bold measures to stabilize and liberalize the economy. Despite their successful implementation, the reforms failed to kick-start the stagnant economy. Further, faced with economic stagnation, the government undertook a bold and ambitious economic reform program in 1994, in tandem with the devaluation of the CFA franc, with the view to invigorating rapid growth (IMF, 1999).

Despite a decade-long adjustment, our empirical analysis indicated that the adjustment reforms instituted before 1994 exerted no effect on the long-run path of the macrovariables and thus we only accounted for the small shift in growth rates that followed the 1994 reforms. The mean annual growth rates of GDP and investment experienced a moderate increase. Moreover, import growth picked up slightly. However, long-run export growth remained unaffected.

2.6.3 Countries with no shift in long-run trends

Gambia: Gambia introduced far-reaching SAPs in 1985/86, followed by many others. Gambia's reform program is generally considered to be one of the most successful ones carried out among African countries (World Bank, 1994, *pp.* 57; Rodrik, 1998). However, the reform program failed to raise the economy to a higher growth path. Part of the reason could be the low growth potential of the country because it has had poor human and physical resources. Not least important is, however, the fact that no protracted economic decline preceded the reforms and thus Gambia, unlike Ghana and Uganda, has not had the benefit of resurgence in economic activity (Rodrik, 1998).

Guinea: Guinea commenced sweeping economic reforms in 1986, the breadth and speed of which was truly impressive (Arulpragasam and Sahn, 1994; IMF, 1999; World Bank, 1994). Compared to most SSA countries, reform in Guinea constituted a more remarkable break from the past (Arulpragasam and Sahn, 1994; World Bank, 1994). Despite the rigorous pursuit of reform, adjustment failed to push the economy to a path of higher growth. Guinea's lack of economic recovery following reforms is reminiscent of the limitations of the traditional Fund-Bank adjustment package (Arulpragasam and Sahn, 1994), namely single-minded faith in the magic of 'getting prices right'.

Madagascar: Madagascar began implementing SAPs in 1981. The program

[19] The effects of the financial stabilization measures taken in 1980/81 were reflected in the nearly stagnant real GDP in 1981/82 and the contraction of -2.5% over 1982/83 and 1983/84 (Kanbur, 1990, *pp.* 33).

sought to correct chronic imbalances in public finance, balance of payments, and monetary situation. Although a stable financial situation was achieved, the structural reforms crafted to restore growth have met with meager success until late 1980s (Dorosh and Bernier, 1994). Hence, the country undertook further economic reforms involving radical changes in trade and exchange rate policies in the late 1980s.

Our analysis shows that, despite more-than-a-decade-long adjustment, the Malagasy economy stayed on its *pre*-reform growth track. GDP growth averaged about 2 percent over the sample period, which appears too small to make any dent on poverty and is quite low even by African standards.

Niger (CFA): Niger launched SAPs in 1983, which was followed by additional loan agreements with the IMF during 1984-86 and with the World Bank in 1986-89 (Dorosh, 1994). Fiscal policy reforms, that cut government capital expenditure by more than 40% in real terms between 1983 and 1986, constituted the heart of the program. Further, Niger successfully implemented macroeconomic and structural reforms (World Bank, 2001, *pp.* 5). In spite of years of adjustment and the considerable progress achieved in implementing the reforms (World Bank, 1995, *pp.* 2), the long-run movements of the macrovariables remained unaffected. GDP growth hovered around 3% over the period 1965-2009.

Tanzania: Tanzania embarked on SAPs in earnest in 1986. Although the program was more extensive than earlier half-hearted adjustment efforts, it could at best be described as gradual as the pace of reform was relatively slow (Sahn *et al.*, 1998). In the first half of the 1990s, the country demonstrated a much lower commitment to reform, which jeopardized the achievements of the late 1980s (Sahn *et al.*, 1998). Since the mid-1990s, Tanzania has been pursuing substantial structural and institutional reforms. The country has had sustained increase in GDP growth rate since the mid-2000s, which, according to some recent studies (World Bank, 2008a), can be partly attributed to the rigorous pursuit of reform since mid-1990s. This is, however, beyond the scope of this paper and requires further analysis.

Zambia: Zambia launched its first SAPs in 1983, which was short-lived and ineffective. A new government assumed power in late 1991 and, unlike the stop-and-go nature of the reforms in the 1980s, introduced far-reaching reforms. Since the early 1990s, Zambia has been implementing extensive and sweeping macroeconomic, structural, and institutional reforms (IMF, 1999, World Bank, 2004). However, the economy remained on its pre-reform growth path. This is in line with World Bank (2004, *pp.* 1) which finds that "despite these reforms, real GDP grew at an average annual rate of just 1.3% between 1992 and 2003. The reasons for such disappointing performance in the face of sweeping policy reforms are not well understood."

2.7 Concluding remarks

In this paper, we investigated the link between IMF-World Bank stabilization-cum-structural adjustment programs (SAPs) and long-run economic performance in 18 Sub-Saharan African (SSA) countries on a country-specific basis for the period 1960 to 2009. We find that only few SSA countries have experienced a sustained increase GDP, export, and investment growth rates. Most of these countries had considerable real exchange rate depreciation, while many of them suffered sizable terms-of-trade declines. The traditional (*first-generation*) IMF-World Bank adjustment package is associated with resurgence of growth in GDP, export, and investment only in two countries (Ghana and Uganda). This stands in stark contrast to the 'adjustment with growth' rationale for adjustment programs.

The dismal growth performance of many African countries, despite having successfully implemented more-than-a-decade-long Fund-Bank structural adjustment programs, is reminiscent of the limitations of the traditional adjustment package, i.e. the single-minded focus on 'getting prices right' and thus the neglect of growth-enhancing policy measures. Payoffs in terms of investment growth were particularly disappointing for many countries. This might suggest that, although fiscal and monetary retrenchment were required for the success of reforms, more public investment in, *inter alia*, physical and human capital would have helped countries move onto a higher growth trajectory (World Bank, 2005, 2008b). Notwithstanding this, some countries enjoyed an increase in long-run export and import growth rates, which is consistent with the fact that many countries undertook bold trade and exchange rate reforms (World Bank, 1994).

Taken as a whole, countries in the CFA franc zone fared much worse than their non-CFA counterparts due to the different adjustment strategies pursued, although some of them experienced growth acceleration in the wake of the 1994 devaluation of the CFA franc. This might suggest that adjustment has done more harm than good when unaccompanied by real exchange rate devaluation. The sharp fall in long-run GDP growth in the CFA zone appears to have been mainly induced by the draconian cutbacks in investment spending necessitated by Fund-Bank program conditionality coupled with the inability to use exchange rate as an instrument of adjustment.

Appendix

Table 2.4: Long-run (steady state) growth rates of the macrovariables (countries that experienced shift in long-run trends): *Second-best* choice of rank

	Burkina Faso*			Cameroon*			
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1994}}$	$\gamma_1+\gamma_{11}^{1994}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1988}}$	$\boldsymbol{\gamma_{11}^{1995}}$	$\gamma_1+\gamma_{11}^{88}+\gamma_{11}^{95}$
Δy_t	0.03 (10.34)***	0.02 (4.2)***	0.05	0.07 (11.17)***	−0.09 (7.85)***	0.06 (4.64)***	0.04
Δinv_t	0.08 (3.86)***	−0.02 (1.89)*	0.06	0.09 (4.61)***	−0.14 (3.33)***	0.12 (2.65)***	0.07
Δex_t	0.05 (2.67)***	0.03 (0.94)	0.08	0.05 (1.80)*	0.02 (0.34)	−0.03 (0.48)	0.04
Δim_t	0.04 (2.87)***	0.01 (0.41)	0.05	0.05 (2.04)**	−0.03 (0.56)	0.04 (0.95)	0.06
Δcg_t	0.06 (3.80)***	−0.01 (0.31)	0.05	0.08 (3.96)***	−0.11 (2.73)***	0.08 (2.30)**	0.05
Δaid_t	0.06 (3.96)***	−0.01 (0.027)	0.05	0.00 (0.00)	0.07 (1.07)	−0.03 (0.45)	0.04

	Côte d'Ivoire*			Ethiopia		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1981}}$	$\gamma_1+\gamma_{11}^{1981}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1992}}$	$\gamma_1+\gamma_{11}^{1992}$
Δy_t	0.08 (11.11)***	−0.07 (6.97)***	0.01	0.02 (3.96)***	0.04 (5.06)***	0.06
Δinv_t	0.12 (3.44)***	−0.14 (2.99)***	−0.02	0.02 (1.73)*	0.07 (3.20)***	0.09
Δex_t	0.06 (5.14)***	−0.02 (1.71)*	0.04	0.02 (1.74)*	0.10 (3.94)***	0.12
Δim_t	0.10 (5.68)***	−0.07 (3.13)***	0.03	0.04 (2.68)***	0.07 (2.63)***	0.11
Δcg_t	0.12 (12.81)***	−0.12 (9.31)***	0.00	0.05 (2.78)***	0.05 (1.71)*	0.10
Δaid_t	0.07 (2.70)***	−0.09 (2.47)**	−0.02	0.08 (3.19)***	−0.01 (0.28)	0.07

	Ghana			Kenya		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1983}}$	$\gamma_1+\gamma_{11}^{1983}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1985}}$	$\gamma_1+\gamma_{11}^{1985}$
Δy_t	0.01 (6.02)***	0.04 (25.79)***	0.05	0.04 (6.79)***	0.00 (0.00)	0.04
Δinv_t	−0.04 (4.03)***	0.10 (7.29)***	0.06	−0.01 (0.46)	0.07 (1.83)*	0.06
Δex_t	−0.05 (4.80)***	0.14 (9.19)***	0.09	0.02 (1.21)	0.03 (1.62)	0.05
Δim_t	−0.04 (9.45)***	0.12 (17.03)***	0.08	−0.06 (2.76)***	0.15 (5.00)***	0.09
Δcg_t	0.06 (8.82)***	−0.01 (1.25)	0.05	0.04 (1.99)**	0.02 (0.67)	0.06
Δaid_t	0.01 (0.71)	0.04 (2.62)***	0.05	0.12 (2.72)***	−0.13 (2.11)**	−0.01

Significance level: *10%; **5%; ***1%

Notes: t-values in parentheses; γ_1 is the long-run growth rate of the macrovariables; γ_{11}^{19yy} measures the shift in long-run growth rates that accompanied the introduction of adjustment programs in 19yy; $\gamma_1+\gamma_{11}^{19yy}$ is the overall long-run growth rate of the macrovariables over the sample period. *CFA franc zone countries.

Source: Author's analysis based on the data described in Section 2.4.

Table 2.5: Long-run (steady state) growth rates of macrovariables (countries that experienced shift in long-run trends): *Second-best* choice of rank

	Malawi			**Mozambique**		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1982}}$	$\gamma_1+\gamma_{11}^{1982}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1987}}$	$\gamma_1+\gamma_{11}^{1987}$
Δy_t	0.06 (3.74)***	−0.04 (1.72)*	0.02	−0.01 (1.03)	0.07 (7.16)***	0.06
Δinv_t	0.06 (1.58)	−0.07 (1.33)	−0.01	0.05 (5.78)***	0.00 (0.00)	0.05
Δex_t	0.05 (3.56)***	0.01 (0.38)	0.06	0.01 (0.86)	0.13 (7.63)***	0.14
Δim_t	0.04 (0.95)	0.01 (0.21)	0.05	−0.01 (1.09)	0.07 (3.59)***	0.06
Δcg_t	0.08 (2.67)***	−0.03 (0.07)	0.05	−0.03 (2.46)**	0.01 (0.62)	−0.02
Δaid_t	0.04 (1.08)	0.02 (0.38)	0.06	0.19 (13.81)***	−0.15 (4.19)***	0.04
	Rwanda			**Senegal***		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1995}}$	$\gamma_1+\gamma_{11}^{1995}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1994}}$	$\gamma_1+\gamma_{11}^{1994}$
Δy_t	0.03 (8.49)***	0.06 (7.18)***	0.09	0.02 (6.79)***	0.02 (4.18)***	0.04
Δinv_t	0.05 (2.65)***	0.12 (2.50)**	0.17	0.04 (2.57)**	0.05 (1.97)**	0.09
Δex_t	0.04 (3.35)***	0.08 (2.55)**	0.12	0.01 (2.36)**	0.00 (0.00)	0.01
Δim_t	0.05 (3.07)***	0.05 (1.40)	0.10	0.02 (2.79)***	0.03 (2.22)**	0.05
Δcg_t	0.04 (5.34)***	0.04 (2.18)**	0.08	0.02 (2.41)**	0.01 (0.82)	0.03
Δaid_t	0.06 (5.17)***	0.00 (0.00)	0.06	0.03 (1.73)*	−0.02 (0.47)	0.01
	Sudan			**Uganda**		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1992}}$	$\gamma_1+\gamma_{11}^{1992}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1987}}$	$\gamma_1+\gamma_{11}^{1987}$
Δy_t	0.03 (6.84)***	0.04 (4.67)***	0.07	0.00 (0.00)	0.07 (23.33)***	0.07
Δinv_t	0.03 (1.90)*	0.13 (4.10)***	0.16	0.01 (0.22)	0.07 (1.85)*	0.08
Δex_t	−0.05 (4.51)***	0.20 (8.36)***	0.15	−0.06 (3.11)***	0.19 (7.12)***	0.13
Δim_t	−0.01 (0.34)	0.12 (2.68)***	0.11	0.01 (0.31)	0.08 (3.12)***	0.09
Δcg_t	−0.01 (2.04)**	0.10 (8.86)***	0.09	0.01 (1.22)	0.03 (2.20)**	0.04
Δaid_t	−0.01 (0.17)	0.17 (1.20)	0.16	0.06 (2.03)**	−0.02 (0.37)	0.04

Significance level: *10%; **5%; ***1%

Notes: t-values in parentheses; γ_1 is the long-run growth rate of the variables; γ_{11}^{19yy} measures the shift in long-run growth rates that accompanied the introduction of adjustment programs in 19yy; $\gamma_1+\gamma_{11}^{19yy}$ is the overall long-run growth rate of the variables over the sample period. *CFA franc zone countries.

Source: Author's analysis based on data described in Section 2.4.

Table 2.6: Long-run (steady state) growth rates of the macrovariables: Data from *World Development Indicators*

	Burkina Faso⁎			**Cameroon**⁎			
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1994}}$	$\gamma_1+\gamma_{11}^{1994}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1988}}$	$\boldsymbol{\gamma_{11}^{1995}}$	$\gamma_1+\gamma_{11}^{88}+\gamma_{11}^{95}$
Δy_t	0.03 (29.06)***	0.02 (8.67)***	0.056	0.07 (12.06)***	−0.09 (9.38)***	0.06 (5.07)***	0.04
Δinv_t	0.06 (5.53)***	0.03 (0.08)	0.09	0.10 (6.41)***	−0.18 (5.63)***	0.15 (4.27)***	0.07
Δex_t	0.04 (1.91)*	0.01 (0.37)	0.05	0.09 (8.16)***	−0.01 (1.59)	−0.01 (0.32)	0.07
Δim_t	0.03 (2.48)***	0.02 (0.97)	0.05	0.08 (24.2)***	−0.12 (10.08)***	0.10 (8.15)***	0.06
Δcg_t	0.06 (10.83)***	−0.02 (1.38)	0.04	0.07 (10.77)***	−0.10 (6.36)***	0.06 (4.17)***	0.03
Δaid_t	0.06 (4.62)***	−0.01 (0.54)	0.05	0.00 (0.00)	0.04 (0.67)	0.02 (0.26)	0.06

	Kenya			**Rwanda**		
Var.	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_{11}^{1985}}$	$\gamma_1+\gamma_{11}^{1985}$	γ	$\boldsymbol{\gamma_{11}^{1995}}$	$\gamma_1+\gamma_{11}^{1995}$
Δy_t	0.05 (12.25)***	−0.02 (3.57)***	0.03	0.03 (4.28)***	0.07 (8.68)***	0.10
Δinv_t	−0.03 (3.43)***	0.08 (5.93)***	0.05	0.05 (3.00)***	0.13 (3.25)***	0.18
Δex_t	0.01 (0.99)	0.04 (3.33)***	0.05	0.03 (2.54)**	0.15 (5.00)***	0.18
Δim_t	−0.04 (3.38)***	0.12 (6.67)***	0.08	0.08 (7.60)***	0.04 (1.59)	0.12
Δcg_t	0.05 (5.00)***	−0.01 (1.00)	0.04	0.06 (3.64)***	−0.01 (0.34)	0.05
Δaid_t	0.09 (3.00)***	−0.06 (1.30)	0.03	0.07 (6.50)***	−0.02 (1.13)	0.05

	Senegal⁎			**Madagascar**	**Gambia**
Var.	γ	$\boldsymbol{\gamma_{11}^{1994}}$	$\gamma_1+\gamma_{11}^{1994}$	$\boldsymbol{\gamma_1}$	$\boldsymbol{\gamma_1}$
Δy_t	0.02 (10.53)***	0.02 (6.00)***	0.04	0.02 (3.19)***	0.04 (13.33)***
Δinv_t	0.04 (4.14)***	0.07 (3.98)***	0.11	0.05 (1.84)*	0.07 (2.59)***
Δex_t	0.00 (0.00)	0.03 (1.27)	0.03	0.015 (1.33)	0.02 (2.50)**
Δim_t	0.01 (0.24)	0.05 (3.13)***	0.06	0.02 (1.11)	0.02 (1.25)
Δcg_t	0.01 (0.78)	0.00 (0.00)	0.01	0.02 (2.25)**	0.04 (1.54)
Δaid_t	0.02 (0.76)	−0.01 (0.30)	0.01	0.03 (1.53)	0.05 (1.17)

Significance level: *10%; **5%; ***1%

Notes: t-values in parentheses; γ_1 is the long-run growth rates of the variables; γ_{11}^{19yy} measures the shift in long-run growth rates that accompanied the introduction of adjustment programs in 19yy; $\gamma_1+\gamma_{11}^{19yy}$ is the over-all long-run growth rate of the variables over the sample period. ⁎CFA franc zone countries.

Source: Author's analysis based on the data described in Section 2.4.

Table 2.7: Long-run (steady state) growth rates of the macrovariables (γ_1) (countries with no shift in long-run trends): *first* and *second-best* choices of rank

	Gambia		**Guinea**		**Madagascar**		**Niger***	
	γ_1		γ_1		γ_1		γ_1	
Var.	$r^* = 3$	$r = 4$	$r^* = 4$	$r = 3$	$r^* = 2$	$r = 3$	$r^* = 4$	$r = 3$
Δy_t	0.03 (10.36)***	0.03 (12.73)***	0.02 (7.93)***	0.02 (0.00)	0.02 (3.03)***	0.02 (8.90)***	0.03 (5.39)***	0.03 (5.19)***
Δinv_t	0.06 (3.23)***	0.07 (15.00)***	0.06 (5.46)***	0.06 (4.87)***	0.04 (1.68)*	0.04 (1.71)*	0.04 (2.48)**	0.03 (2.14)**
Δex_t	0.05 (2.75)***	0.06 (3.79)***	0.03 (9.03)***	0.03 (10.39)***	0.01 (0.97)	0.01 (1.18)	0.02 (5.41)***	0.02 (3.50)***
Δim_t	0.04 (7.09)***	0.03 (5.86)***	0.05 (7.62)***	0.05 (6.81)***	0.02 (0.99)	0.02 (1.06)	0.02 (2.64)***	0.02 (2.31)**
Δcg_t	0.00 (0.00)	0.00 (0.00)	0.03 (7.39)***	0.03 (7.50)***	0.01 (1.84)*	0.01 (1.45)	0.02 (2.22)**	0.02 (1.61)
Δaid_t	−0.03 (0.69)	−0.04 (0.83)	0.02 (0.66)	0.01 (0.26)	0.02 (0.80)	0.02 (0.66)	0.00 (0.00)	−0.01 (0.21)

	Tanzania		**Zambia**	
	γ_1		γ_1	
Var.	$r^* = 4$	$r = 3$	$r^* = 4$	$r = 3$
Δy_t	0.045 (9.62)***	0.045 (9.62)***	0.02 (6.05)***	0.02 (4.04)***
Δinv_t	0.08 (3.95)***	0.08 (4.21)***	0.03 (0.01)	0.03 (1.57)
Δex_t	0.04 (1.79)*	0.04 (2.18)**	0.04 (0.02)	0.04 (1.84)*
Δim_t	0.05 (2.46)**	0.05 (3.05)***	0.04 (0.02)	0.04 (1.46)
Δcg_t	0.04 (1.92)*	0.05 (2.15)**	0.03 (0.01)	0.03 (2.42)**
Δaid_t	0.07 (3.11)***	0.06 (3.37)***	0.05 (0.02)	0.05 (1.88)*

Significance level: *10%; **5%; ***1%

Note: t-values in parentheses; r^* is the *first-best* choice of rank; γ_1 is the overall long-run growth rate of the macrovariables over the sample period; *CFA franc zone countries.

Source: Author's analysis based on the data described in Section 2.4.

