ECONSTOR
*Make Your Publications Visible.*

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Herzer, Dierk; Morrissey, Oliver

**Conference Paper**

# The Long-Run Effect of Foreign Aid on Domestic Output

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Effectiveness of Foreign Aid, No. B19-V2

**Provided in Cooperation with:**
Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Herzer, Dierk; Morrissey, Oliver (2010) : The Long-Run Effect of Foreign Aid on Domestic Output, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2010: Ökonomie der Familie - Session: Effectiveness of Foreign Aid, No. B19-V2, Verein für Socialpolitik, Frankfurt a. M.

This Version is available at:
https://hdl.handle.net/10419/37470

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

# The Long-Run Effect of Foreign Aid on Domestic Output

Dierk Herzer [a,*], Oliver Morrissey [b]

[a] *Johann Wolfgang Goethe-University, Grüneburgplatz 1, 60323 Frankfurt am Main, Germany*

[b] *University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom*

**Abstract**

This paper makes two main contributions. First, we examine the long-run effect of foreign aid on domestic output for 59 developing countries using heterogeneous panel cointegration techniques to control for omitted variable and endogeneity bias and to detect possible cross-country differences in the output effect of aid. The main result is that aid has, on average, a negative long-run effect on output, but there are large differences across countries (in about a third of cases the effect is positive). Second, we use a general-to-specific variable selection approach to systematically search for country-specific factors explaining the cross-country differences in the estimated long-run effect of aid. In contrast to previous studies, we find that aid effectiveness does not depend primarily on factors such as the quality of economic policy, the share of a country's area that is in the tropics, the level of democracy or political stability. The results suggest that the cross-country heterogeneity in the output effect of aid can be explained mainly by cross-country differences in law and order, religious tensions and government size.

*JEL-Classification:* F35; O11; C23; C52

*Keywords:* Foreign aid; Domestic output; Heterogeneous panel cointegration techniques; General-to-specific approach

* Corresponding author. Tel.: + 49 69 798 34783; fax: + 49 69 789 35015. E-mail address: DierkHerzer@gmx.de.

## 1. Introduction

Aid financing, and the policy reform conditions typically associated with aid, is the cornerstone of international development strategies. Developing countries, especially the poorest, have insufficient domestic resources to finance their investment and development needs, while rich countries have a desire to assist countries that are less well off; aid serves these purposes as rich donors provide concessional finance to poor countries. The core premise is that such aid is effective in contributing to economic growth and development. This premise appears to be accepted at a global political level. The 2002 Financing for Development Conference in Monterrey, Mexico, and the 2005 G-8 Summit in Gleneagles, Scotland, endorsed commitments by rich countries to increase the amount of foreign aid significantly, especially to Africa. In the same vein, the Report of the Commission for Africa (2005) advocated a doubling of aid to Africa. These commitments are often reiterated, such as at the L'Aquila Summit in 2009. There appears to be a political belief in the effectiveness of aid, although skeptics could counter that it is merely rhetoric as the commitments have not been fully met. This suggests a prior question to investigate: is there evidence that aid is related to output (thus growth) over time on average? This is one question addressed here.

A large literature on aid effectiveness has attempted to provide empirical evidence without reaching any consensus (sketched below). However, this literature based on cross-country or panel growth regressions suffers from many limitations: failure to account for country heterogeneity in aid effects over time; using growth as the dependent variable but levels as independent variables; and the endogeneity problem with weak instruments. Furthermore, it is unclear which factors among the potentially important determinants of growth (explanatory or control variables) are really important for aid effectiveness. Different studies suggest different variables but rarely have there been rigorous attempts to test these against each other. Consequently, there is uncertainty about which variables act, in fact, as empirically important conditioning variables for the effect of aid in promoting growth. This paper attempts to overcome some of these problems by making three contributions:

(1) Heterogeneous panel cointegration techniques, that are robust to omitted variables and endogenous regressors, are employed to examine the long-run effect of foreign aid on domestic output for 59 developing countries over the period 1971 to 2003 — both for the sample as a whole and for each country individually (demonstrating heterogeneity). Accordingly, in contrast to previous studies, the level of output rather than the growth rate of output is used as the dependent variable.

(2) Having estimated the long-run response of output to aid for each country, a variable selection approach (based on a general-to-specific methodology to identify from a large number of potentially relevant variables those that are the important factors) is employed to explain the cross-country differences in the long-run effect of aid on output.

(3) A methodological contribution of this paper is to use a two-step estimation procedure that combines panel and cross-section methods. The first step involves estimating the long-run effect of aid on output for each country using heterogeneous panel estimators. The second step involves estimating the determinants of aid effectiveness using cross-section regressions with the estimated output effect from the first stage as the dependent variable. The aim is to identify which country-specific factors are empirically important determinants of long-run aid effectiveness.

The paper contributes to understanding of the determinants of aid effectiveness whilst avoiding many of the specification and estimation debates that plague the existing literature. A considerable amount of empirical research since the 1970s has addressed aspects of aid effectiveness (see Hansen and Tarp, 2000) without producing a consensus; rather, views appear quite polarized. Studies such as Burnside and Dollar (2000, 2004), Collier and Dollar (2004) find that aid itself has no effect on growth, although when interacted with policy there is a conditional effect; aid has a positive impact on growth in countries with a good policy environment. Others find that this result is not robust so aid is insignificant irrespective of policy (e.g., Easterly et al., 2004) or find evidence for a positive, albeit small in magnitude, effect of aid on growth even allowing for conditioning on policy (e.g., Dalgaard and Hansen, 2001; Daalgard et al., 2004; Hansen and Tarp, 2001). There are now numerous studies of aid effectiveness with conflicting results; the majority of studies find that aid has a small, conditional positive effect on growth, but there is considerable disagreement on which conditional factors are most important (Morrissey, 2006).

As these studies are based on cross-country econometric growth regressions, it should not be surprising that they yield mixed results. Growth is a complex process that is difficult to explain and aid is only one of many factors that may influence growth; many other factors are potentially important (hence candidates as conditioning variables) for some, but not necessarily all, countries. Furthermore, aid can take many different forms, some of which (e.g. financing for investment) are more likely to have a medium term impact on growth than others, such as technical assistance and humanitarian relief. More importantly, although the empirical aid literature has provided valuable insights into whether and how aid may promote growth, it suffers from the limitations inherent in cross-country growth regressions.

Cross-country (panel) regression models, by definition, are not able to capture the heterogeneity in the relationship between aid and growth across countries, and the estimates may be seriously biased in the presence of such heterogeneity. Several country-specific factors may induce apparent differences in the effect of aid on growth, but these factors cannot be fully controlled for in cross-country regressions (especially if effectively unobservable); the classical omitted variables problem. Panel estimation can account for unobserved country-specific effects, but the homogeneous panel estimators used in the aid literature produce inconsistent and potentially misleading estimates of the average values of the parameters in dynamic models when the slope coefficients differ across cross-section units (Pesaran and Smith, 1995).

Another (closely related) problem with the cross-country approach used in the majority of studies is the potential endogeneity of aid. Aid may go to countries that have just experienced natural disasters or severe economic shocks, which could explain a negative correlation between aid and growth. Alternatively, to the extent that donors reward economically successful countries with increased aid one might see a positive correlation between aid and growth, which again would not reflect a causal effect. The recent literature attempts to control for this endogeneity problem through instrumental variable methods. However, it is well known that instrumental variables regressions may lead to spurious results when the instruments are weak or invalid and it is also well known that it is difficult (and sometimes even impossible) to find variables that qualify as valid instruments (Temple, 1999). This does not mean one must reject the empirical results (it does help to explain the mixed evidence), but implies that results must be interpreted with caution.

A further methodological problem with both cross-country and panel studies is the use of the growth rate of output as the dependent variable while the level of the aid/GDP ratio of the recipient country is used as the explanatory variable. Growth rates show, in general, very little persistence over time, whereas the aid/GDP ratio has exhibited persistent movements with positive and/or negative trends for most developing countries since the 1960s. The empirical implication is that there cannot be a long-run relationship between the growth rate of output and the level of the aid/GDP ratio over time; such unbalanced data (with stationary and nonstationary variables) can, even in cross-country analyses, lead to misleading results (Ericsson et al., 2001). In addition, and equally important, several recent contributions to the theoretical growth literature focus on levels instead of growth rates (e.g., Acemoglu and Ventura, 2002).

The rest of the paper is organized as follows: Section 2 provides a brief literature review to identify the potential ways in which foreign aid can affect domestic output. Section 3 sets out the empirical specification to estimate the long-run effect of aid on output and discusses the data.

Section 4 presents the estimates of the long-run effect of aid on output. Section 5 contains the empirical analysis of the determinants of the estimated long-run aid effectiveness. Section 6 concludes with a discussion of the results.

## 2. How Can Foreign Aid Affect Domestic Output?

The recent aid effectiveness literature starts from a standard growth model; as aid is not a variable in such models, it is introduced as a component of investment (Burnside and Dollar, 2000; Rajan and Subramanian, 2008). The assumption is that in a reduced form context aid finances the investment that determines growth. The potential effect of aid on domestic output can be represented in an aggregate production function of the form:

$$Y_t = B_t K_t^{\theta}, \tag{1}$$

where $Y_t$ is output, $K_t$ is the capital stock, and $B_t$ is a productivity parameter. For simplicity, assume that capital depreciates fully each period, so that the end-of-period capital stock is equal to domestic investment, $K_t = I_t$. Assuming further that domestic investment is the aggregate of public and private investment and that public investment is financed by taxes and foreign aid (since aid is primarily given to the government), the production function can be written as:

$$Y_t = B_t(Itax_t + A_t\phi + Iprivate_t)^{\theta}, \tag{2}$$

where $Itax_t$ is tax-financed public investment, $A_t$ represents foreign aid, $\phi$ is the share of aid that is used for public investment purposes, and $Iprivate_t$ is private investment. Equation (2) now shows succinctly how aid can affect output: it can directly increase output (if $\phi > 0$), and it may also influence output indirectly by affecting the amount of tax-financed public investment, the volume of private investment, and the level of productivity.

Accordingly, the direct effect of aid depends on how much of aid is invested by the government. It is common in the literature to assume that aid is intended to finance investment so that if aid is redirected to government consumption (fungibility) this reduces effectiveness, the direct effect on domestic output (see Burnside and Dollar, 2000). This, however, is misleading as government consumption includes expenditures to 'maintain and operate' investment projects, especially in social sectors. Public investment spending is mostly construction costs, whereas the recurrent costs essential for productive investment are included as consumption. Furthermore, a large proportion of aid is not intended to finance capital investment, so the fungibility argument is

generally misguided. Donors direct much of their aid to social sectors (consumption spending that could be considered as investment in human capital) or technical assistance (which may contribute to capacity building and aggregate productivity, although much of this aid is actually spent in the donor rather than the recipient). In a reduced form context, the primary issue is not what class of spending the aid is allocated to but the type of investment (if any) it supports as this determines the expected lag before any impact on output should be observed.