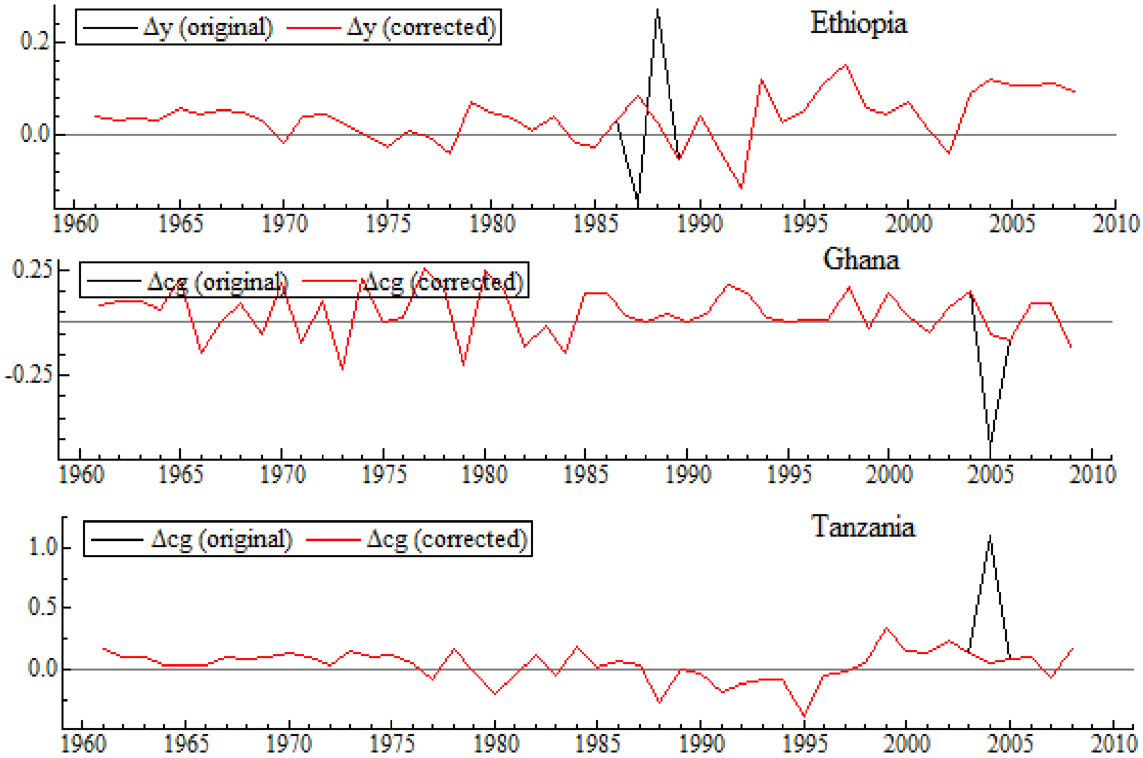

Figure 2.1: Additive outliers (Ethiopia, Ghana, and Tanzania)

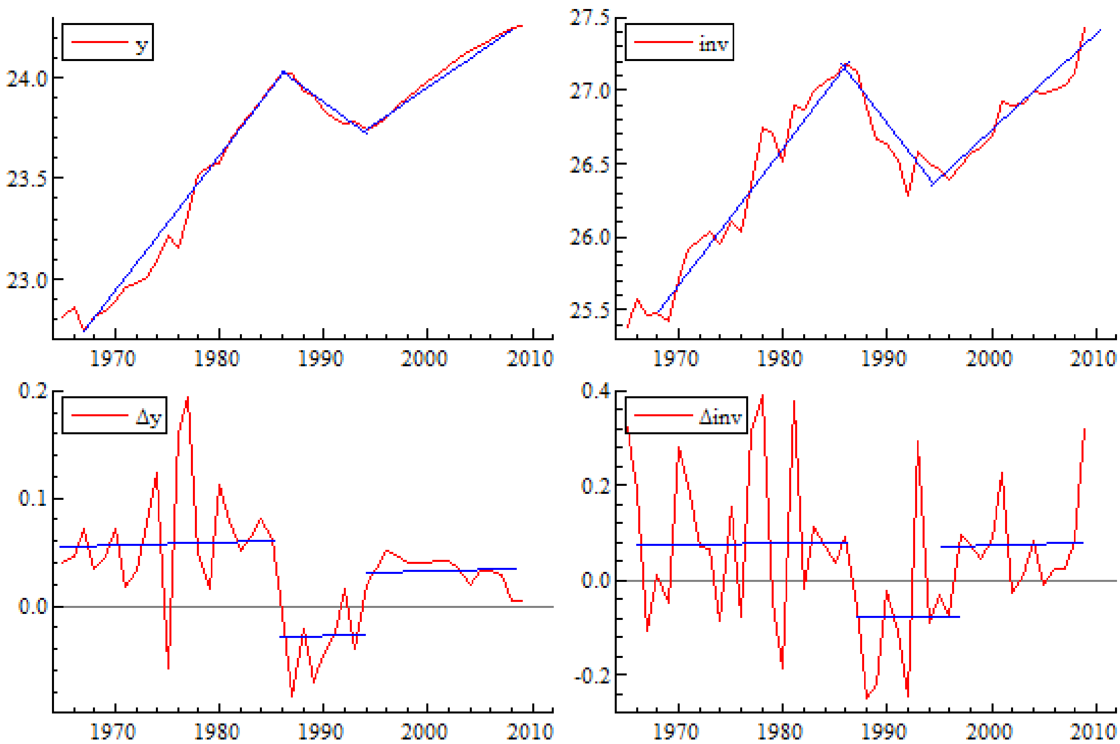

Figure 2.2: Cameroon (GDP and investment in levels and growth rates)

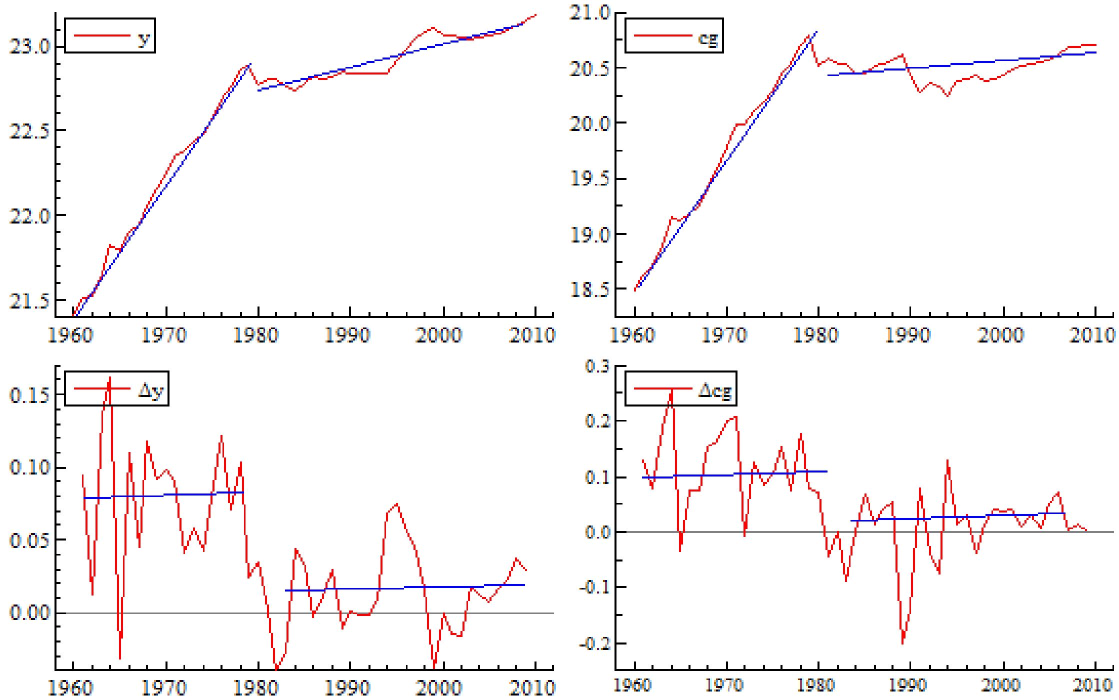

Figure 2.3: Cote d'Ivoire (GDP and government consumption in levels and growth rates)

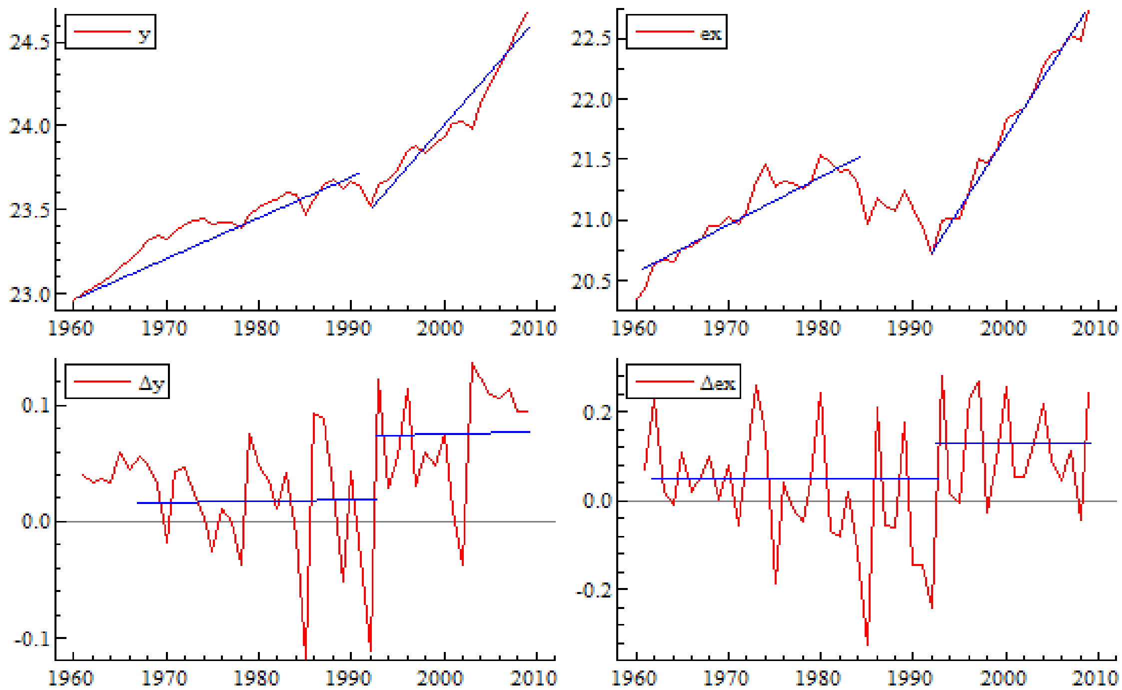

Figure 2.4: Ethiopia (GDP and export in levels and growth rates)

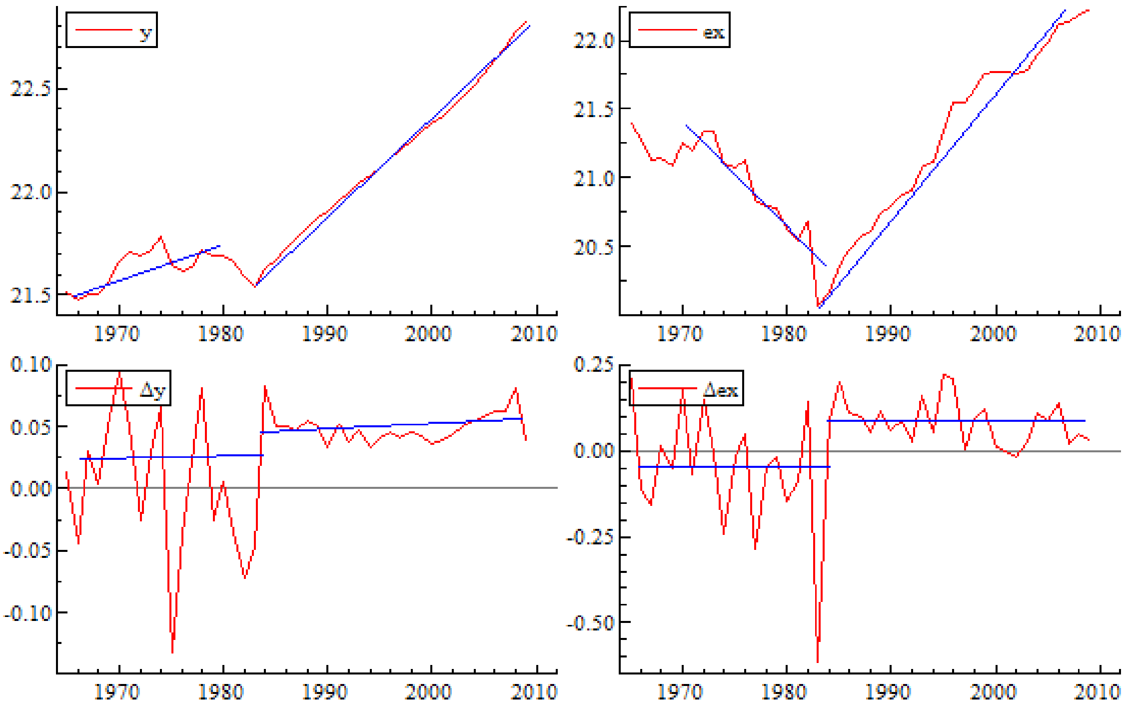

Figure 2.5: Ghana (GDP and export in levels and growth rates)

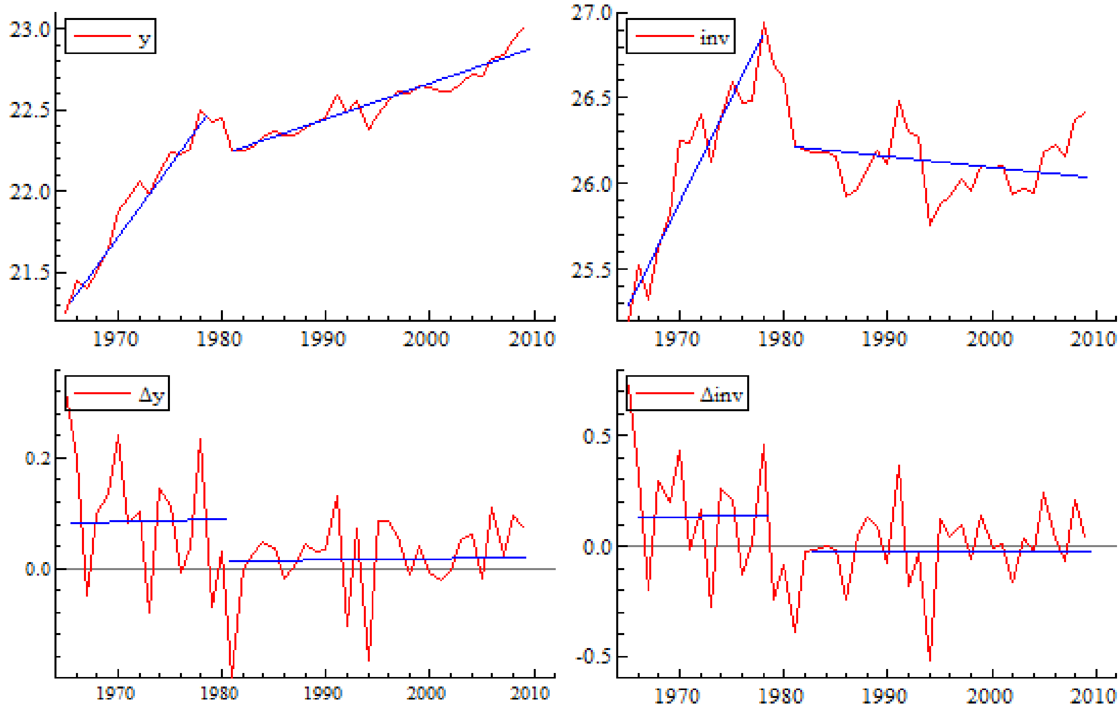

Figure 2.6: Malawi (GDP and investment in levels and growth rates)

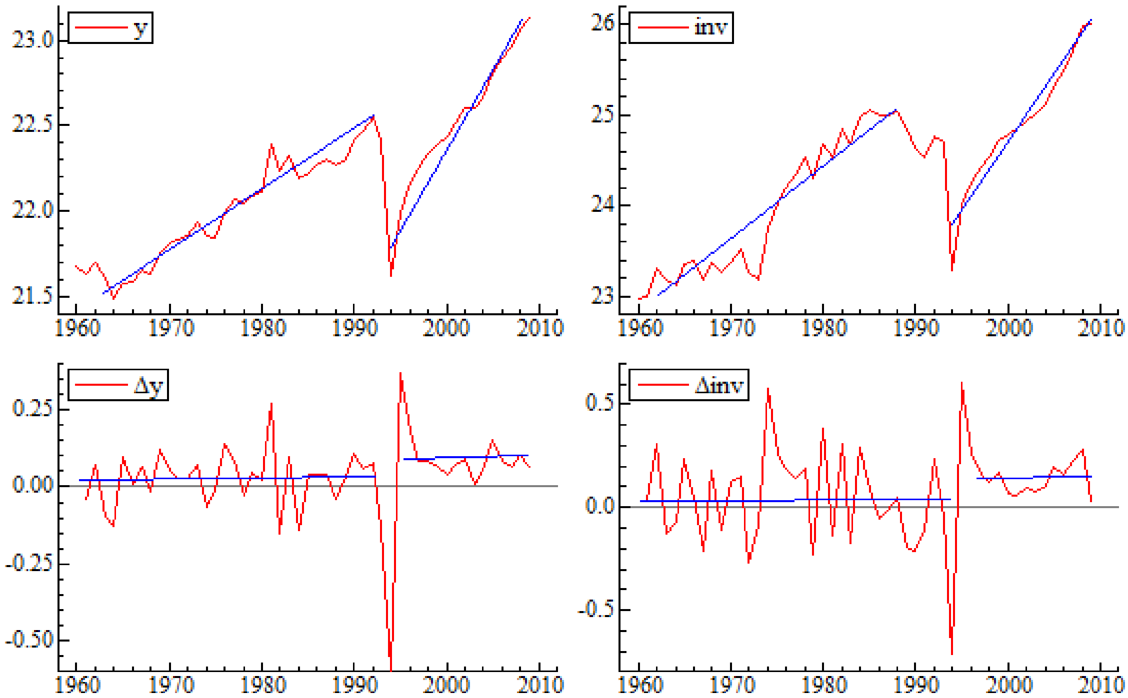

Figure 2.7: Rwanda (GDP and investment in levels and growth rates)

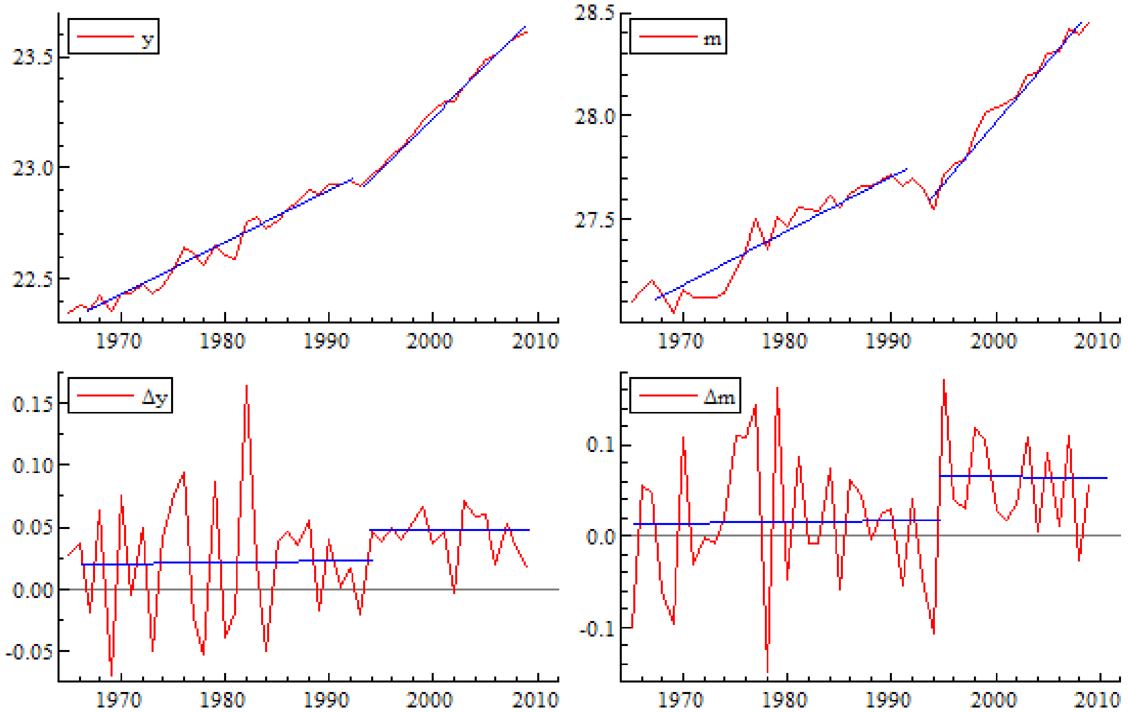

Figure 2.8: Senegal (GDP and import in levels and growth rates)

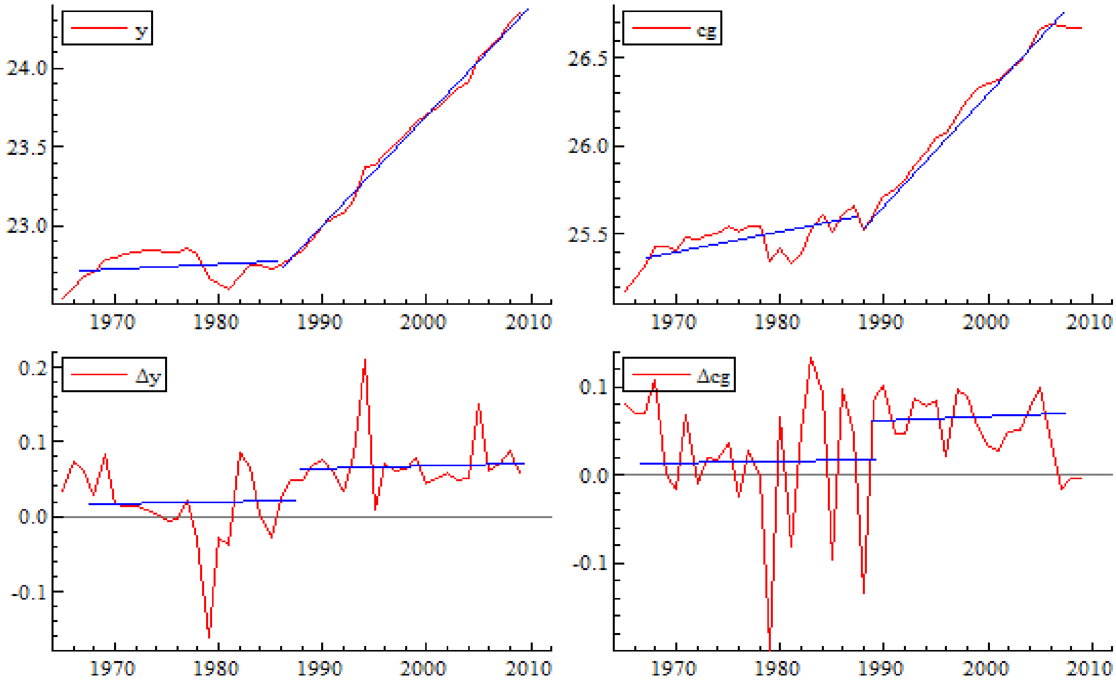

Figure 2.9: Uganda (GDP and government consumption in levels and growth rates)

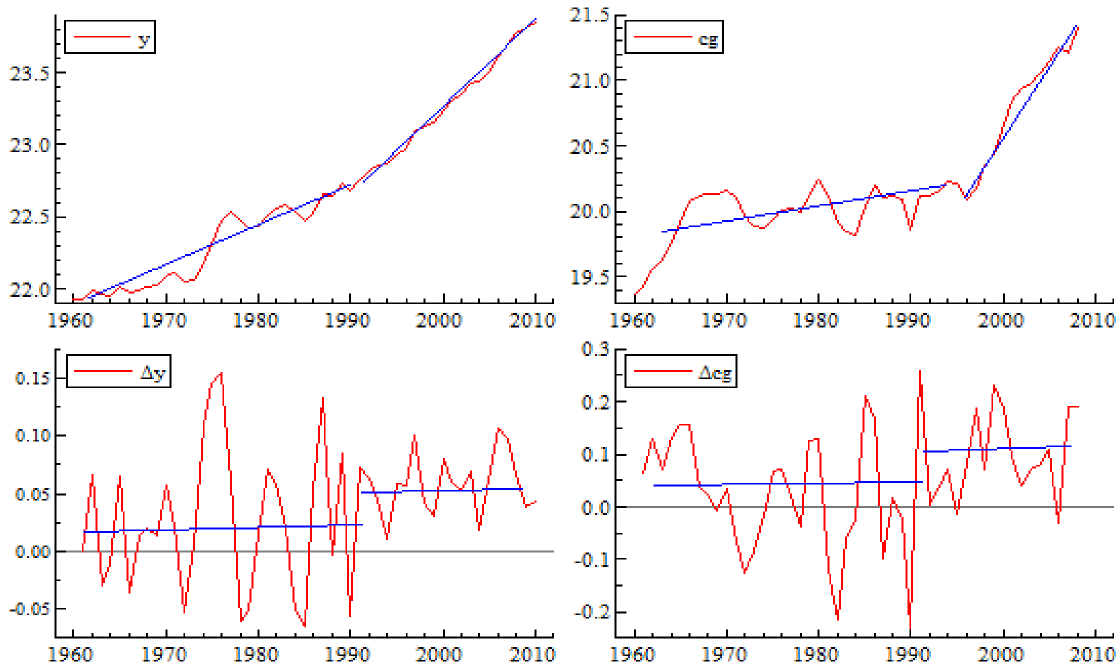

Figure 2.10: Sudan (GDP and government spending in levels and growth rates)

Chapter 3

Social Spending and Aggregate Welfare in Developing and Transition Economies

Fiseha Haile Gebregziabher
Department of Economics, University of Copenhagen
fiseha.haile.gebregziabher@econ.ku.dk

Miguel Niño-Zarazúa
World Institute for Development Economics Research (UNU-WIDER)
miguel@wider.unu.edu

Social Spending and Aggregate Welfare in Developing and Transition Economies*

Fiseha Haile Gebregziabher† *and Miguel Niño-Zarazúa*‡

Abstract

Notwithstanding the unprecedented attention devoted to reducing poverty and fostering human development via scaling up social sector spending, there is surprisingly little rigorous empirical work on the question of whether social spending is effective in achieving these goals. This paper examines the impact of government spending on social sectors (health, education, and social protection) on two major indicators of aggregate human welfare (the *inequality-adjusted* Human Development Index (IHDI) and child mortality), using a panel dataset comprising 55 developing and transition countries from 1990 to 2009. We find that government social spending has a significantly positive causal effect on the IHDI, while government expenditure on health has a significant negative impact on child mortality rate. These results are fairly robust to, among others, the method of estimation, the use of alternative instruments to control for the endogeneity of social spending, the set of control variables included in the regressions, and the use of alternative samples.

Keywords: Social spending; aggregate welfare; *inequality-adjusted* HDI; child mortality.

JEL classification: C33; H51; H52; H53; I31.

3.1 Introduction

Economic growth has been at the heart of development objectives over the past half century. The development of endogenous growth theory (Lucas, 1988; Romer, 1994) has

*This paper has been published in the World Institute for Development Economics Research (UNU-WIDER) Working Paper Series (No. 2014-082). Thanks are due to Yongfu Huang and M. G. Quibria for sharing the Human Development Index (HDI) and *inequality-adjusted* HDI data. We also thank Katarina Juselius for useful comments. Errors and omissions are ours.

†Fiseha Haile Gebregziabher is affiliated with the Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, 1353 Copenhagen, Denmark. E-mail: *fiseha.haile.gebregziabher@econ.ku.dk*.

‡Miguel Niño-Zarazúa is affiliated with the United Nations University World Institute for Development Economics Research (UNU-WIDER), Katajanokanlaituri 6 B, FI-00160 Helsinki, Finland. E-mail: *miguel@wider.unu.edu*.

brought to the fore the importance of social sector policy, which largely focuses on enhancing human development. Social spending and policy strategies that facilitate the process of innovation, knowledge creation, and information are found to have profound effects on the long-run patterns of economic growth and development (Barro, 1991; Rebelo, 1991; Benhabib and Spiegel, 1994; Barro and Sala-i-Martin, 1998). The advancement of human development has also been found to have strong links with poverty reduction (Ravallion and Chen, 1997; Schultz, 1999; Sen, 1999; Squire, 1993).

Continuing investments in the social sectors have been recognized by the international community. In 2000, the Millennium Development Goals (MDGs) were established, which comprise explicit targets to tackle extreme poverty and promote human development. To this end, much of the increase in development assistance has been directed toward the social sectors.[1] There is a widespread consensus that government spending on the social sectors has an important role to play in reducing poverty and fostering human development. In most developing countries, the public sector is a major provider of public goods, notably education, healthcare, water and sanitation, and social protection, which are critical for human capital formation and the overall development process. Government involvement in the social sectors is often justified on the grounds of market failures or positive externalities (Baldacci *et al.*, 2008). Yet, the question of whether government social spending is effective in fostering aggregate human welfare is still a subject of widespread contention.

On the one hand, a strand of the literature finds that social spending is a weak predictor of social outcomes (Flug *et al.*, 1998; Mingat and Tan, 1998; Filmer and Pritchett, 1999; Filmer *et al.*, 2000; Dreher *et al.*, 2008). On the other hand, a number of studies contend that social spending has a beneficial impact on welfare outcomes (Anand and Ravallion, 1993; Bidani and Ravallion, 1997; Gupta *et al.*, 2002, 2003; Mosley *et al.*, 2004; Baldacci *et al.*, 2008). Recent evidence suggests that social spending yields desirable results only in countries with good governance (Rajkumar and Swaroop, 2008).

The mixed results on the effectiveness of social spending can be attributed to several factors. Previous endeavors to disentangle the effect of social spending have been bedeviled by, among others, the dearth of reliable data and measurement problems. Extant studies have also been hampered by fundamental methodological shortcomings, particularly with regard to endogeneity concerns. Moreover, many studies fail to control for crucial mediating factors in the relationship between social spending and welfare outcomes. However, no lesson from the existing literature is more manifest than that of the

[1] Aid to the social sectors, namely health care, education, and the provision of safe water and sanitation facilities, increased from about 5 percent of total aid flows in the late 1960s to around 40 percent in 2011. In real terms, this meant an increase from an average of US$ 2 billion a year in the 1960s to US$ 50 billion in the 2000s, reaching US$ 64 billion in 2011 (OECD, 2012). Despite this trend, two-thirds of overall aid is still disbursed through project aid, with less than 10 percent channeled via government budgets.

considerable sensitivity of empirical estimates to the set of control variables and the choice of estimators. Further, notwithstanding the proliferation of studies assessing the effect of social sector expenditures on health and education indicators, such analysis assumes away the distributional aspects of social spending. Put differently, an improvement in such indicators does not necessarily imply that the poor are on the receiving end of the benefits.

This paper seeks to advance this longstanding debate by investigating the impact of government spending on social sectors (health, education, and social protection) on two major indicators of aggregate human welfare (the *Inequality-adjusted* Human Development Index (IHDI) and child mortality) in a sample of 55 developing and transition economies from 1990 to 2009. Unlike most previous studies, we adopt a wide array of estimation methods and empirical specifications, explicitly address endogeneity issues, control for the robust explanatory variables identified by previous studies, and perform rigorous robustness checks on the main findings. In terms of methodology, this paper lies squarely within the standard cross-country regression of the effectiveness of government spending. As such, our objective is to identify regularities in the impact of government social spending across countries that stand out to be robust across samples and estimation methods. Nonetheless, we acknowledge upfront that empirical results based on cross-country regressions should be interpreted bearing in mind the well-known shortcomings of such analysis (see Rodrik, 2005).

Our approach is similar to that of Gomanee *et al.* (2005c). However, this paper differs from Gomanee *et al.* (2005c) in several respects. To begin with, Gomanee *et al.* primarily investigate the impact of foreign aid on aggregate welfare (proxied by the HDI and infant mortality), while accounting for 'pro-poor' public spending.[2] However, the HDI remains problematic as it fails to take into account distributional inequalities in its components, namely health, education, and income. There are good reasons to expect that greater inequality in these welfare dimensions would be associated with lower development levels (Hicks, 1997; Alkire and Foster, 2010). For this reason, we use the *inequality-adjusted* HDI instead of the *conventional* HDI as an indicator of aggregate welfare.

Overall, we find strong evidence that social spending has a significantly positive causal effect on aggregate welfare. The preferred (System GMM) specification indicates that a 1 percent increase in government social spending, in percent of GDP, leads to a 0.004 points increase in the IHDI, which appears modest, albeit not negligible. The implied long-run effect of a similar increase in social spending is an increase in the IHDI of about 0.057 points. Our results are fairly robust to, among others, the method of estimation, the use of alternative instruments to control for the endogeneity of social spending, the set of control variables included in the regressions, and the use of alternative samples.

[2] Gomanee *et al.* (2005c) define 'pro-poor' public spending as comprising expenditures on health, education, and sanitation.

The remainder of this paper is organized as follows. Section 3.2 provides an overview of the empirical literature on the effect of social spending on welfare outcomes. Section 3.3 discusses the data and presents descriptive statistics. Section 3.4 deals with model specification and econometric methodology. We present the empirical results in Section 3.5 and perform extensive robustness checks on the main findings in Section 3.6. Finally, Section 3.7 concludes with reflections on policy.

3.2 Social spending and aggregate welfare

The dearth of reliable and internationally comparable data on poverty/welfare has led most previous studies to use health and education indicators as proxies. One strand of the literature finds evidence that health spending improves health outcomes. For instance, Anand and Ravallion (1993) and Hojman (1996) show that public spending on health has a significant impact on health status. This has been corroborated by a widely cited study by Bidani and Ravallion (1997), who show that health expenditures have a significantly positive impact on the poor. In the same vein, Gupta *et al.* (2002, 2003) find that health expenditure reduces child mortality. In particular, Gupta *et al.* (2003) show that the effect of health spending on health status among the poor is stronger in low-income countries than in high-income countries, suggesting diminishing marginal returns to social investment.

However, these results are not incontrovertible and a second strand of the literature finds a weak link between health spending and health outcomes. Among others, Kim and Moody (1992), Carrin and Politi (1995), Musgrove (1996), Filmer and Pritchett (1999), and Filmer *et al.* (2000) show that health spending does not yield the expected improvement in health outcomes. Filmer and Pritchett (1999) find that public spending on health has a small and statistically insignificant effect on infant and child mortality, whereas a country's per capita income accounts for most of the cross-national variation in mortality. Along similar lines, Kim and Moody (1992) examine whether healthcare expenditure is a strong predictor of reductions in infant mortality. They find that health resources are not a powerful determinant of infant mortality rates and attribute most of the observed change to socioeconomic resources.[3] Filmer *et al.* (2000) argue that inadequate institutional capacity and market failures are behind the tenuous link between health spending and improvements in health status.

In education, evidence of a positive effect of education spending is found in the influential studies of Psacharopoulos (1994) and Psacharopoulos and Patrinos (2004). Harbison and Hanushek (1992) reviewed studies assessing the impact of education spending on education outcomes in developing countries. Whereas half of the studies reported a positive

[3] Kim and Moody (1992) define socioeconomic resources as comprising Gross National Product (GNP), energy consumption daily caloric supply per capita, percentage of population enrolled in secondary education, urbanization, and with safe water supplies.

and statistically significant effect, the other half found no evidence of any measurable impact of education spending. The latter was confirmed by Hanushek (1995), Mingat and Tan (1992, 1998), Flug *et al.* (1998), Wolf (2004), and Dreher *et al.* (2008), who conclude that education spending has no discernible impact on indicators of education attainment. In contrast, controlling for governance and allowing for nonlinear relationships, Baldacci *et al.* (2008) find that both education and health spending have positive and significant impact on education and health capital. Similarly, Baldacci *et al.* (2003) provide evidence, using a covariance structure model, that social spending is an important determinant of social outcomes, particularly in the education sector (see also Gupta *et al.*, 2002).

The preponderance of evidence from previous studies suggests that the link between social spending and social outcomes is at best tenuous and at worst nonexistent. Recent evidence indicates that the ineffectiveness of social spending can be ascribed to poor governance and institutional inefficiencies. Rajkumar and Swaroop (2008) show that an increase in public spending on health and education translates into the expected improvement in health and education outcomes only in countries with good governance.

Despite the considerable effort geared towards assessing the impact of social spending on health and education indicators, such analysis assumes away the distributional aspects of social spending. Put another way, an improvement in such indicators does not necessarily imply that the poor are on the receiving end of the benefits. Some recent studies investigate the impact of social spending on poverty/welfare. Using quantile regressions for a sample of 38 countries, Gomanee *et al.* (2005a) show that higher pro-poor public spending improves welfare, proxied by the HDI and infant mortality. In particular, they find that social spending is more effective in enhancing welfare in countries with lower values of aggregate welfare. This has been confirmed by Mosley *et al.* (2004) and Mosley and Suleiman (2007), who show that higher levels of pro-poor expenditure are associated with lower levels of poverty. In contrast, using data for 104 countries over the period $1980-2000$, Gomanee *et al.* (2005c) find no evidence to suggest that social spending increases aggregate welfare. Dollar and Kraay (2001) find that many supposedly "pro-poor" policies, such as public expenditure on education and health, do not have any significant impact on the income of the poor. The World Bank's (2004) *World Development Report* contends that the 'weak' link between social spending and indicators of social outcomes is due to the fact that little of the spending goes to the poor, although Governments allocate about a third of their budgets to these sectors.

The mixed results on the effectiveness of social spending can be ascribed to several factors. Firstly, previous endeavors to untangle the effects of social spending have been bedeviled by, among others, the paucity of reliable data and measurement problems. Data on social spending and poverty/welfare are relatively scanty and truncated vis-à-vis other macroeconomic indicators. Secondly, extant studies have been hampered by fundamental methodological shortcomings. In particular, many previous studies do not deal with endogeneity concerns and disregard country-specific effects, which more often than not render empirical estimates biased. Finally, many existing studies fail to control for variables that affect the efficacy of social spending. No lesson from the existing literature is, however, more manifest than that of the considerable sensitivity of empirical estimates to the set of control variables and the choice of estimators. In particular, poor governance and institutional inefficiencies are frequently cited as the main factors explaining the ineffectiveness of government spending. The role of institutions and governance in mediating the nexus between social spending, welfare, and sustained growth is highlighted by, *inter alia*, Abed and Gupta (2002), Gupta *et al.* (2002), Mauro (1998), Rodrik *et al.* (2004), Hausmann *et al.* (2005), and Rajkumar and Swaroop (2008).

The present study builds on the existing empirical work and attempts to correct the aforementioned weaknesses. Unlike many preceding studies, we adopt a wide array of estimation methods and empirical specifications, explicitly address endogeneity issues, control for the robust explanatory variables identified by previous studies, and perform rigorous robustness checks on the main findings.

3.3 Data issues

3.3.1 Social spending

The data on government spending on social sectors (health, education, and social protection) for the period $1990 - 2009$ are taken from the International Monetary Fund's (IMF) *Government Finance Statistics* (GFS: 2011 edition). Although the GFS database provides data dating as far back as 1972, we were able to use only data for the period $1990 - 2009$. Data before 1990 are based on the accounting system described in the 1986 GFS manual (GFSM 1986), while the data from 1990 onwards are based on a revised accounting method outlined in GFSM 2001. The revision resulted in major changes in, *inter alia*, the definitions and classification of expenditures, and the timing at which economic events are to be recorded. For instance, in terms of the classification of expense by economic type, the definition of current and capital spending differs between GFSM 1986 and GFSM 2001. Concerning the functional classification of expense, while expenditures are classified into 14 categories in the revised GFS system, they were categorized into only 10 functional categories in GFSM 1986. Moreover, transactions and other economic flows are recorded on an accrual basis in GFSM 2001, i.e. flows are recorded at the time when a transaction occurs, regardless of the timing of cash flows. In contrast, in the GFSM 1986, transactions are recorded when cash is exchanged. See IMF (2001: Appendix I) for more detailed discussion.

The data from the GFS database are given in local currency units (LCU). To the best of our knowledge, previous studies use either the data in LCU or convert them into one monetary unit using exchange rates. However, this is likely to be misleading because exchange rates do not necessarily reflect the relative purchasing power across countries. Therefore, we transformed the data in LCU into purchasing power parity (PPP) dollars. More specifically, for the purpose of transforming LCU at time t to time t_{base}, which is set at 2005, i.e. $t_{base} = 2005$, we use data on the consumer price index (CPI) from the IMF *World Economic Outlook* (2012), and the PPP exchange rate and the official exchange rate from the World Bank's (2012) *World Development Indicators* (WDI) and *Global Development Finance* (GDF). Denoting the PPP exchange rates for the base year as PPP_{base} (LCU per dollars), the currency transformation into constant dollars or PPP is carried out as follows:

$$V_t^{base} = \frac{CPI_{t_{base}}}{CPI_t \; E_{base}} V_t^{LCU} \tag{1}$$

where V_t^{LCU} is the value in LCU at time t, V_t^{base} is the value in PPP at time t, and E_{base} stands for the official exchange rate when constant US dollars are needed and PPP exchange rate when constant PPP is the value of interest.

Some of the original data, in particular those from the World Bank, are given in current dollars. To transform these values into PPP, the currency values were transformed first into LCU and then Equation (1) was applied to the resulting LCU value. Put in other terms, supposing G_t is the official exchange rate (LCU per dollars), the formula for transformation is:

$$V_t^{base} = \frac{CPI_{t_{base}} G_t}{CPI_t \; E_{base}} V_t^{current} \tag{2}$$

where $V_t^{current}$ is the value in current dollars at time t.

For some of the countries in our sample (see Table C in Appendix A), the government spending data are partly in cash and partly in accrual, which raises a question of comparability. Confining the analysis to include only cash data would substantially reduce the sample size, while using only accrual data would considerably shorten the time span and limit the number of countries that could be included in the sample.[4] In either case, the sample size would be too small to perform any meaningful analysis. It is difficult to convert the cash data into accrual (or vice versa) without making constraining assumptions. In an IMF working paper, Seiferling (2013) suggests that merging both data and including an indicator (dummy) variable in parametric analysis to account for any systematic differences would be acceptable for most data series from the GFS database. This suggests that cash data could be taken as a proxy for accrual data and vice versa. However, given the methodological changes introduced by GFSM 2001, mixing cash and accrual data does not seem a plausible option.

Given that most of the data are in cash, a possible way to circumvent this problem is to extend the cash data using the annual growth rates for the accrual data. This is,

[4] Note that we have excluded countries with less than nine observations for government social spending over the period $1990 - 2009$.

in effect, tantamount to predicting the values of the cash data for periods for which we have only accrual data. The underlying assumption is that the year-to-year growth rates of the cash and accrual data are not systematically different from each other although the actual values might differ. This is, in our view, a far less restrictive assumption than the one suggested by Seiferling (2013).

Another limitation of the GFS database is that the government spending data for some of the sample countries have gaps (see Table C in Appendix A). Hence, we imputed the missing observations using health expenditure data at constant 2005 PPP from WDI (2013).[5] In Section 3.6, we test the robustness of our results by excluding the countries with data that are partly in cash and partly in accrual, and countries with data that have gaps. The results remain robust across the subsample of countries.

Finally, note that we use data on central government (CG) spending as a proxy for general government (GG) spending (central plus subnational). GFS data on social sector spending for the GG are scanty for most countries, whereas there is a more comprehensive coverage for CG accounts. Although a potential solution would be to assemble data for the GG based on data for lower government levels (central as well as state and/or local governments), the latter are missing for most countries and periods. Hence, we use CG data as a proxy measure for GG spending, although the latter would be a more satisfactory measure.[6]

3.3.2 Aggregate welfare

Most, if not all, recent studies on poverty rely on income-based poverty measures, such as the headcount index– the percentage of population living on less than a $1.25 a day (adjusted for purchasing power). Although these measures are purported to be internationally comparable, their reliability as indicators of the depth and severity of poverty has been questioned on the grounds that they fail to capture non-income dimensions of poverty (Reddy and Pogge, 2010). Income (or consumption) poverty captures only one aspect of poverty and is at best a very rough approximation of deprivation in other dimensions and thus multidimensional poverty.

Income can, of course, be instrumental in providing the material resources needed for people to lead a decent life. However, income constitutes only a means to an intrinsic end and as such it says nothing definite about the quality of life that people lead (Anand and Sen, 1992; Sen, 1999). For instance, according to the headcount index, a person with less than a $1.25 a day but who have access to improved healthcare, education,

[5] Data on health expenditure are widely available for these countries and are highly correlated with our social spending data. The correlation between government spending on social sectors (health, education, and social protection) and total health expenditure for periods for which both series are available is fairly high, with correlation coefficients of 0.75 or more for 9 of the 10 countries with data gaps.

[6] Note that using CG spending as a proxy for GG spending is not uncommon in the literature (see, among others, Feyzioglu *et al.* (1998), Baldacci *et al.* (2003), and Acosta-Ormaechea and Morozumi (2013)).

and other social services would be counted as poor, whereas a person having more than a \$1.25 a day but suffering from ill health, is ill-educated and disempowered would be deemed non-poor. Ruggeri-Laderchi, Saith, and Stewart (2003) show that in India, 43% of children and more than half of adults who were poor, judged on the basis of health or education indicators, were counted as non-poor using monetary measures of poverty; similarly, more than half of the nutrition-poor children were not in monetary poverty. Thus, the tendency to place an overriding emphasis on income as a poverty measure may obscure the more intrinsic ends of development. This prompts the need for an indicator that better reflects income as well as non-income dimensions of human development.

In view of the limitations of conventional poverty measures and the paucity of reliable cross-country data on poverty over time, we use the IHDI and child mortality rate as indicators of aggregate welfare. The HDI, an index between 0 and 1, is built from three separate components: (1) *longevity*, measured by life expectancy at birth; (2) *educational attainment*, proxied by a weighted average of adult literacy (with a two-thirds weight) and school enrollment rates (with a one-third weight); and (3) *standard of living*, measured by income (GNI) per capita (adjusted for purchasing power). HDI is a widely used measure of aggregate welfare and is calculated using consistent data and methodology (UNDP, 2011). Note in passing that, because the HDI includes GNI per capita as one of its components, we expect poverty to be inversely correlated with HDI insofar as income poverty is lower in countries with higher GNI.

Note, however, that the HDI, like other welfare/poverty indicators, is far from perfect and has been criticized for not addressing distributional inequalities in education, health, and income across the population. There are good reasons to think that greater inequality in these spheres would be associated with lower development levels (Hicks, 1997; Alkire and Foster, 2010). In response to this criticism, the UNDP introduced a new measure, the *inequality-adjusted* HDI, which incorporates distributional aspects into the *conventional* HDI. IHDI equals HDI if there is no distributional inequality in the above-mentioned three spaces. In other words, the gap between HDI and IHDI reflects inequality in the dimensions of human development; the greater the gap, the greater the inequality. Therefore, the IHDI is our preferred measure.

Annual data on HDI for the period 1990 − 2009 come from the UN's *Human Development Report* (HDR). Data on the corresponding IHDI are from Huang and Quibria (2013).[7] As an additional indicator of aggregate welfare, we follow Burnside and Dollar (1998) and Gomanee *et al.* (2005c) and use child mortality rate, for which data is sufficiently available. Reddy and Pogge (2010) argue that non-income indicators of human

[7]UNDP (2011) calculates IHDI only for the year 2011. Following the procedure to compute HDI and IHDI outlined in UNDP (2011: Technical notes 1 and 2), Huang and Quibria (2013) calculate IHDI for all countries in our sample based on the HDI data from the UN's HDR and data on life expectancy at birth (years), duration of primary and secondary education (years), and GNI per capita (PPP) from WDI (2012).

development, such as child mortality rate, can be equally informative as income-based poverty measures, considering the serious shortcomings of the latter.

It should be noted that the HDI (as well as its inequality-adjusted counterpart) and child mortality are not measures of poverty or deprivation, although they are important welfare indicators in their own right. Unlike measures that are based on absolute poverty line, the HDI does not provide a certain threshold or cutoff point under which people can be considered deprived or poor in the spheres of health, education, and income. However, there is considerable correlation in our sample between IHDI and child mortality, on the one hand, and income poverty measures, on the other. In our data, the correlation at country level between IHDI and the "$1.25 a day" measure is −0.82; between child mortality and the "$1.25 a day" measure it is 0.84. This unveils the substantial information overlap between welfare measures and income-based poverty measures, and indicates that our analysis may have implications for the social spending−poverty nexus.

3.3.3 Descriptive statistics

Table B in Appendix A presents basic summary statistics for the welfare outcomes of interest, i.e. the IHDI and child mortality, and the main explanatory variables included in our empirical model. The summary statistics show that there is considerable variation in IHDI and child mortality rate across the sample countries. For instance, HDI ranges from 0.259 to 0.813, while IHDI from 0.086 to 0.271. Similarly, child mortality rate ranges from 6.6 to 204. The countries also exhibit substantial variation in the size of government social spending, in percent of GDP.[8]

The pairwise correlation coefficients (Table D in Appendix A) reveal that GDP per capita, social spending, trade openness, bureaucratic quality, and democracy are significantly positively correlated with IHDI, whereas IHDI is negatively correlated with terms of trade and age dependency ratio, albeit weakly so with the former. The positive correlation between IHDI and social spending may suggest that an increase in government spending on the social sectors leads to a higher level of welfare. However, this may as well reflect that countries increase social sector spending when faced with pressing demands to improve aggregate welfare. In the child mortality specification, GDP per capita, health spending, the degree of urbanization, female education, access to safe water, and access to improved sanitation facilities are significantly negatively correlated with child mortality rate, whereas the latter is positively correlated with fertility rate.