Two issues arise, one relating to how aid is measured and the other to the treatment of investment. Studies following Burnside and Dollar (2000) use standard measures of total aid, so no attempt is made to identify aid that is directed at investment, and typically omit any investment variable from the estimated model on the basis that some portion of investment is directly financed by aid. This imposes the assumption that aid is a proxy for investment; this is problematic because not all aid finances investment and not all investment is financed by aid, and the respective shares will vary across countries and over time.[1]

Even if aid is fully invested, it may or may not increase output. The inflow of aid will affect government decisions on expenditure, tax effort and even borrowing — phenomena widely recognized in the theoretical and empirical literature on the fiscal consequences of foreign aid (McGillivray and Morrissey, 2004) — which may in turn affect the levels of public and private sector investment, and thereby output. Many effects are possible and the net impact on output is an empirical matter that is likely to vary from country to country (and over time).

Foreign aid can affect output through the level of productivity in at least four ways. First, assuming that productivity is determined (inter alia) by the quality of institutions, aid may affect productivity through effects on institutions. As with investment, the direction of any net effect is far from obvious. While many studies find that aid is associated with weak institutions and low productivity (e.g. Bräutigam and Knack, 2004; Rajan and Subramanian, 2007), deep endogeneity implies severe caution in drawing inferences from such studies; it may simply reflect the tendency of aid to go to poorer countries (with weaker institutions). In contrast, aid that supports capacity-building, improved public sector management and policy reform could have a (perhaps gradual) effect of increasing the quality of governance and institutions. Second, and related, aid can affect productivity if it encourages diversion of resources from productive activities to unproductive rent

[1] Gomanee et al. (2005b) address this concern explicitly by including in the measure of aid only those forms that finance physical or human capital investment and by removing from the measure of public investment the proportion that is explained by aid. In a regression with aid and investment that is not aid-financed, they find that both have a positive impact on growth for a sub-Saharan African sample and that the small magnitude of aid effectiveness is largely explained by very low productivity of investment.

seeking; here the focus is on corruption or even conflict rather than institutions more generally. Again, distinguishing cause and effect, and attributing a causal role to aid, is difficult.[2] One should certainly caution against generalizations. Third, aid can reduce the productivity of an economy through Dutch Disease effects, if aid reduces the competitiveness of the more productive tradable goods sector and thus results in a reallocation of resources towards the less productive non-tradable goods sector. The empirical evidence for Dutch Disease effects of aid is inconclusive, and it is quite possible that aid may generate positive externalities for the tradable sector (see Adam and O'Connell, 2004). Finally, aid can cause an increase in productivity by relaxing the binding foreign exchange constraint and allowing the country to increase its imports of capital goods. Capital goods imports from high-income countries are typically associated with higher productivity in developing countries, since capital goods embody technological know-how (e.g. Almeida and Fernandes, 2008).

In summary, the effect of foreign aid on domestic output is theoretically ambiguous and depends on several factors, the most important of which are effects on the level and productivity of investment. Institutional factors are clearly important, but it is not unreasonable to assume that these largely influence the productivity of investment, howsoever financed. Furthermore, institutions tend to display strong persistence within countries over time. It follows that the effect of aid on output may differ significantly from country to country, and that the output effect of aid may change over time. These issues will be addressed empirically in the analysis that follows.

## 3. Empirical Specification: Aid, Investment and Output

The analysis will examine the long-run effect of foreign aid on domestic output using heterogeneous panel cointegration techniques to control for omitted variable and endogeneity bias and to detect possible cross-country differences in the long-run output effect of aid. This section presents the basic empirical model and the data.

It is common practice in panel cointegration studies to estimate a bivariate long-run relationship (see Herzer, 2008). However, this would be inappropriate for aid given the discussion in the preceding section. It would not be reasonable to assume, or estimate as if, aid is the major determinant of output. However, as it is necessary to employ a very parsimonious specification, it is not unreasonable to assume that investment is a primary determinant of output over time, and aid is

[2] Maren (1997), for example, argues that Somalia's civil war was caused by the desire of different factions to control the large amount of food aid the country was receiving. However, the country needed food aid because of inherent (prior) economic and political governance weaknesses.

the element of investment of particular concern. Moreover, since investment may act as a proxy for a number of unobservable time-varying factors that can affect both aid allocation decisions and output, it should be included in the analysis to control for omitted variable bias.[3] Thus, we consider a model of the form:

$$Y_{it} = a_i + \delta_i t + \beta_1 A_{it} + \beta_2 I_{it} + \varepsilon_{it}, \quad (3)$$

where $Y_{it}$ is the log of real GDP (the output measure) over time periods $t = 1, 2, \ldots, T$ and countries $i = 1, 2, \ldots, N$. Given that it is not possible to identify the proportion of aid that actually finances investment, nor the amount of investment that is not financed by aid, $A_{it}$ is represented by the standard measure of aid — the percentage share of net Official Development Assistance (ODA) in GDP, and $I_{it}$ is total investment as a share of GDP.[4] The $\beta$ coefficients represent the cross-country average of the effects of aid and investment on GDP respectively, which are allowed to be country specific and thus to vary across countries. Moreover, we include country-specific fixed effects, $a_i$, and country-specific deterministic time trends, $\delta_i t$, to control for any country-specific omitted factors that are relatively stable over time or evolve smoothly over time. Data on both GDP and aid come from the World Development Indicators 2009 database,[5] while the investment data are from the Heston, Summers, and Aten (2006) Penn World Table. These measures are imperfect but can be interpreted as the effect of aid on output given investment.

An important feature of equation (3) is the assumption that, in the long-run, permanent changes in the aid/GDP ratio and the investment/GDP ratio are associated with permanent changes in the log-level of GDP. Econometrically, this implies that the individual time series for aid (relative to GDP), investment (relative to GDP) and GDP (in logs) must exhibit unit-root behavior and that $A_{it}$ and $I_{it}$ must be cointegrated with $Y_{it}$. Details of the unit root tests are in Appendix A1; the result is a panel of 59 countries ($N = 59$) for which $A_{it}$, $I_{it}$, and $Y_{it}$ behave as random walks, with data from 1971 to 2003 ($T = 33$).

[3] The parameter for the effect of investment on output may act as a proxy for unobservable country-specific factors.
[4] As the levels of aid and investment are both included in GDP (by accounting convention), using levels would create a problem as effectively the explanatory variables are part of the independent variable. Furthermore, it is necessary to deflate aid and investment; dividing by GDP seems appropriate to give a measure of the relative importance of the levels.
[5] Aid as a share of GDP was calculated by multiplying aid per capita (in current US dollars) by population and dividing by GDP (in current US dollars).

**Table 1 Countries and country summary statistics**

| | Average log of GDP | Average aid/GDP ratio | Average investment/ GDP ratio | | Average log of GDP | Average aid/GDP ratio | Average investment/ GDP ratio |
|---|---|---|---|---|---|---|---|
| Algeria | 24.41 | 0.60 | 17.80 | Liberia | 20.35 | 21.62 | 6.19 |
| Belize | 19.77 | 6.68 | 14.84 | Madagascar | 21.86 | 8.38 | 4.18 |
| Benin | 21.04 | 9.63 | 10.06 | Malawi | 20.85 | 17.94 | 10.31 |
| Bolivia | 22.53 | 6.45 | 10.07 | Malaysia | 24.37 | 0.45 | 24.05 |
| Botswana | 21.48 | 6.88 | 22.21 | Mali | 21.22 | 16.39 | 9.12 |
| Burkina Faso | 21.12 | 12.64 | 10.43 | Mauritania | 20.49 | 21.86 | 14.55 |
| Burundi | 20.31 | 16.13 | 4.90 | Morocco | 23.93 | 2.72 | 13.83 |
| Cameroon | 22.73 | 4.05 | 6.18 | Nicaragua | 21.956 | 12.98 | 10.01 |
| Cent. African Rep. | 20.50 | 12.11 | 9.36 | Niger | 21.13 | 12.96 | 7.87 |
| Chad | 20.75 | 11.76 | 10.21 | Nigeria | 24.21 | 0.49 | 7.24 |
| Chile | 24.36 | 0.20 | 18.39 | Pakistan | 24.41 | 3.00 | 12.92 |
| Congo, Dem. Rep. | 22.58 | 6.97 | 9.09 | Panama | 22.73 | 1.05 | 19.17 |
| Congo, Rep. | 21.54 | 6.14 | 22.13 | Papua N. Guinea | 21.57 | 11.63 | 8.58 |
| Costa Rica | 22.92 | 1.77 | 9.93 | Peru | 24.43 | 1.05 | 17.07 |
| Cote d'Ivoire | 22.82 | 4.23 | 6.98 | Philippines | 24.66 | 1.48 | 15.78 |
| Dominican Rep. | 23.10 | 1.20 | 12.37 | Rwanda | 21.09 | 18.16 | 3.01 |
| Ecuador | 23.20 | 1.15 | 21.49 | Senegal | 21.92 | 9.74 | 5.71 |
| Egypt | 24.68 | 6.44 | 8.14 | Sierra Leone | 20.59 | 13.19 | 4.10 |
| El Salvador | 22.97 | 4.20 | 8.31 | Solomon Islands | 19.07 | 22.79 | 8.59 |
| Gabon | 22.13 | 2.13 | 9.26 | South Korea | 26.08 | 0.41 | 33.60 |
| Gambia | 19.39 | 19.81 | 7.82 | Sri Lanka | 22.89 | 5.71 | 14.77 |
| Ghana | 21.90 | 6.97 | 5.73 | Sudan | 22.69 | 5.38 | 13.28 |
| Guatemala | 23.26 | 1.45 | 8.01 | Swaziland | 20.462 | 5.19 | 11.91 |
| Honduras | 22.09 | 6.78 | 14.21 | Syria | 23.15 | 4.88 | 8.81 |
| India | 26.18 | 0.76 | 11.65 | Thailand | 24.80 | 0.74 | 31.05 |
| Indonesia | 25.18 | 1.50 | 18.48 | Togo | 20.70 | 9.67 | 11.32 |
| Jamaica | 22.59 | 2.94 | 16.94 | Tunisia | 23.11 | 2.68 | 18.68 |
| Kenya | 22.86 | 6.34 | 12.53 | Uruguay | 23.46 | 0.30 | 14.37 |
| Lesotho | 20.02 | 17.44 | 20.64 | Zambia | 21.79 | 13.91 | 16.29 |
| | | | | Zimbabwe | 22.43 | 3.05 | 14.16 |

Table 1 lists the countries along with the average values for $Y_{it}$, $A_{it}$, and $I_{it}$ over the period from 1971 to 2003. As expected, there are large cross-country differences in the values of these parameters. Investment accounts for more than 30 percent of GDP in South Korea and Thailand, while several countries, such as Rwanda, Sierra Leone, and Madagascar, are characterized by extremely low investment rates of about 4 percent. Also striking are the differences in the average aid/GDP ratios. The Solomon Islands is the most aid-dependent economy, with aid amounting to more than 22 percent of GDP, followed by Mauritania, Liberia, Gambia, Rwanda, Malawi, Lesotho, Mali, and Burundi. All these countries have an average ratio of aid to GDP of more than 15 percent, whereas in Chile, Uruguay, South Korea, Malaysia, Nigeria, Algeria, Thailand, and India aid accounts for less than 1 percent of GDP. Interestingly, in many countries (Liberia, Rwanda, Solomon Islands, Gambia, Burundi, Sierra Leone, Malawi, Mauritania, Mali, Niger, Madagascar,

Senegal, Papua New Guinea, Nicaragua, Central African Republic, Burkina Faso, Chad, and Ghana), the aid/GDP ratio is greater than the investment/GDP ratio, suggesting that a large part of aid is used for consumption rather than investment, at least in these countries. Moreover, countries with higher GDP tend to have lower aid/GDP ratios, and vice versa. India, South Korea, and Thailand, for example, are among the largest economies and the least aid-dependent countries in our sample. The Solomon Islands, in contrast, is the country with the lowest GDP and the highest share of aid in GDP.