[8]Among the countries in the sample, the highest and lowest IHDI correspond to Lithuania (in the year 2008) and Yemen (in 1990) respectively. The highest government social spending, in percent of GDP, corresponds to Yemen (in 1990). The highest child mortality rate corresponds to Ethiopia (in 1990) while the lowest to Belarus (in 2009).

3.4 Model specification and econometric methods

3.4.1 Model specification

We use a panel dataset comprising 55 developing and transition economies[9] from 1990 to 2009. Given that most of the data are available on an annual basis and that the number of countries is relatively small[10], we first focus on an annual panel spanning the period 1990 − 2009, which provides more degrees of freedom. The downside of using annual observations is that empirical estimates may be driven by short-term "noise". In addition, the variation in annual observations may be insufficient to reflect the effect of structural variables with little variation over time. It is common in the literature to use time-averaged data to smooth out potential business cycle effects and reduce measurement errors. The robustness of the results from the annual panel is tested in Section 3.5 using data averaged over six three-year and one two-year epochs: starting in 1990 − 1992 and ending with 2008 − 2009.[11]

We estimate two models: the first model estimates the effect of government spending on the social sectors on the IHDI. The model takes the following form:[12]

$$W_{i,t} = \theta_0 W_{i,t-1} + \beta_1 Y_{i,t} + \beta_2 S_{i,t} + \beta_3 I_{i,t} + \beta_4 D_{i,t} + \gamma X + \eta_i + v_t + \varepsilon_{i,t} \quad (3)$$

where the subscripts i and t denote country and year respectively; $W_{i,t}$ stands for IHDI; $W_{i,t-1}$ for one-period lagged IHDI, θ_0 measures the persistence of $W_{i,t}$; $Y_{i,t}$ for real GDP per capita[13]; $S_{i,t}$ for government spending on social sectors (health, education, and social protection), in percent of GDP; β_2 is the key parameter of interest and measures the direct effect of government social spending on IHDI once we have controlled for the relevant explanatory variables; $I_{i,t}$ is a vector comprising indicators of institutional quality; $D_{i,t}$ is an indicator of the level of democratization; X is the vector of control variables that may affect $W_{i,t}$ and $S_{i,t}$; η_i denotes unobserved country-specific and time-invariant effects; v_t is a vector of time dummies capturing universal time trends; and finally, $\varepsilon_{i,t}$ represents the disturbance term. Table A in Appendix A provides variable definitions and description of data sources.

As indicators for institutional quality, we follow Rajkumar and Swaroop (2008) and use the indices of bureaucratic quality and corruption from the International Country

[9] Table C in Appendix A presents the list of countries included in the sample.

[10] Note that the number of countries drops in some of the regressions because of the limited overlap between the control variables included in the regressions.

[11] Although it is more common in the literature to use four- or five-year averaged data, this would significantly reduce the sample size, given that the number of countries in the analysis is relatively small. Hence, we organized the data into three-year periods.

[12] This specification draws mainly on Kosack (2003) and Gomanee *et al.* (2005c).

[13] In the models that employ the annual panel, we use GDP per capita in the preceding period ($Y_{i,t-1}$) to address potential endogeneity problems.

Risk Guide (ICRG). The bureaucratic quality index ranges from 1 to 4 and measures the soundness of institutions and the quality of the civil service. The corruption index, ranging from 0 to 6, measures corruption within the political system, with a score of 0 pointing to very high corruption. The democracy index comes from the Polity IV project (Marshall *et al.*, 2013). The measures of economic policy we use in the regressions are standard in the literature: inflation rate, proxy for a country's monetary policy stance; trade openness, sum of exports and imports as a percentage of GDP; and the share of domestic credit to private sector in GDP, an indicator for the potential role of financial sector development in improving welfare. Many of the countries in our sample are vulnerable to the vagaries of the international economy and particularly to primary commodity price fluctuations. The terms of trade index controls for this effect.

The relationship between $W_{i,t}$ and $S_{i,t}$ is estimated using two functional forms: (i) linear-log specification, where $W_{i,t}$ is linear and $S_{i,t}$ is logarithmic, and (ii) log-log specification, where both variables are log-transformed. The linear-log specification may be preferable because it provides the absolute change in the IHDI associated with a percent change in social spending. The log-log specification has the added convenience of smoothing the data and allowing coefficients to be interpreted as elasticities.[14]

The second model specifies child mortality as a function of government health spending and other covariates:

$$C_{i,t} = \alpha_0 C_{i,t-1} + \delta_1 Y_{i,t} + \delta_1 H_{i,t} + \delta_3 I_{i,t} + \delta_4 D_{i,t} + \Phi M + \eta_i + v_t + \varepsilon_{i,t} \tag{4}$$

where $C_{i,t}$ is the child mortality rate in country i in year t; $C_{i,t-1}$ is one-period lagged $C_{i,t}$, with α_0 capturing the persistence in $C_{i,t}$; $H_{i,t}$ stands for government health spending, in percent of GDP; M is a vector of robust explanatory variables associated with child health; and the remaining variables are as defined previously.[15] Income is one of the crucial determinants of health status (Pritchett and Summers, 1996). Moreover, a number of studies show that higher levels of female education are associated with better health status of children as well as the population in general (Filmer and Pritchett, 1999; Cutler *et al.*, 2006). Hence, following Baldacci *et al.* (2008), we include the share of female students in primary and secondary schools as an indicator for gender equality, which takes account of the institutional factors that may have significant bearing on child health through female education.[16] There is also ample evidence that health status is affected by access to safe water and improved sanitation facilities (Mishra and Newhouse, 2009;

[14] Moreover, when the initial IHDI is higher, specifying IHDI in logs would allow a given change in social spending to have a larger effect on IHDI.

[15] $C_{i,t}$ and $H_{i,t}$ are log-transformed, as is common in the literature. However, all regression results are fairly robust to specifying these variables in levels instead of logs.

[16] Although gender equality for the population as a whole would be a more plausible measure, such data is not available. Moreover, data limitation precluded the use of female adult literacy rate as a proxy for female education (with about 90% of the data missing).

Rajkumar and Swaroop, 2008); the degree of urbanization (Schultz, 1993), and fertility rates (Baldacci *et al.*, 2008; Mishra and Newhouse, 2009).

Note that the availability of data on the aforementioned variables differs considerably (see Table B in Appendix A). Data on the macroeconomic variables, such as GDP per capita, inflation, and trade openness, and the indicators of aggregate welfare are available for almost all countries and time periods, whereas data on social spending and some other variables are relatively limited. Hence, the sample size differs across specifications, depending on the control variables included and the instrumental variables used.

3.4.2 Econometric methods

As a first approximation, we estimate Equations (3) and (4) using ordinary least squares (OLS). However, this presumes, *inter alia*, that social spending is exogenously determined, which cannot reasonably be expected to hold given that both social spending and the outcome variables ($W_{i,t}$ and $C_{i,t}$) may be affected by the same unobserved factors and the possibility of reverse causality. For instance, governments may increase their spending on the social sectors when faced with low levels of welfare and poor health outcomes. In order to address the endogeneity problem, we first instrument for social spending in a two-stage least squares (2SLS) and fixed-effects (FE) framework.[17] The presence of country fixed-effects, η_i, in Equations (3) and (4) suggests that the preferred approach is the FE model, which allows us to mitigate heterogeneity-induced bias and control for fixed-effects-related endogeneity. However, the FE model removes a considerable portion of the variation in the right-hand-side variables, which may exacerbate measurement error.

There is evidence suggesting that Instrumental Variables (IV) methods may suffer from finite sample biases and their use is mostly justified on asymptotic grounds (Baum *et al.*, 2007). Limited Information Maximum Likelihood (LIML) and the continuously updated Generalized Method of Moments (CUE) have been shown to have better finite sample performance than IV estimators, although they provide no gain in asymptotic efficiency. In particular, LIML and CUE tend to perform better than IV methods in the presence of weak instruments (Hahn *et al.*, 2004). In light of the different properties of alternative estimators, we investigate the robustness of the empirical results using a range of econometric methods: 2SLS, FE, LIML, Fuller's modified LIML (henceforth Fuller), and CUE.

It is evident that the inclusion of lagged dependent variables in Equations (3) and (4) would render FE estimates inconsistent because they would be correlated with the transformed errors, even if they are uncorrelated with $\varepsilon_{i,t}$. 2SLS estimations are also

[17] Statistical tests (not shown here) indicate that the random-effects model should be rejected in favor of the fixed-effects specification (Hausman test, p-value = 0.00). Hence, we report only fixed-effects estimates.

likely to suffer from dynamic panel bias because η_i is correlated with the lagged dependent variables. For this reason, we first disregard the lagged dependent variables and estimate the models using 2SLS and FE, the robustness of which will be tested using dynamic panel data estimators. Although the 2SLS estimator does not allow for country fixed-effects, η_i, we capture some of the influence of omitted spatially-correlated fixed effects, such as those emanating from geography and natural endowments, using regional dummy variables for Sub-Saharan Africa (SSA), Asia, and Latin America and Caribbean (LAC).

Notwithstanding that finding reliable instruments is a daunting challenge, we use several instrumentation strategies to robustly identify the casual effect of social spending. In the 2SLS and FE specifications, we use 'external' instruments to control for the endogeneity of social spending, the robustness of which will be tested using 'internal' instruments− lags of the instrumented variables− in a dynamic panel data framework.[18] However, we hasten to note that none of our instrumentation strategies are foolproof.

Following Easterly and Rebelo (1993) and Tanzi (1992), we use the logarithm of population and the share of agriculture in GDP as external instruments. Easterly and Rebelo (1993) find that the scale of the economy (measured by its population) is an important determinant of fiscal policy in general and the level of public spending in particular. They provide evidence in favor of strong scale effects: countries with higher population have lower public spending. High population countries tend to spend less, yet there is no reason to suspect that a country can have higher or lower level of welfare simply because it has more or less people. In our sample, the log of population is highly correlated with the share of social spending in GDP (a correlation coefficient, r, of -0.52).

Another factor that is found to influence the level of social spending is a country's economic structure, which is reflected in the share of agriculture in GDP. Tanzi (1992) argues that the more agricultural a country is, the more difficult it is to raise the tax level and thus to increase public spending.[19] This appears to be borne out by the data. Social spending and agriculture (both in percent of GDP) are significantly negatively correlated ($r = -0.51$).

Further, we use the ICRG index of ethnic tensions as an additional instrument, which has been used in earlier studies (see Dreher *et al.*, 2008). Several studies indicate that ideological and ethnic divisions result in higher public spending by compounding the common pool problem, i.e. by inducing one section of society to neglect the tax burden falling on others (Von Hagen, 2005; Alesina *et al.*, 1999, 2003). We find a considerable

[18] It is commonplace in the recent literature to instrument for social spending using its lagged values. However, the economic motivation behind the use of lagged values as instruments may be questionable because it is tantamount to claiming that contemporaneous social spending affects outcomes but lagged spending does not.

[19] Many studies find that it is difficult to impose heavy taxes on the agricultural sector, although the sector has often been subject to heavy implicit taxes (Ahmad and Stern, 1991; Tanzi, 1992). The reason is that agricultural activity is small scale and spatially spread, which is more so in developing countries (*Ibid.*).

correlation in our sample between social spending and the ethnic tensions index ($r = 0.39$). The question of whether the instruments discussed above are valid and strong is an empirical issue, which is dealt with in Section 3.6.[20]

As a robustness check, we organize the data into three-year periods and estimate Equations (3) and (4) using dynamic panel data estimators.[21] As mentioned above, the presence of lagged dependent variables and country fixed-effects, η_i, poses a challenge that necessitates the use of more sophisticated econometric methods. The recent literature is replete with studies using the Arellano and Bond (1991) first-differenced GMM (Dif-GMM) estimators to circumvent the endogeneity problem by removing η_i using first-differencing or orthogonal deviations and then deploying suitably lagged values of the independent and dependent variables as instruments. Nonetheless, the Dif-GMM estimator suffers from large finite-sample biases and poor precision when the time series are persistent. In such cases, the lagged levels of the series are weakly correlated with the lagged first differences, thereby making the instruments for the first-differenced equations "weak" (Blundell and Bond, 1998). In our case, the coefficients on lagged IHDI and lagged child mortality are quite large (see Tables 2 and 4), which would make lagged levels of IHDI and child mortality weak instruments for first differences.

The System GMM estimator (Sys-GMM) has been designed to work around the weak instrument problem associated with Dif-GMM.[22] Sys-GMM solves a system of level and difference equations. For instance, for the IHDI specification (Equation 3), Sys-GMM estimates the following regression equations:

$$W_{i,t} = \theta_0 W_{i,t-1} + \beta_1 Y_{i,t} + \beta_2 S_{i,t} + \beta_3 I_{i,t} + \beta_4 D_{i,t} + \gamma X + \eta_i + v_t + \varepsilon_{i,t} \tag{5}$$

$$\Delta W_{i,t} = \theta_0 \left(\Delta W_{i,t-1}\right) + \beta_1 \left(\Delta Y_{i,t}\right) + \beta_2 \left(\Delta S_{i,t}\right) + \beta_3 \left(\Delta I_{i,t}\right) + \beta_4 \left(\Delta D_{i,t}\right) + \gamma \left(\Delta X\right) + \Delta v_t + \Delta \varepsilon_{i,t} \tag{6}$$

Suitably lagged differences of the endogenous variables are used as instruments in the level equation (5), while suitably lagged levels of the endogenous variables are used as instruments in the first differenced equation (6).[23] Unlike Sys-GMM, Dif-GMM estimates only Equation (6).

[20] An instrument is valid and strong if it affects the dependent variable only through the instrumented variable and is sufficiently correlated with the latter respectively.

[21] In the annual panel, the number of time series observations, T, is relatively "large" compared to the number of countries, N. However, the GMM estimators are particularly designed for the panel data setting with fixed T and large N, and as T increases they may lose even consistency (Anderson *et al.*, 2010).

[22] Blundell and Bond (1998) developed, building on Arellano and Bover (1995), the System GMM estimator.

[23] See Blundell and Bond (1998) and Roodman (2009a) for more detailed exposition of the Sys-GMM estimator.

The Sys-GMM estimator relies on the following assumptions (Blundell and Bond, 1998)[24]:

$$\begin{aligned}
&E(\eta_i) = E(\varepsilon_{i,t}) = E(\eta_i \varepsilon_{i,t}) = 0 \\
&E(\varepsilon_{i,s}\varepsilon_{i,t}) = 0, \quad s \neq t \\
&E(W_{i,1}\varepsilon_{i,t}) = 0, \quad t = 2, ..., T \\
&E(S_{i,1}\varepsilon_{i,t}) = 0, \quad t = 2, ..., T \\
&E(\Delta W_{i,2} n_i) = 0 \\
&E(\Delta S_{i,2} n_i) = 0
\end{aligned} \tag{7}$$

The last four assumptions in Equation (7) are the initial conditions underlying the Sys-GMM estimator. These new assumptions require that the initial conditions satisfy mean-stationarity[25], which is not trivial. In the IHDI specification, the assumption would be that the initial levels of social spending and IHDI are uncorrelated with all future unobserved shocks in the IHDI, and that the initial changes in the IHDI and social spending are uncorrelated with the unobserved country fixed effects. In other words, changes in any instrumenting variable should be uncorrelated with the fixed effects. This assumption can hold only if the fixed effect and the autoregressive coefficient$-$ θ_0 in Equation (3)$-$ offset each other in expectation across the whole panel (see Blundell and Bond, 1998; Roodman, 2009b). For this condition to be satisfied, θ_0 must have absolute value less than unity, i.e. $|\theta_0| < 1$ (Blundell and Bond, 1998). Tables 2 and 4 show that this is fulfilled. IHDI and child mortality are highly persistent series; however, we specify them as stationary persistent processes.

Sys-GMM significantly improves the accuracy of estimates by exploiting additional moment conditions that are informative even for persistent data (Blundell and Bond, 1998). Thus, we opt for the Sys-GMM estimator given that it addresses some of the finite-sample biases and imprecision inherent in the Dif-GMM. However, as alluded to above, the additional moment conditions of the Sys-GMM estimator do not come without a cost and some restrictions on the initial conditions are required, which are not empirically testable and may be less innocuous. In particular, the instruments for the level equations are valid so long as they are orthogonal to the country fixed effects. In addition, it has recently come to light that Sys-GMM may equally suffer from the weak instrument problem, particularly when the time series is large and substantial unobserved heterogeneity exists (Hayakawa, 2006; Bun and Windmeijer, 2010). Hence, to mitigate this problem and thereby check the sensitivity of the results, we complement the internally generated set of instruments with the external instruments discussed earlier.

[24] For illustrative purposes, we assume that social spending ($S_{i,t}$) is the only endogenous right hand side variable in Equations (5) and (6). Note also that we state only the assumptions on the initial conditions required for the consistency of estimates of the coefficients on lagged IHDI and social spending. Similar initial conditions also apply to the remaining variables in Equations (5) and (6).

[25] This requires the data to be mean-stationary, i.e. the mean of the data should not depend on time t.

Another potential deficiency of the Sys-GMM (and, of course, the Dif-GMM) estimator is that the number of internal instruments grows quadratically as the number of time periods increases. Roodman (2009b) cautions that instrument proliferation can over-fit endogenous variables, biasing coefficient estimates and weakening the Hansen test of the instruments' joint validity. Therefore, we follow Roodman (2009b) and reduce the instrument count by "collapsing" instruments.[26] Furthermore, organizing the data into three-year periods considerably reduces the sample size. All these caveats should be borne in mind when interpreting the Sys-GMM results.

3.5 Results[27]

Section 3.5.1 presents the results on the impact of government social spending on the IHDI, while Section 3.5.2 discusses the results on the nexus between government health spending and child mortality. Note that all regressions include time dummies and all t-statistics (reported in parentheses) are computed based on standard errors that are corrected for arbitrary heteroskedasticity and autocorrelation. Besides, all 2SLS, LIML, Fuller, and CUE regressions include regional dummies for SSA, Asia, and LAC. Further, all Sys-GMM results are based on the two-step estimator, which allows for Windmeijer's (2005) finite sample correction.

3.5.1 Social spending and IHDI

The OLS, 2SLS, and FE regression results are reported in Table 3.1. Consider first the benchmark OLS estimation results in column 1. Most of the coefficients are statistically significant and have the expected sign. Social spending is positively associated with IHDI; however, this does not say anything about causality. In order to deal with the potential endogeneity of social spending, we next estimate Equation (3) using 2SLS and employing the log of population and the ethnic tensions index as instruments. Because potential multicollinearity problems may render the parameter estimates unstable, we first include few covariates (column 2) and then include additional institutional and macroeconomic indicators (column 3).

The Hansen test of over-identifying restrictions indicates that the validity of the instruments cannot be rejected. Under-identification tests (not reported) find that the null hypothesis of zero correlation between the instruments and social spending is strongly rejected (p-value: 0.00). The Kleibergen-Paap Wald F test of weak identification (which, like the standard F-statistic, tests for the strength of the partial correlation between the included endogenous variable and the excluded instruments but makes finite-sample cor-

[26] Roodman (2009b) demonstrates that, in some common cases, collapsing instruments is superior to simply restricting the lag ranges.

[27] The 2SLS, LIML, Fuller-LIML, and CUE estimation results are obtained using the Stata (version 12) command *ivreg*2 (Baum *et al.*, 2007), the FE estimates using *xtivreg*2 (Schaffer, 2010), and the Sys-GMM results using the *xtabond*2 routine (Roodman, 2005).

Table 3.1: IHDI equation (2SLS and FE regressions)

Dependent variable	IHDI				
Regression	(1)	(2)	(3)	(4)	(5)
Method	OLS	2SLS	2SLS	FE	FE
Social spending (*ln*)	0.009	0.01	0.014	0.009	0.012
(% GDP)	(10.07)***	(3.67)***	(6.02)***	(2.32)**	(2.54)**
GDP per capita (*ln*)	0.029	0.034	0.026	0.018	0.017
($t-1$)	(24.44)***	(15.02)***	(13.65)***	(4.97)***	(6.10)***
Openness (*ln*)	0.004	0.004	0.001	0.003	0.003
	(3.07)***	(1.72)*	(0.43)	(1.95)*	(1.99)**
Terms of trade (*ln*)	−0.011	−0.006	−0.012	−0.004	−0.005
	(4.94)***	(1.81)*	(3.76)***	(2.70)***	(2.73)***
Inflation	−0.002		−0.002		0.0002
	(2.22)**		(1.47)		(0.61)
Bureau. quality	0.001		0.003		0.002
	(1.50)		(1.70)*		(2.78)***
Democracy	0.001		0.001		−0.0002
	(4.12)***		(2.66)***		(1.34)
Age dependency (*ln*)	−0.035		−0.033		−0.016
	(6.92)***		(4.29)***		(1.91)*
Number of countries	51	51	51	40	38
Observations	504	529	504	442	417
R-squared	0.92	0.90	0.91	0.90	0.90
Kleibergen-Paap F stat.		33.60	42.17	13.18	9.82
Stock-Wright LM stat.		9.09	18.80	4.64	6.13
(p-value)		0.01	0.00	0.098	0.05
Hansen test[a]		0.22	0.77	0.68	0.69

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Notes: Columns 2 and 3 instrument for social spending using log population and the ethnic tensions index; columns 4 and 5 use ethnic tensions index and central government (CG) budget deficit (% of GDP), as instruments. Resource allocation to the social sectors is likely to be influenced by the fiscal stance of the CG (proxied by budget deficit). The assumption that budget deficit affects IHDI only via social expenditure should be supported by empirical tests, which we provided in this table.

Source: Authors' analysis based on the data described in Appendix A.

rections) comfortably exceeds conventional critical values. Further, the Stock-Wright LM S statistic, which is robust in the presence of weak instruments, confirms the existence of significant correlation between the excluded instruments and the dependent variable. These suggest that the instruments are sufficiently correlated with social spending but not significantly correlated with IHDI once the relevant explanatory variables are controlled for, which confirms the validity of our specifications.[28]

The estimated coefficient on social spending is positive and significant. It indicates that a 1 percent increase in government spending on the social sectors, in percent of GDP, increases IHDI by 0.014 points (column 3). At the same time, a 1 percent increase

[28] As an additional crude way of checking whether the instruments pass the exclusion restriction, we simply included them in the second stage regression and found that they are consistently insignificant.

Table 3.2: IHDI equation (System GMM regressions)

Dependent variable	IHDI		
Regression	(1)	(2)	(3)
Lagged IHDI	0.96	0.93	0.93
	(30.75)***	(35.83)***	(18.86)***
Social spending (ln) (% GDP)	0.003	0.004	0.005
	(2.03)**	(2.33)**	(1.88)*
GDP per capita (ln)	0.002	0.002	0.002
	(1.83)*	(2.63)***	(1.84)*
Openness (ln)	0.001	0.005	0.005
	(0.38)	(2.14)**	(2.14)**
Terms of trade (ln)	−0.005	−0.005	−0.005
	(1.58)	(1.77)*	(1.79)*
Inflation		−0.001	−0.0003
		(1.18)	(0.27)
Bureaucratic quality		0.001	0.001
		(0.63)	(0.48)
Democracy		0.00004	−0.000
		(0.30)	(0.01)
Number of countries	43	42	42
Observations	182	175	175
Number of instruments	30	33	30
Hansen test[a]	0.60	0.68	0.53
Diff-in-Hansen test[a]	0.54	0.59	0.41
Autocorr. (second order)[a]	0.01	0.03	0.03
Autocorr. (third order)[a]	0.48	0.80	0.56

[a] p-values Significance level: *10%; **5%; ***1%

Notes: All regressions except column 3 use both internal (third and longer lags of IHDI, social spending, GDP per capita, and openness) as well as external (log population, agriculture (% of GDP), and the ethnic tensions index) instruments; column 3 uses only the aforementioned internal instruments.