Figures A1-A3 in the Appendix plot $Y_{it}$, $A_{it}$, and $I_{it}$ for the period from 1971 to 2003. As can be seen from Figure A1, GDP increased in all countries, with the exception of Liberia (row 4, column 6) and the Democratic Republic of Congo (row 3, column 4) where GDP declined between 1971 and 2003. Noteworthy are also Sierra Leone (row 6, column 7) and Rwanda (row 6, column 5): in Sierra Leone, GDP showed an increasing trend until 1991; fell from 1991 to 1999, and then rose sharply from 1999 to 2003. Similarly, Rwanda's GDP rose from 1971 to 1991, dropped abruptly between 1991 and 1994 (note that this preceded the genocide), and then rose rapidly from 1994 to 2003. Thus, a characteristic feature of the evolution of GDP in most developing countries is volatility; that is, GDP increased but not steadily. This instability is even more characteristic of aid and investment. Figures A2 and A3 show that $A_{it}$ and $I_{it}$ exhibit positive and/or negative trends as well as strong deviations from these trends. Overall, the time-series evolution is consistent with the possibility that $Y_{it}$, $A_{it}$, and $I_{it}$ are nonstationary and cointegrated, which is also confirmed by several panel-unit root and panel cointegration tests reported in Appendices A1 and A2.

## 4. Panel Cointegration Results

The long-run effect of aid on GDP is estimated using the between-dimension group-mean panel DOLS estimator that Pedroni (2000, 2001) argues has a number of advantages over the within-dimension approach. First, it allows for greater flexibility in the presence of heterogeneous cointegrating vectors, whereas under the within-dimension approach the cointegrating vectors are constrained to be the same for each country. Second, the point estimates provide a more useful interpretation in the case of heterogeneous cointegrating vectors, as they can be interpreted as the mean value of the cointegrating vectors, which does not apply to the within estimators. Third, between-dimension estimators suffer from much lower small-sample size distortions than is the case with the within-dimension estimators.

The panel DOLS regression is given by

$$Y_{it} = a_i + \delta_i t + \beta_{1i} A_{it} + \beta_{2i} I_{it} + \sum_{j=-k_i}^{k_i} \Phi_{1ij} \Delta A_{it-j} + + \sum_{j=-k_i}^{k_i} \Phi_{2ij} \Delta I_{it-j} + \varepsilon_{it}, \quad (4)$$

where $\Phi_{1ij}$ and $\Phi_{2ij}$ are coefficients of lead and lag differences which account for possible serial correlation and endogeneity of the regressors. Thus, an important feature of the DOLS procedure is that it generates unbiased estimates for variables that cointegrate even with endogenous regressors. Consequently, in contrast to cross-section and conventional panel approaches, the approach does not require exogeneity assumptions nor does it require the use of instruments. In addition, the group-mean panel DOLS estimator is superconsistent under cointegration, and is robust to the omission of variables that do not form part of the cointegrating relationship. It is calculated as

$$\hat{\beta}_m = N^{-1} \sum_{i=1}^{N} \hat{\beta}_{mi}, \quad (5)$$

where

$$t_{\hat{\beta}_m} = N^{-1/2} \sum_{i=1}^{N} t_{\hat{\beta}_{mi}} \quad (6)$$

is the corresponding *t*-statistic of $\hat{\beta}_m$ ($m$ = 1, 2) and $\hat{\beta}_{mi}$ is the conventional time-series DOLS estimator applied to the *i*th country of the panel. Stock and Watson (1993) find that this estimator performs well in small samples (as here) compared with other cointegration estimators.

The DOLS procedure is applied to both the raw data and to data that have been demeaned over the cross-sectional dimension; that is, in place of $Y_{it}$, $A_{it}$, and $I_{it}$, we also use

$$Y'_{it} = Y_{it} - \overline{Y_t},$$

$$A'_{it} = A_{it} - \overline{A_t}, \text{ and}$$

$$I'_{it} = I_{it} - \overline{I_t}, \text{ where}$$

$$\overline{Y_t} = N^{-1} \sum_{i=1}^{N} Y_{it},$$

$$\overline{A_t} = N^{-1} \sum_{i=1}^{N} A_{it}, \text{ and}$$

$$\overline{I_t} = N^{-1} \sum_{i=1}^{N} I_{it}, \quad (7)$$

to account for cross-sectional dependence due to common shocks or spillovers among countries at the same time. The estimates are reported in Table 2. As can be seen, the unadjusted and demeaned

data produce similar results. The coefficients on $I_{it}$ are highly significant and positive, as expected, whereas the coefficients on $A_{it}$ are highly significant and negative. More precisely, the demeaned data yield an aid coefficient of -0.0081, implying that, in the long-run, a one percentage point increase in the aid-to-GDP ratio leads to a decrease in GDP by 0.0081 percent.

**Table 2 DOLS estimates of the coefficients on aid and investment**

| | $A_{it}$ | $I_{it}$ |
|---|---|---|
| Unadjusted data | -0.0064** (-6.64) | 0.0235** (16.96) |
| Demeaned data | -0.0081** (-9.38) | 0.0170** (15.80) |

*Notes:* The dependent variable is $Y_{it}$. ** indicate significance at the 1% level. *t*-statistics in parentheses. The number of leads and lags in the individual DOLS regressions was determined by the Schwarz criterion with a maximum of three lags.

Appendix A3 presents a set of sensitivity analysis demonstrating that the negative effect of aid on output is not due to individual outliers, sample-selection bias (if a group of countries in a particular region have a significant effect) or the time period, and that the results are similar using private investment data in place of total investment (albeit for a smaller sample). However, it is worth explicitly considering if the results are robust to alternative estimation techniques. For this purpose, a conditional error-correction model is used, regressing $\Delta Y_{it}$, on $Y_{it-1}$, $A_{it-1}$, and $I_{it-1}$, the first differences of $A_{it}$ and $I_{it}$, the lagged first differences of $A_{it}$, $I_{it}$, and $Y_{it}$, individual time trends, and individual intercepts.[6] Computing the group-mean coefficients (and *t*-statistics), yields the following equation (** (*) indicate significance at the 1% (5%) level, *t*-statistics in parentheses):

$$\begin{aligned}\Delta Y_{it} = & \underset{(-16.29)}{-0.3267^{**}Y_{it-1}} \underset{(-3.60)}{-0.0026^{**}A_{it-1}} \underset{(6.52)}{+0.0043^{**}I_{it-1}} \\ & \underset{(-7.01)}{-0.0037^{**}\Delta A_{it}} \underset{(11.86)}{+0.0063^{**}\Delta I_{it}} \underset{(2.29)}{+0.0015^{*}\Delta A_{it-1}} \underset{(-1.09)}{-0.0010\Delta I_{it-1}} \\ & \underset{(5.11)}{+0.1309^{**}\Delta Y_{it-1}} + a_i + \delta_i t\end{aligned} \tag{8}$$

The estimated coefficient on the lagged dependent variable, $Y_{it-1}$, is negative and highly significant; normalizing on $Y_{it}$, yields the following significant long-run relationship:

$$Y_{it} = a_i + \delta_i t - 0.0080 A_{it} + 0.0132 I_{it}\,. \tag{9}$$

[6] We use the demeaned data to account for the likely cross-sectional dependence through common time effects.

As the estimated aid coefficient is almost identical to the result of the DOLS procedure, it can be concluded that the negative long-run effect of aid on output is robust to different estimation techniques. Thus, the results appear to be robust to potential outliers, sample selection, the specification of the empirical model, the sample period, and different estimation techniques. This negative effect for the sample as a whole does not, however, imply that aid exerts a negative effect on GDP in each individual country.

**Table 3 DOLS country estimates and stability tests**

| Country | $A_{it}$ | *t*-stat | *MeanF* | Country | $A_{it}$ | *t*-stat | *MeanF* |
|---|---|---|---|---|---|---|---|
| Algeria | -0.0682** | -2.54 | 7.49* | Liberia | -0.0311** | -5.46 | 0.94 |
| Belize | 0.0267 | 0.78 | 5.88 | Madagascar | -0.0339 | -0.70 | 2.34 |
| Benin | -0.0118 | -0.80 | 3.62 | Malawi | -0.0028 | -1.23 | 1.12 |
| Bolivia | -0.0720** | -4.93 | 1.62 | Malaysia | 0.0157 | 1.51 | 14.48** |
| Botswana | -0.0344** | -5.13 | 24.25** | Mali | -0.0277** | -4.11 | 4.18 |
| Burkina Faso | 0.0050 | 0.94 | 7.26* | Mauritania | -0.0012 | -1.13 | 1.06 |
| Burundi | 0.0453** | 4.91 | 0.80 | Morocco | -0.0209** | -4.00 | 3.08 |
| Cameroon | -0.0184 | -0.94 | 2.59 | Nicaragua | -0.0088** | -3.50 | 4.85 |
| Central African Rep. | -0.0039 | -0.59 | 0.53 | Niger | -0.032** | -7.20 | 0.55 |
| Chad | 0.0457* | 2.17 | 3.15 | Nigeria | -0.0380 | -2.00 | 3.06 |
| Chile | 0.0592** | 3.65 | 2.88 | Pakistan | -0.0393** | -10.35 | 0.97 |
| Congo, Dem. Rep. | 0.0769** | 3.30 | 3.86 | Panama | -0.0612 | -1.12 | 6.69* |
| Congo, Rep. | -0.0548** | -5.79 | 2.95 | Papua New Guinea | -0.0019 | -0.18 | 15.98** |
| Costa Rica | -0.0681** | -3.44 | 5.58 | Peru | 0.0416** | 3.72 | 3.83 |
| Cote d'Ivoire | -0.0153** | -3.08 | 0.80 | Philippines | 0.0194* | 2.39 | 4.33 |
| Dominican Republic | 0.0283** | 5.83 | 3.04 | Rwanda | -0.0209** | -4.42 | 6.43* |
| Ecuador | -0.0297** | -7.36 | 3.86 | Senegal | -0.0362* | -2.74 | 1.44 |
| Egypt | -0.0303** | -3.52 | 1.21 | Sierra Leone | -0.0560* | -2.70 | 2.27 |
| El Salvador | -0.0473** | -4.27 | 26.21** | Solomon Islands | -0.0434** | -3.10 | 4.65 |
| Gabon | 0.0429 | 0.80 | 1.48 | South Korea | 0.0057 | 0.18 | 5.80 |
| Gambia | 0.0012 | 1.20 | 1.56 | Sri Lanka | -0.0259* | -2.17 | 0.43 |
| Ghana | -0.0023 | -0.11 | 1.94 | Sudan | -0.0244* | -2.27 | 6.88* |
| Guatemala | 0.0741* | 2.99 | 4.44 | Swaziland | -0.0405** | -4.18 | 3.43 |
| Honduras | -0.0152* | -2.12 | 3.41 | Syria | 0.0096 | 1.80 | 2.85 |
| India | 0.0181 | 1.39 | 7.94* | Thailand | 0.0333 | 0.63 | 8.66** |
| Indonesia | -0.0390 | -1.39 | 3.08 | Togo | -0.0007 | -0.09 | 5.69 |
| Jamaica | -0.0343** | -11.19 | 0.74 | Tunisia | -0.0052* | -2.15 | 0.88 |
| Kenya | 0.0336** | 8.67 | 4.46 | Uruguay | 0.0116 | 1.27 | 4.19 |
| Lesotho | 0.0006 | 0.27 | 1.44 | Zambia | 0.0007 | 0.36 | 1.83 |
| | | | | Zimbabwe | 0.0229 | 1.21 | 8.50* |