Source: Authors' analysis based on the data described in Appendix A.

in lagged per capita income is associated with an increase in IHDI in the order of 0.026 points, which is consistent with previous studies finding that higher levels of income are associated with improved welfare (or lower poverty) (Gomanee *et al.*, 2005a; Gomanee *et al.*, 2005c; Mosley *et al.*, 2004; Alvi and Senbeta, 2012). Our estimates also show that IHDI is higher in countries with stronger democratic institutions and better bureaucratic quality. Moreover, a decline in the terms of trade and age dependency ratio are linked with an increase in IHDI.

Note, however, that the 2SLS regressions do not allow for country fixed-effects, which may have a significant bearing on the empirical results. Columns 4 and 5 of Table 3.1 are the FE counterparts of columns 2 and 3 respectively. The Hansen test cannot be rejected at conventional levels of significance, suggesting that there are no signs that the instruments are endogenous. Moreover, the Stock-Wright S statistic rejects (at the 5% level) its null hypothesis, indicating that the endogenous regressors are relevant. However, the Kleibergen-Paap F-statistic in column 5 is slightly below the rule of thumb threshold of 10 proposed by Staiger and Stock (1997). The FE results reveal that increasing social spending by 1 percent would increase IHDI by 0.012 points, which is consistent with the 2SLS estimate. Other policy- and institutions-related variables also affect IHDI. IHDI is positively associated with trade openness and bureaucratic quality, whereas it is negatively correlated with the terms of trade. The FE estimates are by and large similar to the 2SLS estimates, which may suggest that the regional dummies in the 2SLS regressions capture most of the country-specific effects.

We now turn to the Sys-GMM results, which are reported in columns 1 through 3 of Table 3.2.[29] Identification is based on a set of 'internal' as well as 'external' instruments.[30] The validity of the instruments[31] and moment conditions was tested using the Hansen test of over-identifying restrictions and the Arellano-Bond test for autocorrelation. The test results show that the null of no second-order autocorrelation is rejected, which precludes the use of second lags of the endogenous variables as instruments. Hence, we restricted the instrument set to lags 3 and longer of the variables. Table 3.2 shows that all specifications pass the Hansen test, suggesting that the instrument set is valid. The test for the null of no third-order autocorrelation cannot be rejected either.[32] Further, we performed a difference-in-Hansen test for the exogeneity of the subset of additional instruments in the Sys-GMM and found that the specifications cannot be rejected.

The Sys-GMM estimates reveal a substantial degree of inertia in IHDI. The lag of the IHDI is highly significant and has considerable explanatory power, rendering some of the covariates included in the regression insignificant. The coefficient on social spending is positive and significant at the 1 percent level. The estimates imply that a 1 percent increase in social spending, in percent of GDP, increases the IHDI by 0.004 points, which is somewhat lower than those indicated by the 2SLS and FE regressions. Given the inclusion of the lagged dependent variable, it is also possible to calculate the long-run effect of social spending on the IHDI. The estimates in the preferred GMM specification (column 2 of Table 3.2), and coefficients of 0.004 for social spending and 0.93 for lagged IHDI suggest that the long-run effect of a 1 percent increase in social spending is to increase the IHDI by about 0.057 points. The coefficient on GDP per capita suggests that a 1 percent increase in per capita income would increase IHDI by 0.002 points. Modelling

[29] To preserve degrees of freedom, the Sys-GMM regressions exclude age dependency ratio from the regressions. The results when this variable is included can be provided upon request from the authors.

[30] For more details, see the notes at the bottom of Table 3.2.

[31] We opt for the instrument set in collapsed form for a couple of interrelated reasons. First, the number of instruments is sizably lower when the instruments are collapsed than when the lag ranges are limited. Second, in some common cases, collapsing instruments is superior to restricting lag ranges (Roodman, 2009b).

[32] Note that we report only the test results for the nulls of no second- and third-order serial correlations. The null of no autocorrelation of higher orders is not rejected as well.

Table 3.3: Child mortality equation (2SLS and FE regressions)

Dependent variable: *ln*(child mortality rate)					
Regression	(1)	(2)	(3)	(4)	(5)
Method	OLS	2SLS	2SLS	FE	FE
Health spending (*ln*)	−0.136	−0.156	−0.181	−0.076	−0.087
(% GDP)	(9.73)***	(4.77)***	(5.45)***	(2.19)**	(2.56)**
GDP per capita (*ln*)	−0.241	−0.68	−0.215	−0.35	−0.19
	(5.63)***	(22.79)***	(3.10)***	(4.81)***	(2.87)***
Female education (*ln*)	0.003		0.005		−0.011
	(0.36)		(0.50)		(1.78)*
Access to sanitation	−0.012		−0.012		−0.002
	(9.53)***		(6.25)***		(0.57)
Access to safe water	0.005		0.004		−0.004
	(3.09)***		(1.95)*		(1.29)
Degree of urbanization	−0.0005		−0.001		−0.009
	(0.37)		(0.65)		(1.88)*
Fertility rate	0.204		0.224		−0.027
($t-1$)	(8.46)***		(5.64)***		(0.74)
Bureaucratic quality	−0.024		−0.024		−0.025
	(1.27)		(0.89)		(1.27)
Democracy	−0.009		−0.005		0.007
	(2.60)***		(0.89)		(2.60)***
Number of countries	55	55	55	55	44
Observations	537	819	533	643	425
R-squared	0.86	0.78	0.86	0.85	0.91
Kleibergen-Paap F st.		134.36	73.92	21.49	20.1
Stock-Wright LM st.		18.69	19.99	5.89	7.73
(p-value)		0.00	0.00	0.05	0.02
Hansen test (p-value)		0.82	0.85	0.51	0.97

Significance level: *10%; **5%; ***1%

Notes: The 2SLS regressions use log population and agriculture (% GDP) as instruments for health spending; the FE regressions use the second lag of health spending and the military spending of neighboring countries (% of CG spending). Although one may suspect lagged spending to be endog enous, empirical tests do not reject the validity of the instruments.

Source: Authors' analysis based on the data described in Appendix A.

the persistence of IHDI and accounting for the endogeneity of per capita income appear to have resulted in a smaller coefficient on the latter, albeit still highly significant. Turning to the coefficients on the other explanatory variables, trade openness and the terms of trade are positively and negatively associated with the IHDI respectively, although the latter is significant only at the 10 percent level. To check the robustness of the results to alterations in the instrument set, we rerun the regression in column 2 using only internal instruments. Column 3 unveils that the coefficient on social spending remains unaffected, although its significance declined somewhat.

To summarize, social spending has a consistently positive and significant impact on the IHDI. The quantitative impact of social spending implied by the Sys-GMM analysis appears modest, albeit not negligible. Given that there are no previous studies on the effect of social spending on IHDI, it is difficult to place these results in context. Gomanee *et al.* (2005c) is, to the best of our knowledge, the only study that is somehow comparable to ours. The finding that social spending has a significantly positive impact on aggregate welfare (as measured by IHDI as well as HDI (see Section 3.6) is at variance with Gomanee *et al.* (2005c), who find social spending to have an insignificant impact on HDI. However, we believe that our results are plausible because Gomanee *et al.* (2005c) do not resolve, among others, problems revolving around the potential endogeneity of social spending.

3.5.2 Health spending and child mortality

We now turn to the results for child mortality, which are reported in Tables 3.3 and 3.4.[33] Column 1 of Table 3.3 shows the baseline OLS estimates. Most of the coefficients are statistically significant and have the expected sign. The 2SLS results, which are based on regressions that instrument for health spending using the log of population and the share of agriculture in GDP, are shown in columns 2 and 3. The models pass the specification tests and the explanatory variables account for a considerable portion of the variation in child mortality rates. The estimated coefficients in column 3 reveal that health spending and per capita income are important factors explaining the cross-country differentials in child mortality rates. A 1 percent increase in health spending, in percent of GDP, reduces child mortality rate by around 0.18 percent. The coefficient on per capita income is consistent with the robust finding in the literature that income explains a good portion of the variation in child mortality rates across countries and over time (Pritchett and Summers, 1996; Filmer and Pritchett, 1999; Baldacci *et al.*, 2008; Rajkumar and Swaroop, 2008).

Among the other variables, access to improved sanitation is negatively linked with child mortality rate. What is more, access to safe water is positively associated with child mortality, which appears counterintuitive, albeit significant only at the 10 percent level. However, a closer scrutiny reveals the presence of a strong collinearity between access to safe water and access to improved sanitation. Both variables are negatively associated with child mortality rate when they enter the regressions individually, which may suggest that multicollinearity problems rendered the coefficients on these variables imprecise. However, the coefficient on health spending remained stable regardless of whether these variables are allowed to enter the regressions individually or jointly. The positive coefficient on lagged fertility suggests that a decline in fertility rates has a positive impact on child survival rates, which is consistent with Baldacci *et al.* (2008) and other studies examining the fertility-child mortality nexus during demographic transitions (Galor and Weil, 2000; Greenwood and Seshadri, 2002, and Rosenzweig and Schultz, 1983). However, a comparison of the magnitude of the coefficients in columns 2 and 3 reveals that the parameter estimates for some of the variables are relatively unstable, perhaps mainly

[33] The sample size in the child mortality regressions is slightly larger than that in the IHDI regressions because data on child mortality rate are available for all countries in the sample.

due to the presence of multicollinearity problems.[34] Therefore, the 2SLS results should be interpreted cautiously.

It should be noted, however, that unaccounted-for country-specific effects may be biasing the 2SLS results. We thus turn to the FE results in columns 4 and 5.[35] The specification tests indicate that the FE models are by and large well specified. Consistent with the 2SLS results, health spending and per capita income are statistically significant and possess the expected sign. The coefficients indicate that a 1 percent increase in health spending results in a decline in child mortality rate of about 0.09 percent. The same increase in per capita income is associated with an approximately 0.19 percent reduction in child mortality rate. Turning to the other explanatory variables, female education and the degree of urbanization are inversely related to child mortality rates, which are in concordance with our expectations. The negative coefficient on female education suggests that mother's schooling reduces the incidence of child mortality, which is consistent with, among many others, Filmer and Pritchett (1999), Rajkumar and Swaroop (2008), and Baldacci *et al.* (2008). The negative coefficient on the degree of urbanization is in line with Schultz (1993), who finds that child mortality rates are higher for rural, low-income, and agricultural households.

Given that the 2SLS and FE results are not based on dynamic specification and that they may be driven by short-term fluctuations, we test their robustness using dynamic panel data estimators and the three-year averaged data. Columns 1 to 3 of Table 3.4 report the Sys-GMM results. Column 2 uses both internal and external instruments, while column 3 employs only internal instruments.[36] In all specifications, the null hypotheses that the instruments are valid and that there is no serial correlation of order two and higher are not rejected at conventional critical values.

The Sys-GMM estimates for health spending and per capita income are qualitatively similar to those from the 2SLS and FE regressions. The coefficient on the lagged dependent variable indicates that child mortality rate is a highly persistent series and has considerable explanatory power, rendering most of the other covariates insignificant. Column 2 shows that health spending and per capita income have significant negative impact on child mortality rate. The estimates imply that increasing health spending, in percent of GDP, and per capita income by 1 percent would reduce child mortality rate by approximately 0.06 percent and 0.03 percent respectively. Given the inclusion of the lagged

[34] Note, however, that the sample sizes in columns 2 and 3 differ.

[35] The instruments employed in the 2SLS regressions turned out to be weak in the FE specifications and thus we follow Filmer and Pritchett (1999) and Bokhari *et al.* (2007) and use the military expenditures, in percent of total central government expenditures, of neighboring countries as an additional instrument for health spending. Government health spending is a function of own military budget and the latter is supposedly a function of the military budget of neighboring countries. Hence, in the reduced form, health spending is a function of neighboring countries' military spending while it is highly unlikely that the latter is correlated with IHDI.

[36] For more details, see the notes at the bottom of Table 3.4.

Table 3.4: Child mortality equation (System GMM regressions)

Dependent variable	*ln*(Child mortality)		
Regression	(1)	(2)	(3)
Lagged child mortality (*ln*)	0.93 (34.69)***	0.93 (14.92)***	0.95 (12.00)***
Health spending (*ln*) (% GDP)	−0.052 (1.96)**	−0.06 (3.15)***	−0.065 (2.42)**
GDP per capita (*ln*)	−0.025 (3.04)***	−0.028 (2.52)**	−0.054 (1.52)
Female education (*ln*)		−0.002 (0.51)	−0.006 (0.91)
Access to sanitation		−0.001 (0.11)	−0.0004 (0.17)
Access to safe water		0.001 (0.35)	0.001 (0.46)
Degree of urbanization		−0.0001 (0.20)	0.0001 (0.10)
Fertility rate ($t-1$)		0.015 (0.84)	0.0001 (0.00)
Bureaucratic quality		0.003 (0.18)	0.009 (0.31)
Democracy		0.002 (0.82)	0.001 (0.20)
Number of countries	51	44	44
Observations	242	193	195
Number of instruments	26	33	31
Hansen test[a]	0.51	0.29	0.32
Difference-in-Hansen test[a]	0.65	0.37	0.42
Autocorrelation (second order)[a]	0.21	0.82	0.77
Autocorrelation (third order)[a]	0.21	0.13	0.15

[a]Denotes *p*-value. Significance level: *10%; **5%; ***1%

Notes: All except column (3) use internal (second and longer lags of child mortality, health spending, and GDP per capita) and external (log population and agriculture (% GDP)) instruments; column 3 uses only the aforementioned internal instruments.

Source: Computed based on the data described in Appendix A.

dependent variable, it is also possible to calculate the long-run effect of health spending on child mortality. The estimates in the preferred GMM specification (column 2 of Table 3.4) and coefficients of 0.06 for health spending and 0.93 for lagged health spending suggest that the long-term effect of a 1 percent increase in health spending, as a percentage of GDP, is to lower child mortality rate by about 0.86 percent. Column 3 shows that these results are fairly robust to using only internal instruments, although income per capita loses its significance.

All in all, the Sys-GMM results support the proposition that government health spending and income per capita are important contributing factors to reducing the incidence of child mortality.

3.6 Robustness Checks

In this section, we run several robustness checks to test the validity of our results. Sections 3.6.1 and 3.6.2 rerun the main regressions using alternative samples and estimators respectively. Section 3.6.3 redefines social spending and aggregate welfare. Section 3.6.4 uses alternative specifications, while Section 3.6.5 expands the set of control variables. Section 3.6.6 disaggregates social spending. Finally, Section 3.6.7 tests whether the efficacy of social spending hinges on democratic governance. Tables 3.6 − 3.20 in Appendix B report the results. To economize on space, we report only the coefficients of social spending and per capita income. To facilitate understanding, Table 3.5 summarizes Tables 3.6 to 3.20 and reports only the coefficients to social spending (and health spending in the child mortality equation).

3.6.1 Alternative samples

Five of the countries in our dataset (Argentina, Chile, Latvia, Lithuania, and Uruguay) are high-income (HI) with high IHDI and very low child mortality rates, and their inclusion may imply a downward bias for the effectiveness of social spending. Hence, to make sure that these countries are not skewing the results, we rerun our main IHDI and child mortality regressions (columns 3 and 5 of Tables 3.1 and 3.3, and column 2 of Tables 3.2 and 3.4) using a restricted sample that excludes them. As can be seen from columns 1 through 3 of Tables 3.6 and 3.7, the previous result that government social (and health) spending has significant positive impact on IHDI (and negative impact on child mortality in the case of health spending) remains intact when the supposedly 'outlying' countries are omitted.

Moreover, we tested the sensitivity of our results to including only middle income (MI) countries, which constitute about two-thirds of the sample. Columns 4 through 6 of Tables 3.6 and 3.7 report the IHDI and child mortality regression results for the MI subsample. We find no considerable difference in the effect of social spending when the sample is restricted to include only MI countries.

As discussed in Section 3.3, for some of the sample countries, the social spending data were reported partly in cash and partly in accrual, and thus we extended the cash data using the growth rates for the accrual data. To make sure that these countries are not systematically influencing the results, we re-estimate the baseline models by confining the analysis to include only the countries for which we have complete cash data. In addition, the social spending data for some of the countries in the sample had gaps and we imputed the missing observations. We now check the sensitivity of the results to excluding these countries. Tables 3.8 and 3.9 show that the previous finding that social spending improves aggregate welfare is sufficiently robust. As an additional way of checking the sensitivity of the results, we follow Seiferling (2013) and include an indicator (dummy) variable to account for any systematic differences between the cash and accrual data. The estimates (not reported) reveal that there exists no discernible difference between the two.

In sum, the main results for the full sample hold for the subsamples as well. However, splitting the sample considerably reduces the sample size, which may render the empirical estimates unreliable. In some of the regressions, the instruments are 'weak' and in some others the p-values of the Hansen test are relatively small, suggesting that caution should be taken in interpreting the results for the subsamples.

3.6.2 Alternative estimators

To check whether the core results of this paper hold across different estimators, we rerun our models using an alternative set of estimation methods: LIML, Fuller-LIML, and CUE (see Section 3.4.2). Columns $5-7$ of Tables 3.10 and 3.11 show that the key results prove robust to the use of different estimators.

3.6.3 Redefining social spending and aggregate welfare

In the analysis presented above, social spending has been defined relative to a country's GDP, which provides a reasonable measure of government social spending relative to a country's overall resources. Per capita social spending is an alternative measure, which is often deemed more appropriate in assessing the efficacy of social sector spending with respect to specific MDGs. Hence, as a robustness check, we rerun the core regressions using spending per capita instead of spending in percent of GDP. Columns 1 to 4 of Tables 3.10 and 3.11 show the results for IHDI and child mortality respectively. We find that the 2SLS, FE, and Sys-GMM results are broadly consistent with those found previously.

So far, our focus has been on the nexus between social spending and the *inequality-adjusted* HDI. Table 3.12 (columns 1 through 4) shows the results when IHDI is replaced with the *conventional* HDI in the 2SLS, FE, and Sys-GMM specifications. Using HDI leads to qualitatively similar results, although social spending has now a larger impact. This may suggest that distributional inequalities in education, health, and income reduce the effectiveness of government spending in improving aggregate welfare. However, the issue of inequality is not a trivial one and requires further analysis, which is beyond the scope of this paper.

3.6.4 Alternative specifications

An alternative way to assess the effectiveness of social spending is to investigate its impact on the rates of change in IHDI and child mortality rate rather than simple variation in the variables themselves. Such changes are likely to capture the long-term effect of social spending and may better reflect a country's progress towards sustainable development. Thus, we reestimate our main Sys-GMM regressions (columns 2 and 3 of Tables 3.2 and

3.4) using the growth rates of IHDI and child mortality rates as dependent variables.[37] The results are presented in columns 4 and 5 of Table 3.20. Consistent with our previous result, social spending has a significantly positive impact on IHDI growth. The coefficient on the initial level of IHDI is negative and significant, suggesting a convergence effect: countries with low IHDI to begin with experience higher increases in IHDI, *ceteris paribus*. The results for child mortality are also in line with our previous findings.

Columns 1 to 3 of Table 3.20 present the results when IHDI is expressed in logs rather than levels. The estimated coefficients suggest that the main findings of this paper are not sensitive to whether the variables of primary interest are expressed in logs or levels.

3.6.5 Additional control variables

To ensure that omitted variables are not biasing our results, we expand the set of control variables with some potential determinants of IHDI and child mortality. In the IHDI regressions, the additional control variables include: aid flows, in percent of GNI[38]; Gini coefficient, measuring income inequality; the share of domestic credit to private sector in GDP; and corruption index. In the child mortality regressions, the set of additional control variables comprises: aid in percent of GNI, the percentage of population aged under 5, the prevalence of undernutrition, ethnolinguistic fractionalization, and the percentage of the population that is Muslim.[39]

Tables 3.13 to 3.16 show the results when these variables are added to the baseline IHDI and child mortality regressions. The key 2SLS, FE, and Sys-GMM results remain unaffected.[40] Only some of the additional controls are significant at conventional levels. The finding that aid has an insignificant effect on IHDI does not come as much of a surprise. Most of the countries in the sample are middle-income and receive far less official development assistance than do low-income countries.

Turning to the child mortality regressions, the inclusion of additional controls leaves the main results broadly unaffected.[41] Aid (lagged one-period) has a significant negative effect on child mortality rate in the FE specification.[42] Moreover, the positive coefficients on ethnolinguistic fractionalization and 'predominantly Muslim' are consistent

[37] We focus on the GMM specifications, which use the three-year averaged data, because of the limited variation in the annual data.

[38] We include aid lagged one-period to minimize endogeneity problems.

[39] We allow these variables to enter the regressions individually because of the limited overlap between them, which would have led to a considerable loss of degrees of freedom.

[40] Including the Gini coefficient reduces the sample size substantially, which may explain the decline in the significance of social spending in some of the regressions and the decline in instrument strength in some others.

[41] In the FE regression that includes aid, the instruments in our baseline regression are unreliable and thus we use lagged health spending and military spending, in percent of central government expenditure, as instruments.

[42] In contrast, aid has a positive and significant (at the 10 percent level) coefficient in the 2SLS regression. However, the FE result seems plausible given that the 2SLS regression does not control for country fixed-effects.

Table 3.5: Summary of Tables 3.6 − 3.20: Coefficients to social and health spending

Sample countries	Excluding HI countries			Only MI countries		
Method	2SLS	FE	GMM	2SLS	FE	GMM
Table 3.6: IHDI equation: Alternative samples						
Social spending	0.009	0.014	0.004	0.008	0.024	0.003
(% GDP) (*ln*)	(3.41)***	(2.70)***	(2.46)**	(1.94)*	(2.54)**	(3.34)***
Table 3.7: Child mortality equation: Alternative samples						
Health spending	−0.19	−0.09	−0.056	−0.168	−0.074	−0.078
(% GDP) (*ln*)	(5.17)***	(2.66)***	(3.64)***	(4.51)***	(2.11)**	(4.06)***

Sample countries	Only cash data			No data gaps		
Method	2SLS	FE	GMM	2SLS	FE	GMM
Table 3.8: IHDI equation: Alternative samples						
Social spending	0.012	0.015	0.003	0.014	0.01	0.003
(% GDP) (*ln*)	(5.83)***	(2.47)**	(2.02)**	(3.72)***	(2.67)***	(2.09)**
Table 3.9: Child mortality equation: Alternative samples						
Health spending	−0.196	−0.138	−0.059	−0.235	−0.064	−0.06
(% GDP) (*ln*)	(4.39)***	(3.14)***	(3.18)***	(7.08)***	(1.82)*	(3.64)***

Method	2SLS	FE	GMM
Table 3.10: IHDI equation: Social spending per capita			
Social spending	0.014	0.014	0.004
per capita (*ln*)	(6.12)***	(2.54)**	(2.91)***
Tables 3.11: Child mortality equation: Health spending per capita			
Health Spending	−0.181	−0.095	−0.052
per capita (*ln*)	(5.45)***	(2.72)***	(3.42)***

Table 3.12: HDI equation			
Method	2SLS	FE	GMM
Social spending	0.033	0.023	0.013
(% GDP) (*ln*)	(3.66)***	(2.31)**	(2.32)**

Tables 3.13−3.16: IHDI and child mortality equations: Additional control variables

Dependent var.	IHDI			Depen. var.	Child mortality		
Method →	2SLS	FE	GMM	Method →	2SLS	FE	GMM
Control variables↓	Coeffic. to social spending↓			Control var↓	Coeffic. to health spending↓		
Aid (% GNI)	0.014	0.015	0.003	Under-5 pop.	−0.17	−0.103	−0.05
	(6.31)***	(2.68)***	(1.82)*		(5.17)***	(2.70)***	(2.97)***
Gini	0.013	0.012	0.002	Aid (% GNI)	−0.139	−0.055	−0.063
	(4.81)***	(3.49)***	(1.73)*		(3.63)***	(2.17)**	(2.12)**
Finance	0.013	0.012	0.004	Undernutrit.	−0.139	−0.048	−0.068
	(5.82)***	(2.65)***	(2.66)***		(3.63)***	(2.31)**	(3.24)***
Corruption	0.013	0.013	0.004	Fractionaliz.	−0.184		−0.061
	(5.07)***	(2.69)***	(2.14)**		(5.64)***		(3.24)***
				Pred. Muslim	−0.171		−0.059
					(4.75)***		(3.13)***

Method	2SLS	FE	GMM
Table 3.17: IHDI equation: Spending, welfare, and democracy			
Social spending	0.015	0.011	0.004
(% GDP) (*ln*)	(4.91)***	(2.46)**	(1.91)*
Table 3.18: Child mortality equation: Spending, welfare, and democracy			
Health spending	−0.20	−0.104	−0.058
(% GDP) (*ln*)	(5.18)***	(2.87)***	(2.60)***

Table 3.20: IHDI equation: Alternative specifications

Dependent var.	*ln*(IHDI)			Δln(IHDI)
Method	2SLS	FE	GMM	GMM
Social spending	0.064	0.068	0.02	0.021
(% GDP) (*ln*)	(5.10)***	(2.34)**	(2.35)**	(2.63)***

[a]Denotes *p*-value. Significance level: *10%; **5%; ***1%

with, among others, Filmer and Pritchett (1999) and Rajkumar and Swaroop (2008), who show that child mortality rates are higher in countries that are predominantly Islamic and countries with higher ethnolinguistic diversity. Note, however, that some of these results should be interpreted with caution because the instruments appear 'weak' in some of the regressions.

3.6.6 Disaggregating social spending

So far, we have focused on the effectiveness of government social spending, which aggregates expenditures on health, education, and social protection. In this subsection, we investigate the potential differential effects of the different components of social spending on aggregate welfare. Table 3.19 reports the Sys-GMM results. We find that health spending has a significantly positive impact on IHDI, whereas the coefficients on education and social protection expenditures appear insignificant at conventional levels. This result is robust to changes in the set of instruments. This suggests that the previous finding that social spending affects IHDI significantly positively is due to government expenditure on health. However, the number of instruments is quite large when government social spending is disaggregated. Moreover, considering the small size of our sample, the number of variables is large. These suggest caution in interpreting the results.

3.6.7 Social spending, aggregate welfare and democracy

We now consider the possibility that the efficacy of social spending might hinge on democratic governance. A strand of the literature contends that government spending tends to be more effective in countries with democratic institutions that provide an institutionalized check on governments (Svensson, 1999). This suggests that the impact of government spending is stronger the higher the level of democracy. Tables 3.17 and 3.18 add the interaction terms $spending \times democracy$ and $health\ spending \times democracy$ to the baseline IHDI and child mortality regressions respectively.[43] Our approach is similar to that of Burnside and Dollar (2000), who employ interaction terms to answer the question of whether aid has a stronger impact on growth in countries with better policies.

In the IHDI regression, the interaction term enters the 2SLS and Sys-GMM regressions insignificantly, whereas it is significantly negative in the FE specification. In all specifications, social spending has a significantly positive effect on the IHDI. In the child mortality regression, health spending has a consistently negative effect on child mortality rate, while the corresponding interaction term is significant in some of the regressions. To make sure that these results are not artifact of the Polity democracy index, we rerun the regressions using the ICRG index on democratic accountability. We find similar results (not reported). Overall, the results show that there is no strong evidence to suggest that democratic institutions are *sine qua non* for the effective contribution of social spending

[43] In line with most previous studies, democracy is treated as exogenous to government spending.

to aggregate welfare. In other words, government spending on the social sectors may improve aggregate human welfare even in countries with less-advanced democratic institutions. This is consistent with, *inter alia*, Ames (1987), Alesina and Rodrik (1994), and Alesina *et al.* (1996), who contend that even autocratic regimes ultimately need to respond to the demands of their people to avoid social unrest, market disruption, and eventually being overthrown. Note, however, that a more thorough analysis of the nexus between social spending, aggregate welfare, and democracy is necessary before attempting to draw any far-reaching conclusions.

3.7 Conclusions

Despite the widespread attention devoted to promoting human development via increasing social sector spending, the evidence on the relationship between social spending and welfare outcomes remains inconclusive. In this paper, we investigated the impact of government spending on social sectors (health, education, and social protection) on two major indicators of aggregate welfare (the *Inequality-adjusted* Human Development Index (IHDI) and child mortality) using a panel dataset covering 55 developing and transition countries from 1990 to 2009.

Our study provides strong evidence to support the proposition that social spending is a strong predictor of improved aggregate welfare. More specifically, we find that government social spending has a significantly positive causal effect on the IHDI, while government expenditure on health has a significant negative impact on child mortality rate. The preferred (System GMM) specification (column 2 in Table 3.2) indicates that a 1 percent increase in social spending, in percent of GDP, increases IHDI by 0.004 points, which appears modest, albeit not negligible. The implied long-run effect of a similar increase in social spending is an increase in the IHDI of about 0.057 points. The main findings are fairly robust to the method of estimation, the use of alternative instruments to control for the endogeneity of social spending, the set of control variables included in the estimations, and the use of alternative samples.

Moreover, we find results indicating that health spending has a significant positive impact on the IHDI, whereas this is not the case for education and social protection expenditures. It is unclear whether this is due to data limitations or the intrinsic nature, in terms of scope and scale, of the components of government expenditure. Hence, a potential avenue for future research would be to analyze the potential differential effects of the different components of social spending in a rigorous manner. In other words, further work would be needed to have a clearer picture of what kind of social spending works, what does not and why.

Finally, there is no strong evidence that the efficacy of social spending hinges on democratic governance, suggesting that government spending on the social sectors may improve aggregate welfare even in countries with less-advanced democratic institutions.

This is in line with, among others, Ames (1987), Alesina and Rodrik (1994), and Alesina *et al.* (1996), who point out that even autocratic regimes ultimately need to respond to the demands of their people to avoid social unrest, market disruption, and eventually being overthrown. The recent developments in Northern Africa may be illustrative in that respect. Thus, whereas democratic freedoms and the opportunity for people to control their leaders are valid aspirations, there seems to be no apparent reason why social sector policy could not improve aggregate welfare even in contexts of less-advanced democratic institutions.

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

Ł

ŁŁ

Appendices

Appendix A

Table A. Data description and source

Variable	Description and source
GDP per capita	Gross Domestic Product per capita, PPP (constant 2005 international dollar). Source: World Development Indicators (WDI)
Child mortality rate	The number of new born babies out of 1000 that die before reaching the age of 5, if subject to current age-specific mortality rates. Source: WDI
Social spending	Central government (CG) spending (current and capital) on health, edu cation, and social protection. Source: GFS database (2011 edition)
Openness	Sum of exports and imports of goods & services (% GDP). Source: WDI
Terms of trade	The ratio of export price index to import price index. Source: WDI
Inflation	Log of one plus the inflation rate. Source: WDI
Age dependency ratio	The ratio of dependents (people younger than 15 or older than 64) to working age population (those aged 15-64). Source: WDI
Bureaucratic quality	Assesses how much strength and expertise bureaucrats have to govern without drastic alterations in policy or interruptions in government services. Scale is from 0 to 4. Source: ICRG.
Corruption	Index measuring corruption in government, based on the subjective evaluations of experts. Scale is from 0 to 6. Source: ICRG
Democracy index	Comprises two components: democracy (Dc) and autocracy (Ac), ranging from 0 to 10, where 10 corresponds to full democracy and full autocracy respectively. The index is computed by subtracting Ac from Dc and is thus measured on a -10 to 10 scale. Source: Marshall *et al.* (2013).
Aid	Net Official Development Assistance (% GNI). Source: WDI
Gini	Gini coefficient. Source: UNU-WIDER (2013)
Finance	Domestic credit to private sector (% GDP). Source: WDI
Female education	Share of female students in primary & secondary schools. Source: WDI
Degree of urban ization	Percentage of population living in areas defined as urban by each country. Source: WDI
Sanitation	Access to improved sanitation facilities (% population) Source: WDI.
Safe water	Access to improved drinking water source (% population) Source: WDI.
Fertility rate	Number of births per woman. Source: WDI
Democratic acc ountability	This index ranges from 1 to 6 and measures how responsive government is to its people. Source: ICRG
Population	Total population. Source: WDI
Under-5 population	Percentage of population aged under 5. UN Population Division (2013).
Muslim	Percentage of population that is Muslim. Source: La Porta *et al.* (1999)
Undernutrition	Prevalence of undernourishment (% population). Source: WDI
Ethnolinguistic fractionalization	The probability that two randomly selected individuals from a popula tion do not belong to the same ethnolinguistic group. Source: Alesina *et al.* (2003)
Ethnic tensions	The degree of tension within a country stemming from differences in race, nationality, or language: the higher the tension, the lower the rating. Scale is from 1 to 6. Source: ICRG
Agriculture	Agriculture, value added (% GDP). Source: WDI
Military spending	Military expenditures (% total CG spending) Source: WDI
Budget deficit	CG budget deficit (% GDP). Source WDI (2013) and GFS (2011)

Table B. Summary statistics

Panel A: IHDI specification					
Variable	Obs.	Mean	Std. dev.	Min	Max
GDP per capita (2005 PPP)	1017	5360.51	3923.98	433.76	21190.58
Social spending (% of GDP)	785	9.29	6.20	0.14	46.01
HDI	982	0.595	0.133	0.259	0.813
IHDI	962	0.199	0.044	0.086	0.271
Terms of trade	846	104.8	24.3	50.93	251.85
Openness	1007	76.04	39.95	0.31	223.06
Gini coefficient	512	42.43	9.08	20.5	65.77
Inflation	952	28.93	190.75	-17.63	4734.92
Age dependency ratio	1020	67.14	17.83	38.95	120.82
Bureaucratic quality	840	1.89	0.75	0	3.5
Democracy	972	3.95	5.58	-9	10
Panel B: Child mortality specification					
GDP per capita (2005 PPP)	1097	5382.322	3874.504	433.76	21190.58
Health spending (% of GDP)	830	1.88	1.78	0.03	20.82
Child mortality rate	1100	51.49	41.98	6.6	204
Degree of urbanization	1100	49.39	21.82	6.27	92.35
Female education	931	49.4	21.82	30.49	53.97
Fertility rate	1094	47.79	3.17	1.09	8.67
Access to sanitation	1059	68.04	26.39	2.3	100
Access to safe water	1078	83.15	17.43	13.6	100

Table C. Countries with data partly in cash and partly in accrual, and data gaps

Country	Years available		Data gaps	Country	Years available		Data gaps
	Cash	Accrual			Cash	Accrual	
Algeria	2000-2005	2006-09	None	Indonesia	1990-2004	None	2000
Argentina	1990-2001	2002-04	None	Lebanon	1993-99	2000-09	None
Bolivia	1990-2001	2002-07	None	Lithuania	1999-2000	2001-09	None
Chile	1990-2002	2003-09	None	Madagascar	1990-97	2001-09	1998-2001
Costa Rica	1990-2007	None	1992-93; 2004	Malaysia	1990-2009	None	1996-2001
Domincan Republic	1990-2003	None	2001	Mongolia	1992-2002	None	1999
El Salvador	1990-2001	2002-09	None	Morocco	1990-99	None	1996
Ethiopia	1990-2005	None	2000	Romania	1990-2001	2002-07	None
Fiji	1990-1996; 2004-2006	None	1997-2003	Thailand	1990-2004	2005-09	None
				Zambia	1990-2007	None	2000

Panel B. List of countries included in the sample

Algeria, Argentina, Bangladesh, Belarus, Bhutan, Bolivia, Bulgaria, Burundi, Chile, Columbia, Costa Rica, Dominican Republic, Egypt, El Salvador, Ethiopia, Fiji, Georgia, Guatemala, India, Indonesia, Iran, Jamaica, Jordan, Kazakhstan, Kenya, Kyrgyz Republic, Latvia, Lebanon, Lithuania, Madagascar, Malaysia, Maldives, Mauritius, Mexico, Moldova, Mongolia, Morocco, Myanmar, Nepal, Pakistan, Panama, Papua New Guinea, Philippines, Romania, Russian Federation, Seychelles, Sri Lanka, Thailand, Tunisia, Turkey, Uganda, Ukraine, Uruguay, Yemen, Zambia.

Table D. Simple correlations: selected variables

Panel A: IHDI equation								
	GDP*pc*	SS	IHDI	ToT	Trade	Inflation	BQ	Democ.
GDP*pc*	1							
SS	0.38	1						
IHDI	0.82	0.45	1					
ToT	−0.03	0.02	−0.07	1				
Trade	0.24	0.32	0.38	−0.09	1			
Inflation	−0.06	−0.06	−0.05	0.09	−0.21	1		
BQ	0.33	−0.12	0.24	−0.12	0.17	−0.18	1	
Democracy	0.35	0.23	0.40	−0.16	0.21	0.04	0.17	1
Panel B: Child mortality (CM) equation								
	GDP*pc*	HS	CM	Urban.	Educ.	Fertility	Sanit.	Water
GDP*pc*	1							
HS	0.19	1						
CM	−0.69	−0.23	1					
Urbanization	0.69	0.09	−0.69	1				
Education	0.36	0.20	−0.57	0.46	1			
Fertility	−0.61	−0.10	0.85	−0.61	−0.51	1		
Sanitation	0.65	0.15	−0.81	0.70	0.51	−0.71	1	
Water	0.60	0.14	−0.78	0.66	0.39	−0.70	0.83	1

Notes: GDP*pc* stands for GDP per capita; SS for social spending; ToT for terms of trade; BQ for bureaucratic quality; HS for health spending; and Education for female education.

Appendix B

Table 3.6: IHDI equation: Alternative samples

Dependent variable: IHDI						
Sample countries	Excluding HI countries			Only MI countries		
Regression	(1)	(2)	(3)	(4)	(5)	(6)
Method	2SLS	FE	Sys-GMM	2SLS	FE	Sys-GMM
Social spending	0.009	0.014	0.004	0.008	0.024	0.003
(% GDP) (*ln*)	(3.41)***	(2.70)***	(2.46)**	(1.94)*	(2.54)**	(3.34)***
Lagged IHDI			0.95			0.96
			(30.89)***			(24.98)***
GDP per capita	0.026	0.018	0.002	0.038	0.02	0.002
(*ln*)	(12.23)***	(5.50)***	(2.25)**	(9.14)***	(2.89)***	(1.93)*
Number of countries	47	34	38	25	19	33
Observations	461	374	160	246	201	141
R-squared	0.91	0.89		0.90	0.79	
Number of instruments	2	2	33	2	2	33
Kleibergen-Paap F st.	35.24	6.98		13.72	2.8	
Stock-Wright LM st.	8.23	7.37		2.90	10.00	
(p-value)	0.016	0.025		0.23	0.007	
Hansen test[a]	0.58	0.59	0.76	0.56	0.60	0.30
Diff-in-Hansen test[a]			0.57			0.76
Autocor. (2^{nd} order)[a]			0.02			0.03
Autocor. (3^{rd} order)[a]			0.86			0.62

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.7: Child mortality equation: Alternative samples

Dependent variable: *ln*(child mortality rate)						
Sample countries	Excluding HI countries			Only MI countries		
Regression	(1)	(2)	(3)	(4)	(5)	(6)
Method	2SLS	FE	Sys-GMM	2SLS	FE	Sys-GMM
Health spending (% GDP) (*ln*)	−0.19 (5.17)***	−0.09 (2.66)***	−0.056 (3.64)***	−0.168 (4.51)***	−0.074 (2.11)**	−0.078 (4.06)***
Lagged Child mor. (*ln*)			0.90 (15.60)***			0.91 (16.57)***
GDP per capita (*ln*)	−0.198 (2.49)**	−0.157 (2.03)**	−0.026 (2.14)**	−0.251 (3.18)***	0.123 (1.56)	−0.029 (2.74)***
Number of countries	50	39	39	26	21	33
Observations	478	374	172	259	208	145
R-squared	0.84	0.91		0.89	0.92	
Number of instruments	2	2	33	2	2	33
Kleibergen-Paap F st.	65.41	21.11		69.00	11.71	
Stock-Wright LM st.	16.97	8.30		14.17	6.84	
(p-value)	0.00	0.016		0.003	0.03	
Hansen test[a]	0.64	0.82	0.42	0.94	0.17	0.28
Diff-in-Hansen test[a]			0.62			0.06
Autocorr. (2^{nd} order)[a]			0.85			0.20
Autocorr. (3^{rd} order)[a]			0.14			0.27

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.8: IHDI equation: Alternative samples

Dependent variable: IHDI						
Sample countries	Only cash data[b]			No data gaps[c]		
Regression	(1)	(2)	(3)	(4)	(5)	(6)
Method	2SLS	FE	Sys-GMM	2SLS	FE	Sys-GMM
Social spending (% GDP) (*ln*)	0.012 (5.83)***	0.015 (2.47)**	0.003 (2.02)**	0.014 (3.72)***	0.01 (2.67)***	0.003 (2.09)**
Lagged IHDI			0.96 (34.73)***			0.94 (32.93)***
GDP per capita (*ln*) ($t-1$)	0.025 (12.19)***	0.018 (4.42)***	0.002 (2.23)**	0.026 (10.60)***	0.021 (6.83)***	0.002 (2.10)**
Number of countries	40	29	35	39	28	33
Observations	386	329	144	378	316	141
R-squared	0.94	0.87		0.89	0.91	
Number of instruments	2	2	33	2	2	33
Kleibergen-Paap F st.	27.46	7.4		24.32	12.69	
Stock-Wright LM st.	16.22	6.58		9.93	6.27	
(p-value)	0.00	0.037		0.007	0.043	
Hansen test (p-value)	0.35	0.58	0.52	0.93	0.78	0.35
Diff-in-Hansen test[a]			0.48			0.48
Autocorr. (2^{nd} order)[a]			0.01			0.02
Autocorr. (3^{rd} order)[a]			0.94			0.38

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Notes: [b]Countries for which we have complete cash-based data on government spending; [c]Countries with no data gap.

Table 3.9: Child mortality equation: Alternative samples

Dependent variable: *ln*(child mortality rate)						
Sample countries	Only cash data[b]			No data gaps[c]		
Regression	(1)	(2)	(3)	(4)	(5)	(6)
Method	2SLS	FE	Sys-GMM	2SLS	FE	Sys-GMM
Health spending (% GDP) (*ln*)	−0.196 (4.39)***	−0.138 (3.14)***	−0.059 (3.18)***	−0.235 (7.08)***	−0.064 (1.82)*	−0.06 (3.64)***
Lagged Child mor. (*ln*)			0.94 (13.08)***			0.93 (12.90)***
GDP per capita (*ln*)	−0.341 (4.27)***	−0.258 (3.17)***	−0.03 (2.24)**	−0.097 (1.44)	−0.292 (4.97)***	−0.026 (2.32)**
Number of countries	43	34	36	43	33	35
Observations	398	322	155	415	329	154
R-squared	0.85	0.90		0.84	0.92	
Number of instruments	2	2	33	2	2	33
Kleibergen-Paap F st.	73.58	14.65		86.59	15.26	
Stock-Wright LM st.	13.22	9.86		28.70	3.65	
(p-value)	0.001	0.007		0.00	0.16	
Hansen test[a]	0.51	0.90	0.28	0.23	0.72	0.17
Diff-in-Hansen test[a]			0.14			0.18
Autocorr. (2^{nd} order)[a]			0.75			0.66
Autocorr. (3^{rd} order)[a]			0.22			0.25

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Notes: [b]Countries with complete cash-based data. [c]Countries with no data gap.

Table 3.10: IHDI equation: Alternative estimators and specification

Dependent variable: IHDI							
Regression	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Method	2SLS	FE	GMM	GMM	CUE	LIML	Fuller
Lagged IHDI			0.86 (19.57)***	0.85 (17.38)***			
Social spending per capita (*ln*)	0.014 (6.12)***	0.014 (2.54)**	0.004 (2.91)***	0.004 (3.18)***			
Social spending (% GDP) (*ln*)					0.014 (6.10)***	0.014 (6.01)***	0.014 (6.03)***
GDP per capita (*ln*)	0.012 (3.43)***	0.005 (0.95)	0.002 (3.27)***	0.002 (1.71)*	0.027 (13.91)***	0.012 (13.65)***	0.027 (13.68)***
Number of Countries	52	38	42	42	52	52	52
Observations	504	417	175	175	504	504	504
R-squared	0.91	0.89			0.91	0.91	0.91
Number of instruments	2	2	33	30	2	2	2
Kleibergen-Paap F st.	42.98	6.91			42.17	42.17	42.17
Stock-Wright LM st.	18.80	6.13			18.80	18.80	18.80
(p-value)	0.00	0.047			0.00	0.00	0.00
Hansen test[a]	0.74	0.69	0.52	0.44	0.77	0.77	0.77
Diff-in-Hansen test[a]			0.60	0.15			
Autocor. (2^{nd} order)[a]			0.02	0.02			
Autocor. (3^{rd} order)[a]			0.86	0.86			

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.11: Child mortality equation: Alternative estimators and specification

Dependent variable: ln(child mortality rate)							
Regression	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Method	2SLS	FE	GMM	GMM	CUE	LIML	Fuller
Lagged Child mor.			0.86 (13.96)***	0.93 (8.98)***			
Health Spending per capita (ln)	−0.181 (5.45)***	−0.095 (2.72)***	−0.052 (3.42)***	−0.059 (2.30)**			
Health spending (% GDP) (ln)					−0.18 (5.51)***	−0.181 (5.45)***	−0.181 (5.46)***
GDP per capita	−0.034 (0.38)	−0.092 (1.18)	−0.02 (1.76)*	−0.034 (0.98)	−0.219 (3.29)***	−0.215 (3.10)***	−0.216 (3.11)***
Number of countries	55	44	44	44	55	55	55
Observations	533	425	193	195	533	533	533
R-squared	0.86	0.91			0.86	0.86	0.86
Number of instrum.	2	2	33	31	2	2	2
Kleibergen-Paap F	73.92	20.02			73.92	73.92	73.92
Stock-Wright LM	19.99	8.74			19.99	19.99	19.99
(p-value)	0.00	0.01			0.00	0.00	0.00
Hansen test[a]	0.85	0.91	0.32	0.30	0.85	0.85	0.85
Diff-in-Hansen test[a]			0.13	0.37			
Autoco. (2^{nd} order)[a]			0.94	0.59			
Autoco. (3^{rd} order)[a]			0.10	0.19			

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.12: HDI equation

Dependent variable: HDI				
Regression	(1)	(2)	(3)	(4)
Method	2SLS	FE	Sys-GMM	Sys-GMM
Lagged HDI			0.93 (35.55)***	0.93 (18.75)***
Social spending (% GDP) (ln)	0.033 (3.66)***	0.023 (2.31)**	0.013 (2.32)**	0.015 (1.87)*
GDP per capita (ln)	0.083 (13.08)***	0.055 (6.11)***	0.007 (2.62)***	0.007 (1.84)*
Number of countries	52	39	42	42
Observations	518	431	175	175
R-squared	0.93	0.91		
Number of instruments	2	2	33	30
Kleibergen-Paap F stat.	18.64	13.50		
Stock-Wright LM stat.	7.25	5.77		
(p-value)	0.027	0.056		
Hansen test[a]	0.84	0.48	0.36	0.26
Diff-in-Hansen test[a]			0.29	0.23
Autocorr. (2^{nd} order)[a]			0.79	0.56
Autocorr. (3^{rd} order)[a]			0.13	0.31

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Notes: Column 3 uses internal as well as external instruments, while column 4 employs only internal instruments.