*Notes:* The number of leads and lags was determined by the Schwarz criterion with a maximum of three lags. *MeanF* is a Chow-type test for parameter constancy in cointegrating regressions. The 5% (1%) critical value for the stability test (*MeanF*) is 6.22 (8.61) (Hansen, 1992). ** (*) indicate significance at the 1% (5%) level.

The individual country DOLS point estimates of the output effect of aid are presented in Table 3. The most striking feature of these estimates is the heterogeneity in the coefficients, ranging from -0.072 in Bolivia to 0.077 in the Democratic Republic of Congo. Thus, although the long-run

effect of aid on GDP is negative in general or on average in developing countries, aid does not have a negative long-run effect on GDP in all countries. More precisely, for 37 out of 59 countries (and thus in 63 percent of cases) an increase in aid is associated with a decrease in GDP, while in 22 cases (37 percent of the countries) an increase in aid is associated with an increase in GDP. Even within the country groups with negative and positive effects, the individual country estimates show considerable heterogeneity. For example, the point estimates suggest that Chile and Guatemala benefit markedly from aid (although both received quite small or negligible amounts of aid). In contrast, in many countries, such as Lesotho and Togo (both recipients of relatively large amounts of aid) both the positive and negative effects are marginal (close to 0), whereas in many other countries, such as Algeria (negligible aid) and Sierra Leone (large aid) aid has a strong negative effect on GDP. There is a tendency for the output response to be positive (negative) for countries receiving small (large) amounts of aid, but there are exceptions.

In light of the finding that the effect of aid on GDP is not constant across countries, a natural question is whether it is constant over time. To answer this question, we compute for each country-DOLS regression the *MeanF* test developed by Hansen (1992). This test is a Chow-type test for parameter constancy in cointegrating regressions with unknown change points and is designed to detect any gradual changes in the regression coefficients.[7] The results of this test are reported in columns 4 and 8 of Table 3. They show that the null hypothesis of parameter stability is rejected at least at the five percent level in about 20 percent of cases, suggesting that in some countries the effect of aid on GDP has changed over time. The most plausible explanation for this finding is that the effect of aid on GDP depends on political and institutional factors. If policies and institutions affecting the aid-output relationship change substantially over time, then also the effect of aid on output changes over time. The country-specific factors affecting the effect of aid on GDP are examined in the next section.

## 5. Determinants of the Long-Run Effect of Aid on Output

This section systematically searches for country-specific factors that help to explain the cross-country differences in the output effect of aid; that is, determinants of aid effectiveness. These determinants have been investigated by several studies using cross-country growth regressions, some including interaction terms between aid and a small number of potential determinants of aid

[7] Hansen (1992) develops the stability tests using the FMOLS estimator. Because the DOLS estimator is asymptotically equivalent to the FMOLS estimator, the test statistics have the same distributions and are thus applicable to both estimators.

effectiveness. A limitation of the conventional interaction-term approach, however, is the inability to empirically identify which independent variable in the interaction term determines the effect of the other independent variable on the dependent variable. For example, a statistically significant interaction term between aid and economic policy does not necessarily imply that aid effectiveness depends on economic policy. A statistically significant aid-policy interaction term can also be compatible with the growth effect of economic policy being influenced by aid. A different approach employs cross-section regressions with the output effect of aid as the dependent variable to consider a large number of possible determinants of aid effectiveness. By including as many variables as possible relevant to aid effectiveness and using as the dependent variable the *output effect* of aid instead of *output growth*, the approach is less subject to omitted variable and endogeneity bias than the conventional interaction-term approach.

The twenty variables considered to be potentially relevant to aid effectiveness are only summarized here; details and the country composition of the sample is in Appendix A4, with variables, definitions and sources in Appendix Table A6. As the economic policy index as used by Burnside and Dollar (2000) may mask differences in the influence of the components, separate indicators of monetary policy, fiscal policy and trade policy are employed here: inflation, the budget balance and liberalization dates (the Sachs and Warner openness index) respectively. These are the 'policy' variables.

Political and institutional factors are likely to be important. The level of democracy is measured by the democracy index from the POLITY IV data base, with values from 0 (absence of democracy) to 10 (highest level of democracy), while political instability is measured as the number of revolutions and coups. The quality of institutions is obviously important but as different measures capture distinct features, several institutional variables from the International Country Risk Guide (ICRG) are included: Socioeconomic conditions, Investment profile, Internal conflict, External conflict, Corruption, (influence of the) Military in politics, Law and order, Religious and Ethnic tensions, and Bureaucratic quality. The indicators for corruption, internal conflict, external conflict, military in politics, religious tensions, and ethnic tensions are rescaled so that higher values always reflect higher institutional quality.

A number of other factors are considered. Geographical or structural features are captured by the share of a country's area that is in the tropics (Dalgaard et al., 2004). The annual variability of aid may influence effectiveness, so aid uncertainty is measured by the standard deviation of the residuals of a regression of aid (as a percentage of GDP) on two lags of aid, a constant term, and a linear time trend (Lensink and Morrissey, 2000). GDP per capita is included as the marginal

effectiveness of aid may vary with income level (Gomanee et al., 2005a). As adverse effects of rent-seeking or intervention may relate to government size the share of government consumption in GDP is included. Finally, absorptive capacity is measured by the secondary school enrolment rate.

The dependent variable is the estimated long-run effect of aid on GDP from Table 3, $\hat{\beta}_i$. As discussed in Section 4, this effect can be assumed to be time-constant in 80 percent of the countries in our sample and thus be treated as average impact per year. For the remaining 20 percent, the estimated output effect of aid is not constant; nevertheless, it can be interpreted as a time average over the period 1971-2003. Consequently, we also use time averages for the independent variables in that period. Unfortunately, however, we do not have complete data on all variables for all countries, limiting the sample to 45 countries.

Given the large number of variables, all of which are potential determinants of aid effectiveness, we are confronted with the classical problem of variable selection — the problem of identifying those variables that are, in fact, important regressors for explaining the cross-country variations in the long-run effect of aid on GDP. To deal with this problem, we use the general-to-specific variable selection approach suggested by Hoover and Perez (2004). Hoover and Perez show by means of Monte Carlo simulations that their approach is very effective in identifying the true parameters of the data generating process, outperforming other variable selection procedures such as the extreme bounds approaches of Levine and Renelt (1992) and Sala-i-Martin (1997).

Following Hoover and Perez (2004), we start by estimating a general specification in which all variables are included, and simplify it by removing insignificant variables. The variables are first ranked according to their *t*-statistics,[8] then five simplification paths are applied in which each of the five variables with the lowest *t*-statistics is the first to be removed, yielding five equations. From these equations, variables with insignificant coefficients are then eliminated sequentially according to the lowest *t*-values until the remaining variables are significant at least at the 5% level. After removal of each variable, a battery of specification tests is performed, including a Jarque-Bera test (*JB*) for normality of the residuals, a Ramsey RESET test for general nonlinearity and functional form misspecification (*RESET*), and a sub-sample stability test (*STABILITY*) using an *F*-test for the equality of the variances of the residuals of sub-Saharan countries versus the rest of the sample. Furthermore, an *F*-test of the hypothesis that the current specification is a valid restriction of the general specification is used after each step (*RESTRICTION*). In our case, all of these tests are

[8] Since an estimated dependent variable may introduce heteroskedasticity into the regressions (see Saxonhouse, 1976), we use White's heteroskedasticity-consistent standard errors to compute the *t*-statistics.

passed, implying five well-specified parsimonious equations, which are all valid restrictions of the general model. Finally, we construct the non-redundant joint model from each of these equations by taking all specifications and performing the *F*-test for encompassing the other specifications. This procedure yields the final specification in Table 4.

**Table 4 General-to-specific approach: final specification**

| Independent variables | Coefficients (*t*-statistics) |
|---|---|
| *Rel* | 0.0100** (2.93) |
| *Law* | 0.0137* (2.32) |
| *Size* | -0.0025* (-2.04) |
| Diagnostic tests | |
| Adj. $R^2$ | 0.21 |
| *JB* ($\chi^2_{(2)}$) | 0.61 [0.74] |
| *RESET* ($\chi^2_{(1)}$) | 0.12 [0.73] |
| *STABILITY* | $F(24, 19) = 1.02$ [0.97] |
| *RESTTRICTION* | $F(18, 24) = 0.88$ [0.61] |

*Notes:* Reported *t*-statistics (in parentheses) are based on White's heteroskedasticity-consistent standard errors. ** (*) indicate significance at the 1% (5%) level. *JB* is the Jarque-Bera test for normality, *RESET* is the usual test for general nonlinearity and misspecification, *STABILITY* is an *F*-test for the equality of the variances of sub-Saharan countries versus the rest of the sample, and *RESTRICTION* is an *F*-test of the hypothesis that the model is a valid restriction of the general model. Numbers in brackets behind the values of the diagnostic test statistics are the corresponding *p*-values. The sign of the coefficient on Rel is positive, since a higher number indicates lower religious tensions.

The final model passes all the diagnostic tests. The assumption of normally distributed residuals cannot be rejected, and the Ramsey RESET test does not suggest nonlinearity or misspecification. The model also passes the *F*-test for parameter stability and the *F*-test that the final model is a valid restriction of the general model. In addition, the recursive residuals in Figure A5 persistently lie within the error bounds of –2 and +2 standard errors, suggesting that no outliers are present. Consequently, statistically valid inferences can be drawn from the results in Table 4.