Table 3.13: IHDI equation: Additional explanatory variables

Dependent variable: IHDI								
Method	2SLS				FE			
Regression	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Social spending (% GDP) (*ln*)	0.014 (6.31)***	0.013 (4.81)***	0.013 (5.82)***	0.013 (5.07)***	0.015 (2.68)***	0.012 (3.49)***	0.012 (2.65)***	0.013 (2.69)***
GDP per capita (*ln*) ($t-1$)	0.024 (9.33)***	0.029 (18.31)***	0.027 (14.57)***	0.027 (14.13)***	0.022 (6.97)***	0.016 (5.75)***	0.024 (6.65)***	0.017 (5.85)***
Aid (% GNI) ($t-1$)	0.000 (0.02)				0.0001 (0.97)			
Gini (*ln*)		−0.007 (1.28)				0.006 (2.34)**		
Finance			−0.004 (3.12)***				−0.003 (3.59)***	
Corruption				0.003 (2.35)**				0.001 (1.36)
Number of count.	52	52	52	52	33	31	38	38
Observations	451	279	503	504	364	239	417	417
R-squared	0.90	0.93	0.92	0.92	0.89	0.92	0.90	0.90
Kleibergen-Paap F	49.64	18.97	45.69	35.79	7.38	11.44	9.77	9.40
Stock-Wright LM	19.70	21.12	17.63	15.78	7.68	9.78	6.57	6.81
(p-value)	0.00	0.00	0.00	0.00	0.02	0.008	0.04	0.03
Hansen test[a]	0.95	0.10	0.96	0.66	0.33	0.57	0.59	0.64

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.14: IHDI equation: Additional controls

Dependent variable: IHDI				
Method	System GMM regressions			
Regression	(1)	(2)	(3)	(4)
Lagged IHDI	0.94 (36.83)***	0.96 (34.49)***	0.95 (30.80)***	0.92 (37.43)***
Social spending (% GDP) (*ln*)	0.003 (1.82)*	0.002 (1.73)*	0.004 (2.66)***	0.004 (2.14)**
GDP per capita (*ln*)	0.002 (1.60)	0.011 (1.61)	0.002 (2.68)***	0.002 (2.35)**
Aid (% GNI) ($t-1$)	−0.0001 (0.32)			
Gini (*ln*)		−0.002 (0.86)		
Finance			−0.002 (2.35)**	
Corruption				0.0003 (0.40)
Number of countries	37	40	42	42
Observations	153	144	175	175
Number of instruments	34	34	34	34
Hansen test (p-value)	0.58	0.56	0.53	0.61
Diff-in-Hansen test[a]	0.38	0.44	0.59	0.52
Autocorr. (second-order)[a]	0.02	0.03	0.03	0.03
Autocorr. (third-order)[a]	0.65	0.86	0.86	0.74

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.15: Child mortality equation: Additional controls

Dependent variable: ln(child mortality rate)								
Method	2SLS					FE		
Regression	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Health Spending (% GDP) (ln)	−0.17 (5.17)***	−0.139 (3.63)***	−0.139 (3.63)***	−0.184 (5.64)***	−0.171 (4.75)***	−0.103 (2.70)***	−0.055 (2.17)**	−0.048 (2.31)**
GDP per capita (ln)	−0.217 (2.54)**	−0.213 (2.52)**	−0.213 (2.52)**	−0.217 (3.06)***	−0.239 (3.44)***	−0.198 (2.92)***	−0.192 (2.25)**	−0.186 (2.63)***
Under-5 pop.	0.096 (3.65)***					−0.005 (0.33)		
Aid (% GNI) ($t-1$)		0.11 (1.91)*					−0.005 (2.75)***	
Undernutrition (ln)			0.13 (2.33)**					−0.026 (0.54)
Fractionalization				0.174 (1.75)*				
Predom. Muslim					0.004 (4.23)***			
Number of countries	55	55	55	55	55	43	34	42
Observations	522	435	434	532	466	414	302	367
R-squared	0.87	0.86	0.87	0.86	0.89	0.90	0.92	0.92
Kleibergen-Paap F	19.93	35.89	36.38	79.89	79.95	17.69	46.81	56.46
Stock-Wright LM	15.62	13.48	10.2	21.71	16.38	8.53	3.97	6.05
(p-value)	0.00	0.00	0.006	0.00	0.001	0.01	0.14	0.049
Hansen test[a]	0.21	0.15	0.23	0.95	0.23	0.85	0.90	0.42

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.16: Child mortality equation: Additional controls (Sys-GMM regressions)

Dependent variable: ln(Child mortality)					
Regressions	(1)	(2)	(3)	(4)	(5)
Lagged Child mor. (ln)	0.95 (14.81)***	0.87 (15.62)***	0.92 (15.06)***	0.92 (15.00)***	0.94 (13.02)***
Health Spending (% GDP) (ln)	−0.05 (2.97)***	−0.063 (2.12)**	−0.068 (3.24)***	−0.061 (3.24)***	−0.059 (3.13)***
GDP per capita (ln)	−0.025 (2.30)**	−0.021 (1.16)	−0.024 (1.85)*	−0.029 (2.63)***	−0.029 (2.59)***
Under-5 population	−0.01 (0.81)				
Aid (% GNI) ($t-1$)		0.001 (0.71)			
Undernutrition (ln)			−0.003 (0.16)		
Fractionalization				0.032 (0.64)	
Predominantly Muslim					0.0002 (0.55)
Number of countries	43	38	42	43	44
Observations	188	163	173	192	193
Number of instruments	34	34	34	34	34
Hansen test (p-values)	0.45	0.25	0.45	0.26	0.17
Diff-in-Hansen test[a]	0.70	0.52	0.79	0.31	0.34
Autocor. (2^{nd} order)[a]	0.89	0.35	0.45	0.90	0.77
Autocor. (3^{rd} order)[a]	0.14	0.07	0.08	0.11	0.14

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.17: IHDI equation: Spending, welfare, and democracy

Dependent variable: IHDI				
Regression	(1)	(2)	(3)	(4)
Method	OLS	2SLS	FE	Sys-GMM
Lagged IHDI				0.93 (40.44)***
Social Spending (% GDP) (*ln*)	0.009 (9.83)***	0.015 (4.91)***	0.011 (2.46)**	0.004 (1.91)*
Democracy	0.0003 (1.95)*	0.001 (1.86)*	0.0002 (1.22)	0.0002 (0.46)
Social spending×democracy (% GDP)	0.0001 (2.11)**	−0.0001 (0.60)	−0.0003 (2.74)***	−0.0001 (0.49)
Number of countries	51	51	38	42
Observations	504	504	417	175
R-squared	0.92	0.91	0.91	
Number of instruments		2	2	34
Kleibergen-Paap F statistic		19.88	10.64	
Stock-Wright LM statistic		12.97	5.89	
(p-value)		0.00	0.05	
Hansen test (p-value)		0.78	0.49	0.71
Diff-in-Hansen test[a]				0.87
Autocor. (2^{nd} order)[a]				0.50
Autocor. (3^{rd} order)[a]				0.12

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.18: Child mortality equation: Spending, welfare, and democracy

Dependent variable: *ln*(Child mortality rate)				
Regression	(1)	(2)	(3)	(4)
Method	OLS	2SLS	FE	Sys. GMM
Lagged Child mor. (*ln*)				0.93 (13.69)***
Health Spending (% GDP) (*ln*)	−0.148 (9.80)***	−0.20 (5.18)***	−0.104 (2.87)***	−0.058 (2.60)***
Democracy	−0.008 (2.27)**	−0.004 (0.14)	0.009 (3.07)***	0.002 (0.83)
Health spending×democracy (% GDP)	0.004 (1.96)**	0.007 (1.93)*	0.006 (2.42)**	0.001 (0.39)
Number of countries	55	55	44	44
Observations	537	533	425	193
R-squared	0.86	0.86	0.91	
Number of instruments		2	2	34
Kleibergen-Paap F statistic		42.43	18.54	
Stock-Wright LM statistic		12.49	8.63	
(p-value)		0.00	0.01	
Hansen test (p-value)		0.25	0.66	0.27
Diff-in-Hansen test[a]				0.43
Autocorr. (2^{nd} order)[a]				0.90
Autocorr. (3^{rd} order)[a]				0.12

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Table 3.19: IHDI equation: Disaggregating social spending

Dependent variable: IHDI [Sys-GMM regressions]			
Regression	(1)	(2)	(3)
Lagged IHDI	0.99 (41.77)***	0.97 (31.82)***	0.93 (21.03)***
Health spending (% GDP) (*ln*)	0.003 (3.37)***	0.004 (2.11)**	0.004 (2.18)**
Education spending (% GDP) (*ln*)	0.001 (0.40)	−0.002 (1.22)	−0.003 (1.27)
Social protection spending (% GDP) (*ln*)	−0.001 (1.43)	−0.0002 (0.22)	−0.0003 (0.24)
GDP per capita (*ln*)	0.002 (3.91)***	0.002 (1.84)*	0.004 (3.16)***
Number of countries	41	40	40
Observations	184	162	162
Number of instruments	39	49	46
Hansen test (p-value)	0.50	0.88	0.81
Difference-in-Hansen test[a]	0.30	0.48	0.63
Autocorrelation (second-order)[a]	0.40	0.30	0.89
Autocorrelation (third-order)[a]	0.06	0.75	0.11

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

Notes: Column 1 includes only the variables shown here; Columns 2 and 3 include all variables in Table 3.2; Columns 1 & 2 add second and longer lags of health, education, and social protection spending to the instrument set in Table 3.2 (column 2); column 3 uses only the aforementioned internal instruments.

Table 3.20: IHDI and Child mortality equations: Alternative specifications:

Dependent variable	*ln*(IHDI)			Δln(IHDI)	Δln(CM)
Regression	(1)	(2)	(3)	(4)	(5)
Method	2SLS	FE	Sys-GMM	Sys-GMM	Sys-GMM
Initial IHDI/CM (*ln*)				−0.094 (3.77)***	−0.11 (2.25)**
Lagged IHDI			0.91 (38.53)***		
Social spending (% GDP) (*ln*)	0.064 (5.10)***	0.068 (2.34)**	0.02 (2.35)**	0.021 (2.63)***	
Health spending (% GDP) (*ln*)					−0.058 (2.81)***
GDP per capita (*ln*)	0.135 (12.93)***	0.074 (3.90)***	0.011 (2.23)**	0.012 (2.76)***	−0.022 (1.67)*
Number of countries	51	38	42	39	44
Observations	504	417	175	166	193
Number of instruments	2	2	33	33	33
R-squared	0.91	0.85			
Kleibergen-Paap F st.	42.17	9.82			
Stock-Wright LM st.	13.95	5.22			
(p-value)	0.00	0.07			
Hansen test (p-value)	0.63	0.99	0.26	0.30	0.21
Diff-in-Hansen test[a]			0.40	0.56	0.16
Autocor. (2^{nd} order)[a]			0.03	0.02	0.93
Autocor. (3^{rd} order)[a]			0.99	0.98	0.15

[a]Denotes p-value. Significance level: *10%; **5%; ***1%

References

[1] Abed, G. T., and Gupta, S. (2002). *Governance, corruption, and economic performance*. Washington, DC: International Monetary Fund.

[2] Abegaz, B. (1999). Aid and reform in Ethiopia. In *Aid and Reform in Africa* (*pp.* 167-226), Devarajan, S., Dollar, D., and Holmgren, T. (Eds.). Washington, D.C.: World Bank.

[3] Acosta-Ormaechea, S., and Morozumi, A. (2013). Can a government enhance long-run growth by changing the composition of public expenditure? IMF Working paper no. 162. Washington, DC: World Bank.

[4] AfDB. (1997). Ethiopia: Structural adjustment programme project completion report (SAP I). The African Development Bank.

[5] ———. (2002). Cameroon: Project performance evaluation report (SAP II). The African Development Bank.

[6] Ahmad, E., and Stern, N. (1991). *The theory and practice of tax reform in developing countries*. Cambridge: Cambridge University Press.

[7] Ahmed, M. M., Babiker, R., and Bell, M. (2004). An analysis of fiscal policies in the Sudan: A pro-poor perspective. Khartoum: UNDP.

[8] Alderman, H. (1994). Ghana: Adjustment's star pupil? In *Adjusting to policy failure in African economies* (*pp.* 23-52), Sahn, D. E. (Ed.). Ithaca, NY: Cornell University Press.

[9] Alemu, G. (2010). Global financial crisis: Ethiopia Phase 2. Discussion paper no. 16. Overseas Development Institute: London.

[10] Alesina, A., and Rodrik, D. (1994). Distributive politics and economic Growth. *The Quarterly Journal of Economics*, 109(2): 465-490.

[11] Alesina, A., Baqir, R., and Easterly, W. (1999). Public goods and ethnic divisions. *Quarterly Journal of Economics*, 114(4): 1243-1284.

[12] Alesina, A., Devleeschauwer, A., Easterly, W., Kurlat, S., and Wacziarg, R. (2003). Fractionalization. *Journal of Economic Growth*, 8(2): 155-194.

[13] Alesina, A., Ozler, S., Roubini, N., and Swagel, P. (1996). Political instability and economic growth. *Journal of Economic Growth*, 1: 189-211.

[14] Alkire, S., and Foster, J. (2010). Designing the Inequality-adjusted Human Development Index. Human Development research paper no. 28. New York, NY: UNDP.

[15] Alvi, E., and Senbeta, A. (2012). Does foreign aid reduce poverty? *Journal of International Development*, 24(8): 955-976.

[16] Ames, B. (1987). Political survival: Politicians and public policy in Latin America. Berkeley: University of California Press.

[17] Anand, S., and Ravallion, M. (1993). Human development in poor countries: on the role of private incomes and public services. *Journal of Economic Perspectives*,

7: 133-150.

[18] Anand, S., and Sen, A. (1992). Human development index: Methodology and measurement. Background Paper for *Human Development Report* 1993. New York: Oxford University Press.

[19] Anderson, T. W., Kunitomo, N., and Matsushita, Y. (2010). On asymptotic optimality of the LIML estimator with possibly many instruments. *Journal of Econometrics*, 157: 191-204.

[20] Arellano, M., and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58: 277-297.

[21] Arellano, M., and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68: 29-51.

[22] Arndt, C., Jones, S., and Tarp, F. (2010). Aid, growth, and development: Have we come full circle? *Journal of Globalization and Development*, 1: 1-26.

[23] Arulpragasam, J., and Sahn, D. E. (1994). Policy failure and the limits of rapid reform: lessons from Guinea. In *Adjusting to policy failure in African economies* (*pp.* 23-52), Sahn, D. E. (Ed.). Ithaca, NY: Cornell University Press.

[24] Bacha, E. (1990). A three-gap model of foreign transfers and the GDP growth rate in developing countries. *Journal of Development Economics*, 32(2): 279-296.

[25] Bai J., Lumsdaine R. L., and Stock J. H. (1998). Testing for and dating common breaks in multivariate time series. *Review of Economic Studies*, 65, 395-432.

[26] Baldacci, E., Clements, B., Gupta, S., and Cui, Q. (2008). Social spending, human capital, and growth in developing countries. *World Development*, 36(8): 1317-1341.

[27] Baldacci, E., Guin-Siu Maria, T., and De Mello, L. (2003). More on the effectiveness of public spending on health care and education: a covariance structure model. *Journal of International Development*, 15: 709-25.

[28] Barro, R. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, 106(2): 407-443.

[29] Barro, R., and Lee, J-W. (2005). IMF programs: Who is chosen and what are the effects? *Journal of Monetary Economics*, 52: 1245–1269.

[30] Barro, R. J., and Sala-i-Martin, X. (1998). *Endogenous Growth.* MIT Press

[31] Baum, C. F., Schaffer, M. E., and Stillman, S. (2007). Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *Stata Journal*, 7(4): 465-506.

[32] Benhabib, J., and Spiegel, M. M. (1994). The role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics*, 34: 143-173.

[33] Bidani, B., and Ravallion, M. (1997). Decomposing social indicators using distributional data. *Journal of Econometrics*, 77: 125-139.

[34] Blandford, D., Friedman, D., Lynch, S., Mukherjee, N., and Sahn, D. E. (1994). Oil boom and bust: The harsh realities of adjustment in Cameroon. In *Adjusting to policy failure in African economies* (*pp.* 131-163), Sahn, D. E. (Ed.). Ithaca, NY: Cornell University Press.

[35] Blundell, R., and Bond S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87: 115-143.

[36] Boone, P. (1996). Politics and the effectiveness of foreign aid. *European Economic Review*, 40: 289-329.

[37] Borchgrevink, A. (2008). Limits to donor influence: Ethiopia, aid and conditionality. *Forum for Development Studies*, 35(2): 195-220.

[38] Botchwey, K., Collier, P., Gunning, J. W., and Hamada, K. (1998). *External evaluation of the ESAF: Report by a group of independent experts*. Washington, D.C.: International Monetary Fund.

[39] Bun, M., and Windmeijer, F. (2010). The weak instrument problem of the System GMM estimator in dynamic panel data models. *Econometrics Journal*, 13(1): 95-126.

[40] Burnside, C., and Dollar, D. (1998). Aid, growth, the incentive regime, and poverty reduction. In *The World Bank: Structure and policies*, Christopher, L., and Vines, D. (Eds.). Oxford: Oxford University Press.

[41] Burnside, C., and Dollar, D. (2000). Aid, policies, and growth. *American Economic Review*, 90: 847-868.

[42] Carrin, G., and Politi, C. 1995. Exploring the health impact of economic growth, poverty reduction, and public health expenditure. Tijdschrift voor Economie en Management 40: 227-246.

[43] Cassen R (and associates). (1994). *Does aid work*? Clarendon Press: Oxford.

[44] Chang, C., Fernandez-Arias, E., and Serven, L. (1998). Measuring aid flows: A new approach. Washington, D.C.: World Bank.

[45] Chenery, H. B., and Strout, A. (1966). Foreign assistance and economic development. *American Economic Review*, 56: 679-733.

[46] Clemens, M. A., Radelet, S., Bhavnani, R. R., and Bazzi, S. (2012). Counting chickens when they hatch: Timing and the effects of aid on growth. *Economic Journal*, 122: 590-617.

[47] Collier, P., and Dehn, J. (2001). Aid, shocks, and growth. Working paper no. 2688. Washington, D.C.: World Bank.

[48] Corbo, V., Goldstein, M., and Kahn, M. (1987). *Growth-oriented adjustment programs*. Washington, D.C.: International Monetary Fund and The World Bank.

[49] Cornia, G. A., Jolly, R., and Stewart, F. (1988). *Adjustment with a human face*, 2 vols. Oxford: Clarendon Press.

[50] Cromwell, E. (1992). Malawi. In *Structural adjustment and the African farmer*,

Duncan, A., and Howell, J. (Eds.). London: James Currey.

[51] Cutler, D., Deaton, A., and Lleras-Muney, A. (2006). The determinants of mortality. *Journal of Economic Perspectives*, 20(3): 97-120.

[52] Dalgaard, C.-J., and Erickson., L. (2009). Reasonable expectations and the first Millennium Development Goal: How much can aid achieve? *World Development*, 37 (7): 1170-1181.

[53] Dalgaard, C.-J., and Hansen, H. (2001). On aid, growth and good policies. *Journal of Development Studies*, 37: 17-41.

[54] Dalgaard, C.-J., Hansen, H., and Tarp, F. (2004). On the empirics of foreign aid and growth. *The Economic Journal*, 114(496): 191-216.

[55] Dennis, J. G. (2006). *CATS in RATS: Cointegration analysis of time Series*, Version 2. Illinois: Estima, Evanston.

[56] Dennis, J. G., Hansen, H., Johansen, S., and Juselius, K. (2006). *CATS in RATS: Cointegration Analysis of Time Series*, Version 2. Illinois: Estima.

[57] Dercon, S. (2000). Growth and poverty in Ethiopia in the 1990s: an economic perspective. Centre for the Study of African Economics, Oxford University.

[58] Devarajan, S., and De Melo, J. (1990). Membership in the CFA zone: Odyssean journey or trojan horse? Working paper 482, Washington, D.C.: The World Bank.

[59] Devarajan, S., and Rodrik, D. (1991). Do benefits of fixed exchange rates outweigh their costs? The Franc zone in Africa. Working paper 3727, Cambridge, MA: National Bureau of Economic Research.

[60] Doornik, J. A., and Hendry, D. F. (2001). *Empirical econometric modelling using PcGive*, Volumes I, II and III. London: Timberlake Consultants Press.

[61] Doornik, J. A., Hendry, D. F., and Nielsen, B. (1998). Inference in cointegrating models: UK M1 revisited. *Journal of Economic Surveys*, 12(5): 533-572.

[62] Doornik, J. A., Hendry, D. F., and Pretis, F. (2013). Step-indicator saturation. Discussion paper 658, Department of Economics, University of Oxford.

[63] Dollar, D., and Kraay, A. (2001). Growth is good for the poor. Policy research Working Paper no. 2587. Washington, DC: World Bank.

[64] Dorosh, P. (1994). Economic fallout from a uranium boom: Structural adjustment in Niger. In *Adjusting to Policy Failure in African Economies* (*pp.* 164-195), Sahn, D. E. (Ed.). Ithaca, NY: Cornell University Press.

[65] Dorosh, P., and Bernier, R. (1994). Staggered reforms and limited success: Structural adjustment in Madagascar. In *Adjusting to Policy Failure in African Economies* (*pp.* 302-365), Sahn, D. E. (Ed.). Ithaca, NY: Cornell University Press.

[66] Dreher, A. (2006). IMF and economic growth: The effects of programs, loans, and compliance with conditionality. *World Development*, 34(5): 769-788.

[67] Dreher, A., Nunnenkamp, P., and Thiele, R. (2008). Does aid for education educate children? Evidence from panel data. *World Bank Economic Review*, 22(2): 291-314.

[68] Duruflé, G. (1989). Structural disequilibrium and adjustment politics in the Ivory Coast. In *Structural adjustment in Africa*, Campbell, B., and Loxley, J. (Eds.). London: Macmillan.

[69] Easterly, W. (1999). The ghost of the financing gap: testing the growth model used in the international financial institutions. *Journal of Development Economics*, 60(2): 423-438.

[70] ———. (2003). Can foreign aid buy growth? *Journal of Economic Perspectives*, 17: 23-48.

[71] ———. (2005). What did structural adjustment adjust? *Journal of Development Economics*, 76(1): 1-22.

[72] Easterly, W., Levine, R., and Roodman, D. (2004). New data, new doubts: A comment on Burnside and Dollar's 'aid, policies, and growth'. *American Economic Review*, 94: 774-780.

[73] Easterly, W., and Rebelo, S. (1993). Fiscal policy and economic growth: An empirical investigation. *Journal of Monetary Economics*, 32: 417-458.

[74] Elbadawi, I., Ghura, D., and Uwujaren, G. (1992). World Bank adjustment lending and economic performance in Sub-Saharan Africa in the 1980s. Working paper 1000, Washington, D.C.: The World Bank.

[75] Engle, R. and Granger, C. (1987). Co-integration and error-correction: Representation, estimation and testing. *Econometrica*, 55: 251-276.

[76] Faini, R., De Melo, J., Senhadji, A., and Stanton, J. (1991). Growth-oriented adjustment programs: A statistical analysis. *World Development,* 19(8): 957-967.

[77] Faruqee, R., and Hussain, I. (1994). Adjustment in seven African countries. In *Adjustment in Africa: lessons from country case studies* (*pp.* 1-10), Hussain, I., and Faruqee, R. (Eds.). Washington, D.C.: The World Bank.

[78] Feyzioglu, T., Swaroop, V., and Zhu, M. (1998). A panel data analysis of the fungibility of foreign aid. *The World Bank Economic Review*, 12: 29-58.

[79] Filmer, D., Hammer, J. S., and Pritchett, L. (2000). Weak links in the chain: A diagnosis of health policy in poor countries. *World Bank Research Observer*, 15: 199-224.

[80] Filmer, D., and Pritchett, L. (1999). The impact of public spending on health: does money matter? *Social Science and Medicine*, 49: 1309-1323.

[81] Fischer, S. (1989). Vice-President and Chief Economist of the World Bank, in a speech to the G-24 Deputies, April 3.

[82] Flug, K., Spilimbergo, A., and Wachtenheim, E. (1998). Investment in education: Do economic volatility and credit constraints Matter? *Journal of Development Economics*, 55: 465-481.

[83] Franses, P. H., and Haldrup, N. (1994). The effect of additive outliers on tests for unit roots and cointegration. *Journal of Business and Economic Statistics*, 12(4):

471-478.

[84] Furtado, X., and Smith, W. J. (2009). Ethiopia: Retaining sovereignty in aid relations. In *The politics of aid: African strategies for dealing with donors*, Whitfield, L. (Ed.). Oxford: Oxford University Press.

[85] Galor, O., and Weil, D. N. (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American Economic Review*, 90(4): 806-828.

[86] Gebregziabher, F. H. (2014). The long-run macroeconomic effects of aid and disaggregated aid in Ethiopia. *Journal of International Development*, 26(4), 520-540.

[87] Geda, A. (2001). Macroeconomic performance in post-Derg Ethiopia. *Northeast African Studies*, 8(1): 159-204.

[88] ———. (2005). The political economy of growth in Ethiopia. In *The political economy of economic growth in Africa*, Ndulu, B. J., O'Connell, S., Azam, J. P., Bates, R. H., Fosu, A. K., Gunning, J. W., Njinkeu, D. (Eds.). Cambridge: Cambridge University Press.

[89] Geda, A., and Degefe, B. (2005). Explaining African growth performance: The case of Ethiopia. Working paper no. 3, African Economic Research Consortium.

[90] Gomanee, K., Girma, S., and Morrissey, O. (2005a). Aid, public spending and human welfare: Evidence from quantile regressions. *Journal of International Development*, 17(3): 299-309.

[91] Gomanee, K., Girma, S., and Morrissey, O. (2005b). Aid and growth in Sub-Saharan Africa: Accounting for transmission mechanisms. *Journal of International Development*, 17(8): 1055-1075.

[92] Gomanee, K., Morrissey, O., Mosley, P., and Verschoor, A. (2005c). Aid, government expenditure and aggregate welfare. *World Development*, 33(3): 355-370.

[93] Gonzalo, J. (1994). Five alternative methods of estimating long-run equilibrium relationships. *Journal of Econometrics*, 60: 203-233.

[94] Greenwood, J., and Seshadri, A. (2002). The U.S. demographic transition. *American Economic Review*, 92(2): 153-159.

[95] Gupta, S., Davoodi, H., and Alonso-Terme, R. (2002). Does corruption affect income inequality and poverty? *Economics of Governance*, 3(1): 23-45.

[96] Gupta, S., Verhoeven, M., and Tiongson, E. (2002). The effectiveness of government spending on education and health care in developing and transition economies. *European Journal of Political Economy*, 18(4): 717-737.

[97] Gupta, S., Verhoeven, M., and Tiongson, E. (2003). Public spending on health care and the poor. *Health Economics*, 12(8): 685-696.

[98] Hadjimichael M. T., Nowak, M., Sharer, R., and Tahari, A. (1996). Adjustment for growth: The African experience. Occasional paper 143, Washington, D.C.: International Monetary Fund.

[99] Hahn, J., Hausman, J., and Kuersteiner, G. (2004). Estimation with weak instruments: Accuracy of higher-order bias and MSE approximations. *Econometrics Journal*, 7(1): 272-306.

[100] Hansen, H., and Tarp, F. (2001). Aid and growth regressions. *Journal of Development Economics*, 64(2): 547-570.

[101] Hanushek, E. (1995). Interpreting recent research on schooling in developing countries. *World Bank Research Observer*, 10(2): 227-246.

[102] Harbison, R., and Hanushek, E. (1992). Educational performance of the poor: Lessons from rural Northeast Brazil. New York: Oxford University Press.

[103] Harrigan, J. (1991). Malawi. In *Aid and power: The World Bank and policy-based lending*, Vol. 2, Mosley, P., Harrigan, J., and Toye, J. (Eds.). London: Routledge.

[104] ———. (1997). Modeling the impact of World Bank policy-based lending: The case of Malawi's agricultural sector. *The Journal of Development Studies*, 33(6): 848-873.

[105] ———. (2003). U-Turns and full circles: Two decades of agricultural reform in Malawi 1981-2000. *World Development,* 31(5): 847-862.

[106] Hausmann, R., Pritchett, L., and Rodrik, D. (2005). Growth accelerations. *Journal of Economic Growth*, 10(4): 303-329.

[107] Hayakawa, K. (2006). Efficient GMM estimation of dynamic panel data models where large heterogeneity may be present. Hi-Stat Discussion Paper no. 130, Hitotsubashi University.

[108] Hendry, D. F. and Juselius, K. (2001). Explaining cointegration analysis: Part II. *Energy Journal*, 22(1): 75-120.

[109] Heston, A., Summers, R., and Aten, B. (2011). Penn World Table Version 7.0. Center for international comparisons of production, income, and prices, University of Pennsylvania, Philadelphia.

[110] ———. (2012). Penn World Table Version 7.0, Center for international comparisons of production, income and prices at the University of Pennsylvania, Philadelphia.

[111] Hicks, D. A. (1997). The Inequality-adjusted Human Development Index: A constructive proposal. *World Development*, 25(8): 1283-1298.

[112] Hojman, D. E. (1996). Economic and other determinants of infant and child mortality in small developing countries: the case of Central America and the Caribbean. *Applied Economics*, 28: 281-290.

[113] Hoover, K., Juselius, K., and Johansen, S. (2008). Allowing the data to speak freely: The macroeconometrics of the cointegrated vector autoregression. *American Economic Review, Papers and Proceedings*, 98: 251-255.

[114] Huang, Y., and Quibria, M. G. (2013). The global partnership for inclusive growth. WIDER Working Paper no. 059. Helsinki: UNU-WIDER.

[115] Hungnes, H. (2005). Identifying the deterministic components in GRaM for

Ox Professional: User manual and documentation, mimeo. Available at http://www.hungnes.net/GRaM.

[116] ———. (2010). Identifying structural breaks in cointegrated vector autoregressive models. *Oxford Bulletin of Economics and Statistics*, 72 (4): 551-565.

[117] Husain, I., and Faruqee, R. (1994). *Adjustment in Africa: lessons from country case Studies*. Washington, D.C.: The World Bank.

[118] Imam, P., and Salinas G. (2008). Explaining episodes of growth accelerations, decelerations, and collapses in Western Africa. Working Paper 287. Washington, D.C.: IMF.

[119] IMF. (1996). IMF approves new three-year ESAF loan for Ethiopia. Press release number 96/51, Washington, D.C.: International Monetary Fund.

[120] ———. (1997). IMF approves three-year ESAF loan for Cameroon. Press release number 97/38, Washington, D.C.: International Monetary Fund.

[121] ———. (1998). Côte d'Ivoire: ESAF policy framework paper, 1998-2000. Washington, D.C.: International Monetary Fund.

[122] ———. (1998). Ghana: ESAF policy framework paper, 1998-2000. Washington, D.C.: International Monetary Fund.

[123] ———. (1998). Rwanda: ESAF policy framework paper, 1998/99-2000/01. Washington, D.C.: International Monetary Fund.

[124] ———. (1999). Burkina Faso: ESAF policy framework paper, 2000-2002. Washington, D.C.: International Monetary Fund.

[125] ———. (1999). Mozambique: ESAF policy framework paper, 1999-2002. Washington, D.C.: International Monetary Fund.

[126] ———. (1999). Senegal: ESAF policy framework paper, 1999-2001. Washington, D.C.: International Monetary Fund.

[127] ———. (1999). The IMF's ESAF: Is it working? Washington, D.C.: International Monetary Fund.

[128] ———. (2001). *A Manual on Government Finance Statistics*. Washington, DC: International Monetary Fund.

[129] Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12: 231-254.

[130] ———. (1991). Estimating and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica*, 59: 1551-1580.

[131] ———. (1996). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford: Oxford University Press.

[132] ———. (2000). Modelling of cointegration in the vector autoregressive model. *Economic Modelling*, 17: 359-373.

[133] ———. (2002). A small sample correction of the test for cointegrating rank in the vector autoregressive model. *Econometrica*, 70: 1929-1961.

[134] Johansen, S., and Juselius, K. 1994. Identification of the long-run and short-run structure: an application to the ISLM model. *Journal of Econometrics*, 63: 7-36.

[135] ———. (2014). An asymptotic invariance property of the common trends under linear transformations of the data. *Journal of Econometrics*, 178: 310-315.

[136] Johansen, S., Mosconi, R., and Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. *The Econometrics Journal*, 3(2): 216-249.

[137] Jones, B. F., and Olken, B. A. (2008). The anatomy of start-stop growth. *Review of Economics and Statistics*, 90(3): 582-587.

[138] Juselius, K. (2006). *The cointegrated VAR model: Methodology and applications.* Oxford: Oxford University Press.

[139] Juselius, K. and MacDonald, R. (2004). International parity relationships between the USA and Japan. *Japan and the World Economy*, 16: 17-34.

[140] Juselius, K., Møller N., and Tarp, F. (2014). The long-run impact of foreign aid in 36 African countries: Insights from multivariate time series analysis. *Oxford Bulletin of Economics and Statistics*, 76(2): 153-184.

[141] Juselius, K., Reshid, A., and Tarp, F. (2013). The real exchange rate, foreign aid and macroeconomic transmission mechanisms in Tanzania and Ghana. Working paper no. 90, UNU-WIDER.

[142] Kanbur, R. (1990). Poverty and the social dimensions of structural adjustment in Côte d'Ivoire. Working paper 2, Washington, D.C.: The World Bank.

[143] Kosack, S. (2003). Effective aid: How democracy allows development aid to improve the quality of life. *World Development*, 31(1): 1-22.

[144] Kim, K., and Moody, P. M. (1992). More resources better health? A cross-national perspective. *Social Science and Medicine*, 34: 837-842.

[145] Kyle, S. (1994). Structural adjustment in a country at war: The case of Mozambique. In *Adjusting to policy failure in African economies* (*pp.* 234-259), Sahn, D. E. (Ed.). Ithaca, NY: Cornell University Press.

[146] La Porta, R., Lopez-de-Silanes, F., Shleifer, A., and Vishny, R. (1999). The quality of government. *Journal of Law, Economics, and Organization*, 15(1): 222-279.

[147] Lele, U. (1990). Structural adjustment, agricultural development, and the poor: Some lessons from the Malawian experience. *World Development*, 18(9): 1207-1219.

[148] Lipton, M. (1987). Limits of price policy for agriculture: Which way for the World Bank? *Development Policy Review,* 5: 197-215.

[149] Lucas, R. E. Jr. (1988). On the mechanic of economic development. *Journal of Monetary Economics*, 22(1): 3-42.

[150] Marshall, M. G., Gurr, T. D., and Jaggers, K. (2013). Polity IV project: Dataset users' manual. College Park, MD: University of Maryland.

[151] Mauro, P. (1998). Corruption and the composition of government expenditure.

Journal of Public Economics, 69: 263-279.

[152] Mavrotas, G. (2005). Aid heterogeneity: looking at aid effectiveness from a different angle. *Journal of International Development*, 17(8): 1019-1036.

[153] Mingat, A., and Tan, J.-P. (1992). Education in Asia: A comparative study of cost and financing. Washington, DC: World Bank.

[154] Mingat, A., and Tan, J.-P. (1998). The mechanics of progress in education: Evidence from cross-country data. Working paper no. 2015. Washington, DC: World Bank.

[155] Mishra, P., and Newhouse, D. (2009). Does health aid matter? *Journal of Health Economics*, 28: 855-872.

[156] Mosley, P. (1991). Kenya. In *Aid and power: The World Bank and policy-based lending*, Vol. 2, Mosley, P., Harrigan, J., and Toye, J. (Eds.). London and New York: Routledge.

[157] Mosley, P., Hudson, J., and Verschoor, A. (2004). Aid, poverty reduction and the 'new conditionality'. *The Economic Journal*, 114: 217-243.

[158] Mosley, P., and Suleiman, A. (2007). Aid, agriculture, and poverty in developing countries. *Review of Development Economics*, 11: 139-158.

[159] Mosley, P., and Weeks, J. (1993). Has recovery begun? Africa's adjustment in the 1980s revisited. *World Development,* 21(10): 1583-1606.

[160] Musgrove, P. (1996). Public and private roles in health: theory and financing patterns. Discussion paper no. 339. Washington, DC: World Bank.

[161] Mwega, F. M., and Ndung'u, N. S. (2008). Explaining African economic growth performance: The case of Kenya. In *The political economy of economic growth in Africa, 1960-2000*, (Vol 2: *pp.* 325-368), Ndulu, B., O'Connell, S., Azam, J-P., Bates, R. H., Fosu, A. K., Gunning, J. W., Njinkeu, D. (Eds.). Cambridge: Cambridge University Press.

[162] Ndulu, B., O'Connell, S., Azam, J-P., Bates, R. H., Fosu, A. K., Gunning, J. W., Njinkeu, D. (2008). The political economy of economic growth in Africa 1960-2000, Vol. 2. Cambridge: Cambridge University Press.

[163] Nielsen, H. B. (2004). Cointegration analysis in the presence of outliers. *The Econometrics Journal*, 7(1): 249-271.

[164] ———. (2008). Influential observations in cointegrated VAR models: Danish money demand 1973-2003. *The Econometrics Journal*, 11: 1-19.

[165] OECD (2012). Aid statistics. Paris: OECD.

[166] Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica,* 57, 1361-1401. Erratum (1993) *Econometrica*, 61: 248-249.

[167] Political Risk Services. (2007). International country risk guide. Available from <www.prsgroup.com>.

[168] Przeworski, A., and Vreeland, J. R. (2000). The effect of IMF programs on economic

growth. *Journal of Development Economics*, 62(2): 385-421.

[169] Pritchett, L., and Summers, L. H. (1996). Wealthier is healthier. *Journal of Human Resources*, 31(4): 841-868.

[170] Psacharopoulos, G. (1994). Returns to investment in education: a global update. *World Development*, 22(9): 1325-1343.

[171] Psacharopoulos, G., and Patrinos, H. A. (2004). Returns to investment in education: a further update. *Education Economics*, 12(2): 111-134.

[172] Rahbek, A., Hansen, E., and Dennis, J. G. (2002). ARCH innovations and their impact on cointegration rank testing. Working Paper no. 22, Centre for Analytical Finance, University of Copenhagen.

[173] Rajan, R. G., and Subramanian, A. (2008). Aid and growth: What does the cross-country evidence really show? *The Review of Economics and Statistics*, 90: 643-665.

[174] Rajkumar, A., and Swaroop, V. (2008). Public spending and outcomes: Does governance matter? *Journal of Development Economics*, 86: 96-111.

[175] Ravallion M., and Chen S. (1997). What can new survey data tell us about recent changes in distribution and poverty? *World Bank Economic Review*, 11(2): 357-382.

[176] Rebelo, S. T. (1991). Long-run policy analysis and long-run growth. *Journal of Political Economy*, 99(3): 500-521.

[177] Reddy, S., and Pogge, T. (2010). How not to count the poor. In Anand, S., and Stiglitz, J. (Eds.), *Measuring global poverty.* Oxford: Oxford University Press.

[178] Riddell, R. (2007). *Does foreign aid really work?* Oxford: Oxford University Press.

[179] Rodrik, D. (1990). How should structural adjustment progams be designed? *World Development,* 18(7): 933-941.

[180] ———. (1998). Trade performance and economic performance in Sub-Saharan Africa. Working paper 6562, Cambridge, MA: NBER.

[181] Rodrik, D. (2005). Why we learn nothing from regressing economic growth on policies. Mimeograph, Kennedy School of Government.

[182] Rodrik, D., Subramanian, A., and Trebbi, F. (2004). Institutions rule: The primacy of institutions over geography and integration in economic development. *Journal of Economic Growth*, 9(2): 131-165.

[183] Roodman, D. (2005). Xtabond2: Stata module to extend Xtabond dynamic panel data estimator. Available at http://econpapers.repec.org/software/bocbocode/s435901.htm.

[184] Roodman, D. (2007). The anarchy of numbers: Aid, development, and cross-country empirics. *The World Bank Economic Review*, 21 (2): 255-277.

[185] Roodman, D. M. (2009a). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, 9 (1): 86-136.

[186] Roodman, D. M. (2009b). A note on the theme of too many instruments. *Oxford*

Bulletin of Economics and Statistics, 71: 135-158.

[187] Romer, P. M. (1994). The origins of endogenous growth. *Journal of Economic Perspectives*, 8(1): 3-22.

[188] Rosenzweig, M. R., and Schultz, T. P. (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *Journal of Political Economy*, 91(5): 723-746.

[189] Ruggeri-Laderchi, C., Saith, R., and Stewart, F. (2003). Does it matter that we don't agree on the definition of poverty? A comparison of four approaches. QEH Working paper no. 107. Oxford: University of Oxford.

[190] Sahn, D.E., and Arulpragasam, J. (1991). The stagnation of smallholder agriculture in Malawi: A decade of structural adjustment. *Food Policy*, 16: 219-234.

[191] Sahn, D. E., and Arulpragasam, J. (1994). Adjustment without structural change: The case of Malawi. In *Adjusting to policy failure in African economies* (*pp.* 196-233), Sahn, D. E. (Ed.). Ithaca, NY: Cornell University Press.

[192] Sahn, D. E., Dorosh, P., and Younger, S. D. (1998). *Structural adjustment reconsidered: Economic policy and poverty in Africa.* Cambridge: Cambridge University Press.

[193] Sarris, A. H., and Van den Brink, R. (1994). From forced modernization to perestroika: Crisis and adjustment in Tanzania. In *Adjusting to policy failure in African economies* (*pp.* 260-301), Sahn, D. E. (Ed.). Ithaca, NY: Cornell University Press.

[194] Seiferling, M. (2013). Recent improvements to the *Government Finance Statistics* yearbook database in response to analytical needs. Working paper no. 15. Washington, DC: International Monetary Fund.

[195] Sen, A. N. 1999. *Development as freedom.* New York: Alfred A. Knopf.

[196] Schaffer, M. E. (2010). Xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. Available at http://ideas.repec.org/c/boc/bocode/s456501.html.

[197] Schatz, S. P. (1994). Structural adjustment in Africa: A failing grade so far. *The Journal of African Studies*, 32(4): 679-692.

[198] Schultz, T. P. (1993). Investments in the schooling and health of women and men: Quantities and returns'. *The Journal of Human Resources*, 28 (4): 694-734.

[199] Schultz, T. P. (1999). Health and schooling investments in Africa. *Journal of Economic Perspectives*, 13(3): 67-88.

[200] Spanos, A. (2009). The pre-eminence of theory versus the European CVAR perspective in macroeconometric modeling. *Economics – The Open-Access Open-Assessment E-Journal*, 3: 21-39.

[201] Squire, L. (1993). Fighting Poverty. *American Economic Review*, 83(2): 377-382.

[202] Staiger, D., and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65: 557-586.

[203] Stiglitz J. (2003). *Globalization and its discontent.* New York: Norton paperbacks.

[204] Summers, L. H., and Pritchett, L. H. 1993. The structural-adjustment debate. *The American Economic Review*, 83(2): 383-389.

[205] Svensson, J. (1999). Aid, growth and democracy. *Economics and Politics*, 11(3): 275-297.

[206] Swamy, G. (1994). Kenya: Patchy, intermittent commitment. In *Adjustment in Africa: lessons from country case Studies* (*pp.* 193-237), Husain, I., and Faruqee, R. (Eds.). Washington, D.C.: The World Bank.

[207] Tanzi, V. (1992). Structural factors and tax revenue in developing countries: A decade of evidence. In *Open economies, structural adjustment and agriculture* (*pp.* 267-285) Goldin, I., and Winters, L. A. (Eds.). Paris: CEPR and OECD.

[208] Temple, J. (2010). Aid and conditionality. In *Handbook of Development Economics*, Rodrik, D., and Rosenzweig, M. (Eds.). Amsterdam: Elsevier.

[209] UNDP. (2011). *Sustainability and Equity: A Better Future for All.* New York, NY: Palgrave Macmillan.

[210] Von Hagen, J. (2005). Political economy of fiscal institutions. GESY Discussion paper no. 149. Germany: Mannheim University.

[211] Wacziarg, R., and Welch, K. H. (2008). Trade liberalization and growth: New evidence. *World Bank Economic Review*, 22(2), 187-231.

[212] Wade, R. H. (2001). Capital and revenge: The IMF and Ethiopia. *Challenge*, 44(5): 67-75.

[213] Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics*, 126(1): 25-51.

[214] Wolf, A. (2004). Education and economic performance: Simplistic theories and their policy consequences. *Oxford Review of Economic Policy*, 20(2): 315-333.

[215] World Bank (1988a). Adjustment lending: An evaluation of ten years of experience. Policy and research series no. 1, Washington, D.C.: The World Bank.

[216] ———. (1990). Sudan: reversing the economic decline. Report No. M14-SU, Washington, D.C.: The World Bank.

[217] ———. (1994). *Adjustment in Africa: Reforms, results and the road ahead.* Oxford: Oxford University Press.

[218] ——. (1995). Cameroon: Project completion report. Washington, D.C.: World Bank.

[219] ——. (1995). Niger: Economic recovery credit. Washington, D.C.: World Bank.

[220] ———. (1995). *Malawi: Agriculture sector memorandum: strategy options in the 1990s*, Vol. 1. Washington, D.C.: The World Bank.

[221] ———. (1997). Accelerating Malawi's growth: long-term prospects and transitional problems. Washington, D.C.: World Bank.

[222] ——. (2001). Niger: Country assistance progress report. Washington, D.C.: World

Bank.
[223] ——. (2002). A case for aid: Building a consensus for development assistance. Washington, D.C.: World Bank.
[224] ———. (2004). *Making services work for poor people.* World Development Report 2004. Washington, DC: World Bank.
[225] ——. (2004). Zambia: Policies for growth and diversification. Washington, D.C.: World Bank.
[226] ——. (2005). *Economic Growth in the 1990s: Learning from a Decade of Reform.* Washington, D.C.: World Bank.
[227] ——. (2008a). *Sustaining and Sharing Economic Growth in Tanzania.* Washington DC: The World Bank.
[228] ——. (2008b). *The Growth Report: Strategies for Sustained Growth and Inclusive Development.* Washington, D.C.: World Bank.
[229] ———. (2011). World Development Indicators. Washington, DC: World Bank.
[230] ———. (2012). World Development Indicators. Washington, DC: World Bank.