All coefficients have the expected sign: a higher level of law and order (*Law*) is associated with greater aid effectiveness whereas government size (*Size*) and religious tensions (*Rel*) are negatively related to the output effect of aid (the sign of the coefficient on *Rel* is positive as a higher number indicates lower religious tensions). The results suggest that cross-country variations in religious tensions, law and order, and government size are important factors in explaining the cross-country differences in the long-run effect of aid on GDP.

The output effect of aid appears not to depend (directly) on the quality of monetary policy (measured by inflation), fiscal policy (measured by the budget balance), trade policy (measured by the Sachs and Warner openness index), the share of a country's area that is in the tropics, the level

of democracy, political instability, socioeconomic conditions, investment profile, internal and external conflicts, corruption, the influence of the military in politics, ethnic tensions, the institutional strength and quality of the bureaucracy, aid uncertainty, GDP per capita, and absorptive capacity (measured by the secondary school enrolment rate). All these variables turned out to be insignificant and hence were removed from the general model.

**Table 5 General specification and correlation coefficients**

| Regressors | *t*-statistics | Correlation coefficients between the variables of the final model and each variable that was omitted from the final model (*t*-statistics) | | |
|---|---|---|---|---|
| *Rel* | 2.50* | | | |
| *Law* | 1.91 | *Rel* | *Law* | *Size* |
| *Size* | -1.89 | | | |
| *Inflation* | 0.75 | 0.23 (1.52) | -0.14 (-0.91) | -0.15 (-0.98) |
| *Budget* | -0.54 | 0.12 (0.81) | 0.25 (1.68) | -0.13 (0.83) |
| *Openness* | -1.61 | 0.18 (1.18) | 0.50** (3.76) | -0.26 (-1.74) |
| *Tropics* | -1.01 | 0.26 (1.77) | -0.32* (-2.18) | -0.23 (-1.56) |
| *Demo* | 0.39 | 0.19 (1.26) | 0.10 (0.64) | -0.10 (-0.64) |
| *Instability* | -0.38 | 0.09 (0.59) | -0.30* (-2.09) | -0.15 (-0.10) |
| *Socio* | 1.18 | 0.06 (0.39) | 0.53** (4.08) | 0.01 (0.06) |
| *Invprof* | 0.16 | 0.23 (1.58) | 0.49** (3.70) | 0.04 (0.27) |
| *Intconf* | -0.67 | 0.22 (1.48) | 0.67** (5.97) | 0.37* (2.63) |
| *Extconf* | 0.74 | 0.14 (0.94) | 0.11 (0.75) | -0.10 (-0.64) |
| *Corr* | -0.77 | 0.25 (1.70) | 0.47** (3.50) | 0.14 (0.91) |
| *Military* | -0.54 | 0.31* (2.12) | 0.39** (2.79) | 0.25 (1.72) |
| *Ethnic* | -0.94 | 0.29 (1.96) | 0.40** (2.88) | 0.20 (1.35) |
| *Bureau* | 0.55 | 0.04 (0.23) | 0.51** (3.82) | 0.07 (0.46) |
| *Uncertain* | 1.40 | 0.03 (0.20) | -0.18 (-1.19) | 0.35* (2.44) |
| *GDPpc* | 0.13 | 0.34* (2.37) | 0.19 (1.23) | -0.07 (-0.45) |
| *Capacity* | 0.12 | 0.15 (0.98) | 0.18 (1.19) | -0.16 (-1.09) |
| Diagnostic tests | | | | |
| Adj. $R^2$ | 0.10 | | | |
| *JB* ($\chi^2_{(2)}$) | 3.30 [0.19] | | | |
| *RESET* ($\chi^2_{(1)}$) | 0.49 [0.49] | | | |
| *STABILITY* | F(19, 24) = 1.30 [0.54] | | | |

** (*) indicate significance at the 1% (5%) level. Reported *t*-statistics in the second column are based on White's heteroskedasticity-consistent standard errors. *JB* is the Jarque-Bera test for normality, *RESET* is the usual test for general nonlinearity and misspecification, *STABILITY* is an *F*-test for the equality of the variances of sub-Saharan countries versus the rest of the sample. Numbers in brackets behind the values of the diagnostic test statistics are the corresponding *p*-values. The indicators for corruption, internal conflict, external conflict, military in politics, religious tensions, and ethnic tensions are rescaled so that higher values always reflect higher institutional quality.

Table 5 provides some information about the performance of these excluded variables. The second column reports the *t*-statistics of the variables of the final model and each variable that was omitted from the final specification — that is, we present the *t*-statistics of the variables of the general specification. The last three columns give an indication of the extent to which the omitted

variables are collinear with the regressors of the final model, showing the pair-wise correlation coefficients and their *t*-statistics.

When all variables are included together in the regression, the statistical significance of the variables of the final model decreases (as expected), while several variables that were omitted from the final specification, such as openness (*Openness*), internal conflict (*Intconf*), and corruption (*Corr*), have the 'wrong' sign. This suggests that several of the omitted variables are correlated with the variables in the final model, in turn implying that some of the excluded variables might play an important indirect role in the aid-GDP relationship by affecting the included variables or being affected by them. The pair-wise correlation coefficients show that law and order is significantly correlated with many of the other variables: openness, the share of a country's area that is in the tropics (*Tropics*), political instability (*Instability*), socioeconomic conditions (*Socio*), investment profile (*Invprof*), internal conflict, corruption, military in politics (*Military*), ethnic tensions (*Ethnic*), and bureaucratic quality (*Bureau*). Government size is positively correlated with internal conflict and aid uncertainty (*Uncertain*), while religious tension has positive correlations with GDP per capita (*GDPpc*) (implying higher values are associated with lower tension) and military in politics (which is rescaled, so less military involvement is associated with less tensions).

Thus, the findings suggest that cross-country variations in the long-run effect of aid on GDP can be explained primarily by cross-country differences in religious tensions, law and order, and government size. However, this does not imply that all other variables are irrelevant for exploiting the potential of aid to increase domestic output. There are several factors — such as openness, political instability, and corruption — that are related to the direct determinants of the long-run effect of aid on GDP and thus are likely to play an important indirect role in the long-run relationship between aid and GDP.

## 6. Conclusion and Discussion

This paper first examined the nature of the relationship between foreign aid and economic development using panel cointegration techniques designed to deal with problems plaguing previous cross-country studies on aid effectiveness: omitted variables, cross-country heterogeneity, endogeneity, and neglected long-run level relationships between foreign aid and domestic output. Employing data for 59 developing countries over the period 1971 to 2003, aid has, on average, a negative long-run effect on output. This finding is robust to potential outliers, sample selection, the specification of the empirical model, the sample period, and different estimation techniques.

Nevertheless, there are large differences in the long-run effect of aid on output across countries. More specifically, an increase in the aid/GDP ratio is associated with a long-run decrease in GDP in almost two-thirds of the countries, while in just over a third of the cases an increase in the aid share is associated with a long-run increase in GDP. In general, irrespective of the sign, the effect is small. As investment was found to have a significant positive effect on long-run GDP, it is possible that this captures any long-run positive effect of aid on GDP through financing investment (Lensink and Morrissey, 2000; Gomanee et al., 2005b).

Using the estimated 'output effect' of aid for each country, a general-to-specific variable selection approach was then applied to identify important country-specific factors explaining the cross-country differences in the long-run effect of aid on output. In contrast to previous studies, the results suggest that aid effectiveness does not depend primarily on the quality of economic policy, location in the tropics, the level of democracy or corruption, political stability or absorptive capacity. Instead, the results suggest that the cross-country heterogeneity in the output effect of aid can be explained mainly, or most directly, by cross-country differences in law and order, religious tensions and government size. However, there are several factors — such as openness, location in the tropics, political instability, corruption, and bureaucratic accountability — that are highly correlated with law and order, religious tensions, and/or government size, suggesting these factors may play an indirect role in long-run effectiveness of aid.

Of the three primary determinants of aid effectiveness religious tension has the most significant coefficient; aid appears to be more effective in countries with lower religious tension. The most significant correlates of religious tension are GDP per capita (higher income is associated with lower tension) and the role of the military in politics (less involvement is associated with less tension); interestingly, religious tension does not appear to be correlated with ethnic tension. It is unlikely to be religious tensions themselves that undermine aid effectiveness, but a combination of the three variables: in low income countries that are militarized and have religious tension aid is unlikely to contribute to output. To some extent donors show awareness of this, as they monitor the use of aid to ensure it is not diverted to support military spending. This provides a warning in countries where aid is being used to support the war on terror; aid is unlikely to impact on output where the primary concern is security.

The finding that aid is less effective in countries with large consumption/GDP ratios, and that government size is positively correlated with internal conflict and aid uncertainty, supports arguments that governments with high levels of rent-seeking behavior are less likely to use aid effectively. Aid may induce self-interested individuals to engage in rent-seeking activities aimed at

appropriating part of any resource windfall, and, as a result, scarce resources are withdrawn from other productive activities, as several theoretical models predict (e.g. Hodler, 2007; Economides, 2008). One reason is that revenues, including aid, are more likely to be diverted to consumption spending rather than public investment, limiting any increase in output. This adverse effect may be heightened in the presence of internal conflict: distributive conflicts, of which war is the extreme, diminish the productivity of an economy in a number of ways, such as by generating uncertainty in the economic environment and/or by destroying the economic and political institutions and infrastructure that facilitate an efficient allocation of resources. Furthermore, the instability associated with distributive conflicts may make aid receipts more volatile and the associated aid uncertainty is likely to have a greater impact on reducing investment spending than on consumption (Lensink and Morrissey, 2000). In these cases donors need to carefully monitor the use of aid to ensure that productive activities are supported with reliable flows of income.

The analysis suggests that the most important variable capturing the quality of institutions is 'law and order', including the rule of law and protection of property rights. It is worth emphasizing that this variable is significantly correlated with most of the explanatory variables included (10 of the 17 other measures); it can be interpreted as capturing the effect of institutions on productivity and hence on aid effectiveness. There are numerous channels through which these effects can operate, and causality should not be inferred. The correlates suggest that countries that are more open with lower tensions, lower conflict and better indicators of institutional quality will have higher values of law and order, and such countries are less likely to be in the tropics. Thus, measures that improve institutions are likely to enhance aid effectiveness, and a wide range of potential interventions, from policy to capacity-building, are included.

A final conclusion is that the negative effect of aid found for many countries need not remain negative; it can become positive over time when certain country-specific factors determining the effect of aid change (i.e. are improved). Of these factors perhaps the most important is the quality of institutions. Institutions, especially enforcement of law and order (including property rights), can limit the appropriation of aid by rent seeking governments and thus ensure that aid is not wasted through profligate consumption but invested in productive activities. The results suggest a primacy of institutions over aid: a bad institutional environment not only depresses economic activity, as found in several studies, but also prevents aid from raising the standard of living. To improve aid effectiveness it is necessary to support institutional development, and there is evidence that this can be successful.

**References**


Acemoglu, D., Ventura, J., 2002. The world income distribution. *The Quarterly Journal of Economics* 117, 659–994.

Adam, C., O'Connell, S., 2004. Aid versus trade revisited: donor and recipient policies in the presence of learning-by-doing. *The Economic Journal* 114, 150–173.

Almeida, R., Fernandes, A., 2008. Openness and technological innovations in developing countries: evidence from firm-level surveys. *Journal of Development Studies* 44, 701–727.

Bräutigam, D.A., Knack, S., 2004. Foreign aid, institutions, and governance in sub-Saharan Africa. *Economic Development and Cultural Change* 52, 255–285.

Breitung, J., 2005. A parametric approach to the estimation of cointegrating vectors in panel data. *Econometric Reviews* 24, 151–173.

Burnside, C., Dollar, D., 2000. Aid, policies, and growth. *American Economic Review* 90, 847–868.

Burnside, C., Dollar, D., 2004. Aid, policies, and growth: Revisiting the evidence. *World Bank Policy Research Working No. 3251*.

Chauvet, L., Guillaumont, P., 2004. Aid and growth revisited: Policy, economic vulnerability, and political instability. In: Tungodden, B., Stern, N., Kolstad, I. (Eds.), *Toward pro-poor policies. Aid, institutions, and globalization*. Washington and New York: World Bank and Oxford University Press, 95–109.

Collier, P., Dollar, D., 2002. Aid allocation and poverty reduction. *European Economic Review* 46, 1475–1500.

Collier, P., Dollar, D., 2004. Development effectiveness: what have we learnt? *The Economic Journal* 114, F244–F271.

Commission for Africa 2005. *Our common interest: report of the Commission for Africa*. London: Commission for Africa.

Daalgard, C.-H., Hansen, H., 2001. On aid, growth, and good policies. *Journal of Development Studies* 37, 17–41.

Daalgard, C.-H., Hansen, H., Tarp, F., 2004. On the empirics of foreign aid and growth. *Economic Journal* 114, 191–216.

Easterly, W., Levine, R., Roodman, D., 2004. Aid policies and growth: comment, *American Economic Review* 94, 774–780.

Economides, G., Kalyvitis, S., Phillippopoulos, A., 2008. Does foreign aid distort incentives and hurt growth? Theory and evidence from 75 aid-recipient countries. *Public Choice* 134, 463–488.

Ericsson, N.R, Irons, J.S., Tryon. R.W., 2001. Output and inflation in the long-run. *Journal of Applied Econometrics* 16, 241–253.

Everhart, S.S. Sumlinski, M.A., 2001. Trends in private investment in developing countries, statistics for 1970 – 2000 and the impact on private investment of corruption and the quality of public investment. *International Finance Corporation Discussion Paper No. 44*, World Bank, Washington D.C.

Gallup, J.L., Sachs, J.D., Mellinger, A.D., 1999. Geography and economic development. *International Regional Science Review* 22, 179–232.

Gomanee, K., Girma, S., Morrissey, O., 2005a. Aid, Public Spending and Human Welfare: Evidence from Quantile Regressions. *Journal of International Development* 17, 299-309.

Gomanee, K., Girma, S., Morrissey, O., 2005b. Aid and growth in sub-Saharan Africa: accounting for transmission mechanisms. *Journal of International Development* 17, 1055–1076.

Hansen, B.E., 1992. Tests for parameter instability in regressions with I(1) processes. *Journal of Business and Economic Statistics* 10, 321–35.

Hansen, H., Tarp, F., 2000. Aid effectiveness disputed. *Journal of International Development* 12, 375–398.

Hansen, H., Tarp, F., 2001. Aid and growth regressions. *Journal of Development Economics* 64, 547–570.

Herzer, D., 2008. The long-run relationship between outward FDI and domestic output: Evidence from panel data. *Economics Letters* 100, 146–149.

Heston, A., Summers, R., Aten, B., 2006. *Penn World Table Version 6.2*. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

Hodler, R., 2007. Rent seeking and aid effectiveness. *International Tax and Public Finance* 14, 525–541.

Holly, S., Pesaran, M.H., Yamagata, T., 2009. A spatio-temporal model of house prices in the US. *Journal of Econometrics* (forthcoming).

Hoover, K., Perez, S.J., 2004. Truth and robustness in cross-country growth regressions. *Oxford Bulletin of Economics and Statistics* 66, 765–798.

Im, K.S., Pesaran, M.H., Shin, Y., 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115, 53–74.

Islam, M.N., 2005. Regime changes, economic policies and the effect of aid on economic growth. *Journal of Development Studies* 41, 1467–1492.

Johansen, S., 1988. Statistical analysis of cointegrating vectors. *Journal of Economic Dynamics and Control* 12, 231–254.

Kosack, S., 2003. Effective aid: How democracy allows development aid to improve the quality of live. *World Development* 31, 1–22.

Larsson, R., Lyhagen, J., Löthegren, M., 2001. Likelihood-based cointegration tests in heterogeneous panels. *Econometrics Journal* 4, 109–142.

Lensink, R., Morrissey, O., 2000. Aid instability as a measure of uncertainty and the positive impact of aid on growth. *Journal of Development Studies* 36, 31–49.

Levine, R.E., Renelt., D., 1992. A sensitivity analysis of cross-country growth regressions. *American Economic Review* 82, 942–963.

McGillvary, M., Morrissey, O., 2004. Fiscal effects of aid. In: Addison, T., Roe, A. (Eds.), *Fiscal policy for development: Poverty, Reconstruction, and Growth*. Palgrave Macmillan for UNU-Wider: Basingstoke, 72–96.

Maren, M., 1997. *The road to hell: The Ravaging effects of foreign aid and international charity*. New York: The Free Press.

Morrissey, O., 2006. Aid or trade, or aid and trade? *The Australian Economic Review* 39, 78–88.

Pedroni, P., 1999. Critical values for cointegration tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics* 61, 653–670.

Pedroni, P., 2000. Fully modified OLS for heterogeneous cointegrated panels. *Advances in Econometrics* 15, 93–130.

Pedroni, P., 2001. Purchasing power parity tests in cointegrated panels. *The Review of Economics and Statistics* 83, 727–731.

Pedroni, P., 2004. Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric Theory* 20, 597–625.

Pesaran, M.H. 2006. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica* 74, 967–1012.

Pesaran, M.H., 2007. A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics* 22, 265–312.

Pesaran, M.H., Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics* 68, 79–113.

Rajan, R.R., Subramanian, A., 2007. Does aid affect governance? *American Economic Review* 97, 322–327.

Rajan, R.R., Subramanian, A., 2008. Aid and growth: what does the cross-country evidence really show? *The Review of Economics and Statistics* 90, 643–665.

Reinsel, G.C., Ahn, S.K., 1992. Vector autoregressive models with unit roots and reduced rank structure: estimation, likelihood ratio test and forecasting. *Journal of Time Series Analysis* 13, 353–375.

Sachs, J.D., Warner, A., 1995. Economic reform and the process of global integration. *Brookings Papers on Economic Activity* 1, 1–118.

Sala-I-Martin, X.X., 1997. I just ran two million regressions. *American Economic Review* 87, 178–183.

Saxonhouse, G., 1976. Estimated parameters as dependent variables. *American Economic Review* 66, 178–183.

Singh, R.D., 1985. State intervention, foreign economic aid, savings and Growth in LDCs: some recent evidence. *Kyklos* 38, 216–232.

Svensson, J., 1999. Aid, growth and democracy. *Economics and Politics* 11, 275–297.

Stock, J.H., Watson, M.W., 1993. A simple estimator of cointegrating vectors in higher-order integrated systems. *Econometrica* 61, 783–820.

Temple, J., 1999. The new growth evidence. *Journal of Economic Literature* 37, 112–156.

Wacziarg, R., Welch, K.H., 2008. Trade liberalization and growth: New evidence. *World Bank Economic Review* 22, 187–231.

**Appendix A1 Panel unit-root tests**

To examine the time series properties of the data, we use the panel unit root test of Im, Pesaran, and Shin (2003) (IPS), which is based on the augmented Dickey-Fuller (ADF) regression for the *i*th cross-section unit:

$$\Delta x_{it} = z_{it}'\gamma + \rho_i x_{it-1} + \sum_{j=1}^{p_i} \varphi_{ij} \Delta x_{it-j} + \varepsilon_{it}, \tag{A.1}$$

where $p_i$ is the lag order and $z_{it}$ represents deterministic terms, such as fixed effects or fixed effects combined with individual time trends. The IPS test tests the null hypothesis of a unit root for all $i$, $H_0 : \rho_i = 0$, against the alternative of (trend) stationary, $H_1 : \rho_i < 0$, $i = 1, 2, \ldots, N_1$; $\rho_i = 0$, $i = N_1 + 1, N_1 + 2, \ldots, N$, using the standardized $t$-bar or IPS statistic

$$\Gamma_{\bar{t}} = \frac{\sqrt{N}\left[\bar{t}_{NT} - \mu\right]}{\sqrt{v}}, \tag{A.2}$$

where $\bar{t}_{NT}$ is the average of the $N$ (= 59) cross-section ADF $t$-statistics, and $\mu$ and $v$ are, respectively, the mean and variance of the average of the individual $t$-statistics, tabulated by Im, Pesaran, and Shin (2003).

However, the standard IPS test can lead to spurious inferences if the errors, $\varepsilon_{it}$, are not independent across $i$ — for example, due to common shocks or spillovers between countries. Therefore, we also employ the cross-section augmented IPS test proposed by Pesaran (2007), which is designed to filter out the cross-section dependency by augmenting the ADF regression with the cross-section averages of lagged levels and first-differences of the individual series. Accordingly, the cross-section augmented ADF (CADF) regression is given by

$$\Delta x_{it} = z_{it}'\gamma + \rho_i x_{it-1} + \sum_{j=1}^{p_i} \varphi_{ij} \Delta x_{it-j} + \alpha_i \bar{x}_{t-1} + \sum_{j=0}^{p_i} \eta_{ij} \Delta \bar{x}_{t-j} + v_{it}, \tag{A.3}$$

where $\bar{x}_t$ is the cross-section mean of $x_{it}$, $\bar{x}_t = N^{-1} \sum_{i=1}^{N} x_{it}$. The cross-section augmented IPS statistic is the simple average of the individual CADF statistics and is defined as

$$CIPS = t\text{-bar} = N^{-1} \sum_{i=1}^{N_i} t_i, \tag{A.4}$$

where $t_i$ is the OLS $t$-ratio of $\rho_i$ in Equation (A.3). Critical values are tabulated by Pesaran (2007).

The test results for the variables in levels and in first differences are presented in Table A1 Both the IPS and the CIPS test statistics are unable to reject the hypothesis that all countries have a unit root in levels. Since the unit root hypothesis can be clearly rejected for the first differences, it can be concluded that $Y_{it}$, $I_{it}$, and $A_{it}$ are integrated of order one, I(1).

**Table A1 Panel unit root tests**

| Variables | Deterministic terms | IPS statistics | CIPS statistics |
|---|---|---|---|
| Levels | | | |
| $Y_{it}$ | $c$, $t$ | -1.36 | -2.01 |
| $I_{it}$ | $c$, $t$ | -0.76 | -2.49 |
| $A_{it}$ | $c$, $t$ | 0.51 | -1.94 |
| First differences | | | |
| $\Delta Y_{it}$ | $c$ | -9.56** | -2.47** |
| $\Delta I_{it}$ | $c$ | -12.79** | -3.02** |
| $\Delta A_{it}$ | $c$ | -12.79** | -2.74** |

*Notes: c* (*t*) indicates that we allow for different intercepts (and time trends) for each country. Three lags were selected to adjust for autocorrelation. The IPS statistic is distributed as $N(0, 1)$. The relevant 1% (5%) critical value for the CIPS statistics is -2.73 (-2.61) with an intercept and a linear trend, and -2.23 (-2.11) with an intercept.** denote significance at the 1% level.

To select a panel of countries for which $A_{it}$, $I_{it}$, and $Y_{it}$ behave as random walks, those countries for which the individual time series do not pass a screening for a unit root via the ADF and the KPSS tests were eliminated. This sample selection procedure yields a sample of 59 countries ($N = 59$) from 66 countries for which data on all variables are available from 1971 to 2003 ($T = 33$). The excluded countries are Argentina, Brazil, Hong Kong, Mexico, Trinidad and Tobago, Turkey, and Venezuela, since the unit root-tests suggest that aid is stationary and the log of GDP of these countries behaves like a random walk.

**Appendix A2 Cointegration tests**

We first test for cointegration using the Larsson et al. (2001) approach, which is based on Johansen's (1988) maximum likelihood estimation procedure. Like the Johansen time series cointegration test, the Larsson et al. panel test treats all variables as potentially endogenous, thus avoiding the normalization problems inherent in residual-based cointegration tests. Moreover, in contrast to residual-based cointgration tests, the Larsson et al. procedure allows the determination of the number of cointegrating vectors. It involves estimating the Johansen vector error-correction model for each country separately and then computing the individual trace statistics

$LR_{iT}\{H(r)|H(p)\}$. The null hypothesis is that all of the *N* countries in the panel have a common cointegrating rank, i.e. at most *r* (possibly heterogeneous) cointegrating relationships among the *p* (= 3) variables:

$$H_0 : rank(\Pi_i) = r_i \le r \qquad \text{for all } i = 1, \ldots, N, \tag{A.5}$$

whereas the alternative hypothesis is that all the cross-sections have a higher rank:

$$H_1 : rank(\Pi_i) = p \qquad \text{for all } i = 1, \ldots, N, \tag{A.6}$$

where $\Pi_i$ is the long-run matrix of order $p \times p$. To test $H_0$ against $H_1$, a panel cointegration rank trace test statistic is computed by calculating the average of the individual trace statistics:

$$\overline{LR}_{NT}\{H(r)|H(p)\} = \frac{1}{N}\sum_{i=1}^{N} LR_{iT}\{H(r)|H(p)\}, \tag{A.7}$$

and then standardizing it as follows:

$$\Psi_{\overline{LR}}\{H(r)|H(p)\} = \frac{\sqrt{N}\left(\overline{LR}_{NT}\{H(r)|H(p)\} - E(Z_k)\right)}{\sqrt{Var(Z_k)}} \Rightarrow N(0,1), \tag{A.8}$$

where the mean $E(Z_k)$ and variance $Var(Z_k)$ of the asymptotic trace statistic are tabulated by Breitung (2005) for the model we use — the model with a constant and a trend in the cointegrating relationship.

However, the Johansen trace statistics are biased toward rejecting the null hypothesis in small samples. To avoid the Larsson et al. test, as a consequence of this bias, also overestimating the cointegrating rank, we compute the standardized panel trace statistics based on small-sample corrected country-specific trace statistics. Specifically, we use the small-sample correction factor suggested by Reinsel and Ahn (1992) to adjust the individual trace statistics as follows:

$$LR_{iT}\{H(r)|H(p)\} \times \left[\frac{T - k_i \times p}{T}\right], \tag{A.9}$$

where $k_i$ is the lag length of the models used in the test.

A potential problem with Larsson et al. approach, however, is that it does not take into account potential error cross-sectional dependence, which could bias the results. To test for cointegration in the presence of possible cross-sectional dependence we follow Holly et al. (2009) and adopt a residual-based two-step approach in the style of Pedroni (1999). Unlike Pedroni, we use the common correlated effects (CCE) estimation procedure developed by Pesaran (2006) in the

first-step regression. This procedure allows for cross-section dependencies that potentially arise from multiple unobserved common factors by augmenting the cointegrating regression with the cross-section averages of the dependent variable and the observed regressors as proxies for the unobserved factors. Accordingly, the cross-section augmented cointegrating regression we estimate for each country is given by

$$Y_{it} = a_i + \delta_i t + \beta_{1i} A_{it} + \beta_{2i} I_{it} + g_{1i} \overline{Y_t} + g_{2i} \overline{A_t} + g_{3i} \overline{I_t} + e_{it}, \tag{A.10}$$

where $\overline{Y_t}$, $\overline{A_t}$, and $\overline{I_t}$ are the cross-section averages of $Y_{it}$, $A_{it}$, and $I_{it}$ in year *t*. In the second step, we compute the cross-section augmented IPS statistic for the residuals from the individual CCE long-run relations, $\hat{\mu}_{it} = Y_{it} - \hat{\delta}_i t - \hat{\beta}_{1i} A_{it} - \hat{\beta}_{2i} I_{it}$, including an intercept. Thus, we account for unobserved common factors that could be correlated with the observed regressors in both steps.

The results of these tests are presented in Table A2. For completeness, we also report the standard panel and group ADF and PP test statistics suggested by Pedroni (1999, 2004). As can be seen, all tests strongly suggest that $Y_{it}$, $A_{it}$, and $I_{it}$ are cointegrated. The standardized trace statistics clearly support the presence of one cointegrating vector. Also, the CIPS, the ADF, and the PP statistics reject the null hypothesis of no cointegration at least at the 5% level, implying that there exists a long-run relationship between output, aid, and investment.

**Table A2 Panel cointegration tests**

| | Cointegration rank | | |
|---|---|---|---|
| | $r = 0$ | $r = 1$ | $r = 2$ |
| Panel trace statistics | 9.43** | 0.22 | -2.52 |
| CIPS statistic | | -3.00** | |
| Panel ADF statistic | | -3.78** | |
| Group ADF statistic | | -3.05** | |
| Panel PP statistic | | -3.18** | |
| Group PP statistic | | -2.07* | |

*Notes:* ** (*) indicate a rejection of the null hypothesis of no cointegration at the 1% (5%) level. The panel trace statistics, the ADF statistics, and the PP statistics are asymptotically normally distributed. The relevant 5% (1%) critical value for the CIPS statistic is -2.11 (-2.23). The number of lags was determined by the Schwarz criterion with a maximum number of three lags. The panel statistics pool the autoregressive coefficients across different countries during the unit root test on the residuals of the static cointegrating regressions, whereas the group statistics are based on averaging the individually estimated autoregressive coefficients for each country. The panel ADF statistic is analogous to the Levin et al. (2002) panel unit root test. The group ADF statistic is analogous to the IPS panel unit root test. The PP statistics are panel versions of the Phillips-Perron (PP) *t*-statistics.

## Appendix A3 Sensitivity results for long-run estimates

To verify that the negative effect of aid on output is not due to individual outliers the DOLS regression (with demeaned data to account for the likely cross-sectional dependence through common time effects)[9] is re-estimated excluding one country at a time from the sample. The sequentially estimated group-mean coefficients and their *t*-statistics are presented in Figure A4. As they are relatively stable between -0.0095 and -0.007 and always significant at the one percent level, we conclude that the average negative effect is not the result of individual outliers.

**Table A3 DOLS estimation with regional country groups excluded from the sample**

| | $A_{it}$ | $I_{it}$ | Number of countries in the sample |
|---|---|---|---|
| Excluding North Africa | -0.0064** (-8.07) | 0.0162** (14.78) | 55 |
| Excluding sub-Saharan Africa | -0.0092** (-7.65) | 0.0218** (14.77) | 29 |
| Excluding South America | -0.0084** (-9.91) | 0.0154** (14.12) | 54 |
| Excluding Central America and the Caribbean | -0.0068** (-7.62) | 0.0148** (13.25) | 50 |
| Excluding East Asia | -0.0096** (-10.26) | 0.0189** (14.75) | 51 |
| Excluding South Asia and the Middle East | -0.0080** (-8.56) | 0.0164** (14.29) | 55 |

** indicate significance at the 1% level. *t*-statistics in parentheses. The countries included in each region are: North Africa: Algeria, Egypt, Tunisia, Morocco; sub-Saharan Africa: Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Congo, Dem. Rep., Congo, Rep., Cote d'Ivoire, Gabon, Gambia, Ghana, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Niger, Nigeria, Rwanda, Rwanda, Sierra Leone, Sudan, Swaziland, Togo, Zambia, Zimbabwe; South America: Bolivia, Chile, Ecuador, Peru, Uruguay; Central America and the Caribbean: Belize, Costa Rica, Dominican Republic, El Salvador, Guatemala, Honduras, Jamaica, Nicaragua, Panama; East Asia: Indonesia, Malaysia, Papua New Guinea, Philippines, Solomon Islands, South Korea, Thailand; South Asia and the Middle East: India, Pakistan, Sri Lanka, Syria

The negative long-run relationship between aid and GDP may be due to sample-selection bias if a group of countries in a particular region have a significant effect on the results. To investigate this equation (4) is re-estimated excluding (in turn) countries from North Africa, sub-Saharan Africa, South America, Central America and the Caribbean, East Asia, and South Asia and the Middle East. The resulting group-mean values for $\beta_1$ (and $\beta_2$) are reported in Table A3. Regardless of which of these regions is excluded from the sample, the long-run relationship between aid and GDP remains negative and highly significant. From this, it can be concluded that the estimates do not suffer from sample selection-bias. Another conclusion that can be drawn from the results is that the effect of aid on GDP is not lower in sub-Saharan Africa compared to other regions; estimating the output effect for sub-Saharan Africa separately, the DOLS group-mean value is -0.0070 (with a *t*-value of -5.63), which is very close to the result for the total sample (-0.0081).

A potential specification problem is that the estimates may be biased if the investment variable includes public investment that is aid-financed. In this case, the estimated $\beta_1$ coefficient would not capture the share of aid that is used for public investment purposes and thus underestimates the true effect of aid on GDP. To investigate this the DOLS regression is estimated with private investment data in place of total investment. The data on private investment (as a share of GDP) are from the International Finance Corporation's *Trends in private investment in developing countries*, authored by Everhart and Sumlinski (2001). Unfortunately, these data are available only for 18 of the 59 countries of our sample over a sufficiently long period of time (25 years from 1975 to 1999).[10] Nevertheless, the long-run effect of aid on GDP remains negative and statistically significant at the 1% level, as the results in Table A4 show.

**Table A4 DOLS estimates of the coefficients on aid and private investment**

| $A_{it}$ | $Iprivate_{it}$ |
|---|---|
| -0.0138** (-3.71) | 0.0230** (9.64) |

*Notes:* The dependent variable is $Y_{it}$. ** indicate significance at the 1% level. *t*-statistics in parentheses. The individual DOLS regressions were estimated with one lead and one lag. The sample period is 1975 -1999 and the number of countries in the sample is 18.

To check whether the results are sensitive to the sample period the DOLS regression is re-estimated for the sub-periods from 1971 through 1995 and 1979 through 2003. The results are presented in Table A5. Once again, the coefficient on $A_{it}$ is always negative and statistically significant.

**Table A5 DOLS estimation for different sub-periods**

| | $A_{it}$ | $I_{it}$ |
|---|---|---|
| 1971-1995 | -0.0072** (-6.56) | 0.0191** (16.42) |
| 1979-2003 | -0.0088** (-9.57) | 0.0168** (12.96) |

The dependent variable is $Y_{it}$. ** indicate significance at the 1% level. *t*-statistics in parentheses. The individual DOLS regressions were estimated with one lead and one lag.

[9] In the following, we use the demeaned data to account for the likely cross-sectional dependence through common time effects.

[10] We include all countries of our sample for which continuous private investment data are available from 1975 to 1999. These countries are: Belize, Chile, Dominican Republic, Ecuador, El Salvador, Guatemala, India, Kenya, Malawi, Malaysia, Morocco, Pakistan, Paraguay, Philippines, South Korea, Thailand, Tunisia, and Uruguay. The data have been demeaned with respect to common time effects.

## Appendix A4 Data and sample for determinants of aid impact

*Data*

Twenty variables are considered potentially relevant to aid effectiveness; their definitions and sources are listed in Table A6 below. The quality of economic policy in Burnside and Dollar (2000) is measured by a weighted average of inflation, the budget balance as a share of GDP, and the Sachs and Warner openness index. As this aggregate policy index may, however, mask potential differences in the influence of monetary policy, fiscal policy, and trade policy on aid effectiveness, the individual components of the policy index are included separately in the analysis. The Sachs and Warner openness index is constructed on the basis of the liberalization dates provided by Sachs and Warner (1995) and Wacziarg and Welch (2008). Data on inflation and the budget balance are from the World Development Indicators.

The fourth variable is the share of a country's area that is in the tropics. Dalgaard et al. (2004), for example, find that the effect of aid on economic growth is negatively associated with the fraction of a country that is located in the tropics, and offer two possible explanations. First, tropical countries tend to grow slower than countries with a more temperate climate because of lower agricultural productivity and the high prevalence of diseases (such as AIDS, malaria, and tuberculosis). For the same reasons, the effect of aid on growth tends to be stronger in countries with a temperate climate than in tropical countries, implying that aid effectiveness depends directly on the climate. Second, climatic circumstances may have influenced the evolution of other slow-changing characteristics, like institutions. Thus, the fraction of tropical area may be seen as a rough indicator for institutional quality in the broadest sense, and aid effectiveness depends indirectly on institutional quality. The variable used is the percentage of land area in the tropics from Gallup et al. (1999) (available at http://www.cid.harvard.edu/ciddata/ciddata).

The long-run effect of aid on GDP may depend on the level of democratization. Svensson (1999) argues that democratic institutions (such as political parties, elected representatives, free speech) provide a recurrent and institutionalized check on governments, forcing them to use aid for productive purposes rather than for unproductive government consumption, and finds that the effect of aid on growth is stronger for countries with higher democracy scores. Similarly, the empirical results in Kosack (2003) suggest that the effect of aid on the quality of life (measured by the Human Development Index) is positively related to the level of democracy; democracy is measured by the democracy index from the POLITY IV data base, with values from 0 (absence of democracy) to 10 (highest level of democracy) (http://www.systemicpeace.org/polity/polity4.htm).

A measure of political instability is used to capture the idea that in a troubled environment, with violent conflicts and frequently changing governments, aid is less effective. Chauvet and Guillaumont (2004) and Islam (2005) find that the growth effect of aid is negatively associated with political instability. The variable used is the number of revolutions and coups, calculated from the Center for Systemic Peace (http://www.systemicpeace.org/inscr/inscr.htm) on the basis of Coups d'état data.

**Table A6 Variables and sources for determinants of aid effect**

| Variables | Definition | Source |
|---|---|---|
| *Inflation* | Percentage change in the consumer prices. Data averaged over the period 1971 to 2003. | World Development Indicators 2009 |
| *Budget* | Overall budget balance as a percentage of GDP. Data averaged over the period 1971 to 2002. | World Development Indicators 2004 |
| *Openness* | Sachs and Warner openness index. Data averaged over the period 1971 to 2003. | Sachs and Warner (1995); Wacziarg and Welch (2008) |
| *Tropics* | The share of a country's area that is in the tropics. | Gallup et al. (1999); http://www.cid.harvard.edu/ciddata/ciddata |
| *Demo* | Democracy. Data averaged over the period 1971 to 2003. | POLITY IV; http://www.systemicpeace.org/polity/polity4.htm |
| *Instability* | Political instability. Measured by the number of revolutions and coups. Data averaged over the period 1971 to 2003. | Center for Systemic Peace; http://www.systemicpeace.org/inscr/inscr.htm |
| *Socio* | Socioeconomic conditions. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Invprof* | Investment profile. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Intconf* | Internal conflict. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Extconf* | External conflict. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Corr* | Corruption. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Military* | Military in politics. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Rel* | Religious tensions. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Law* | Law and order. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Ethnic* | Ethnic tensions. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Bureau* | Bureaucratic quality. Data averaged over the period 1984 to 2003. | Political Risk Services Group |
| *Uncertain* | Aid uncertainty. Measured by the standard deviation of the residuals of a regression of aid (as a percentage of GDP) on two lags of aid, a constant term, and a linear time trend. Estimated over the period 1971 to 2003. | World Development Indicators 2009 |
| *GDPpc* | Real per capita GDP in constant 2000 US dollars. Data averaged over the period 1971 to 2003. | World Development Indicators 2009 |
| *Size* | Government size. Measured by the share of general government consumption in GDP. Data averaged over the period 1971 to 2003. | World Development Indicators 2009 |
| *Capacity* | Absorptive capacity. Measured by the secondary school enrolment rate. Data averaged over the period 1991 to 2005. | World Development Indicators 2009 |
| Dependent variable: $\hat{\beta}_i$ | Long-run effect of aid on GDP. Individual DOLS estimates of the coefficient on $A_{it}$ over the period 1971 to 2003. | Table 6 |

The long-run output effect of aid will depend on the quality of institutions (in ways that may not be captured by the measures already discussed). Institutions lower transaction costs by reducing uncertainty and establishing a stable economic structure to facilitate interactions, thus helping to allocate resources to their most efficient uses. Without institutions, both individuals and governments do not have incentives to invest in physical or human capital or adopt more efficient technologies, implying that resources are misallocated and opportunities for efficient use of aid go unexploited. Among the few studies addressing this issue, Burnside and Dollar (2004) find evidence that the growth effect of aid and institutional quality are positively related, whereas the results in Collier and Dollar (2002) surprisingly suggest that the relationship between the growth effect of aid and institutional quality is negative (although only significant at the 10% level). Both employ a single composite measure of institutional quality. It may be useful to consider several aspects of institutional quality to identify those institutional factors that are most important for aid effectiveness. Several institutional variables are available from the International Country Risk Guide (ICRG), published by the Political Risk Services Group.[11] They are defined as follows:

- Socioeconomic conditions — this index quantifies socioeconomic pressures at work in society that could constrain government action or fuel social dissatisfaction and thus de-stabilize the political regime.
- Investment profile — this measure assesses the factors affecting the risk to investment that are not covered by other political, economic or financial risk components, such as contract viability or payment delays.
- Internal conflict — the internal conflict measure is an assessment of political violence within a country (such as civil war, terrorism, or civil disorder) and its actual or potential impact on governance.
- External conflict — the external conflict measure assesses the risk to the incumbent government from foreign action, ranging from non-violent external pressure to violent external pressure.
- Corruption — this index assesses the level of corruption within the political system.
- Military in politics — this measure assesses the influence of the military in politics.
- Religious tensions — this is a measure of the extent to which society and/or governance is dominated by a single religious group that seeks to replace civil law by religious law and to exclude other religions from the political and social process.

[11] See https://www.prsgroup.com/prsgroup_shoppingcart/pc-75-7-icrg-historical-data.aspx

- Law and order — this index assesses the strength and impartiality of the legal system and popular observance of the law.
- Ethnic tensions — this measure is an assessment of the degree of tensions within a country attributable to racial, nationality, or language divisions.
- Bureaucratic quality — this is an assessment of the institutional strength and quality of the bureaucracy in terms of acting as a shock absorber to minimize revisions of policy when governments change.

It is important to note that the indicators for corruption, internal conflict, external conflict, military in politics, religious tensions, and ethnic tensions are rescaled so that higher values always reflect higher institutional quality.[12]

Aid uncertainty is a possible determinant of the long-run effect of aid on GDP. If aid receipts are observed to vary significantly from year to year (aid uncertainty), whether due to macroeconomic uncertainty (aid responding to shocks), disbursement (donor) or absorption (recipient) difficulties, this can undermine budget and economic planning and reduce the quantity, as well as efficiency, of domestic investment. Lensink and Morrissey (2000) find that aid uncertainty has a negative effect on economic growth, although aid itself has a positive effect, which could suggest that the effectiveness of aid is reduced by increased aid uncertainty. Following Lensink and Morrissey (2000), aid uncertainty is measured by the standard deviation of the residuals of a regression of aid (as a percentage of GDP) on two lags of aid, a constant term, and a linear time trend. Data on aid come from the World Development Indicators (as noted in Section 3).

GDP per capita might be important for explaining cross-country differences in the output effect of aid to the extent that aid tends to decline as GDP rises, implying that the marginal effectiveness of aid is highest in low-income countries. Gomanee et al. (2005a) find that the effect of aid on welfare (measured by the Human Development Index and the infant mortality rate) is higher for low-income countries. Data on real per capita GDP are taken from World Development Indicators.

Government size may be a factor in explaining the cross-country differences in the output effect of aid. Economides et al. (2008) develop a growth model in which aid promotes rent-seeking behavior at the expense of productive government activities, and the rent-seeking effect — and thus the growth effect of aid — depends on both the amount of aid and the size of the recipients

country's public sector. The rationale for the latter is that rent seeking triggered by aid transfers takes place via state coffers, and this process is facilitated when the size of the public sector is large. Moreover, large government size may be associated with more regulation and government intervention, and this may also reduce the effectiveness of aid, as suggested by Singh (1985). Empirical support for this hypothesis is provided by Economides et al. (2008), who find that the growth effect of aid is larger for countries with small public sectors. Following common practice, government size is measured by the share of government consumption in GDP (from World Development Indicators).

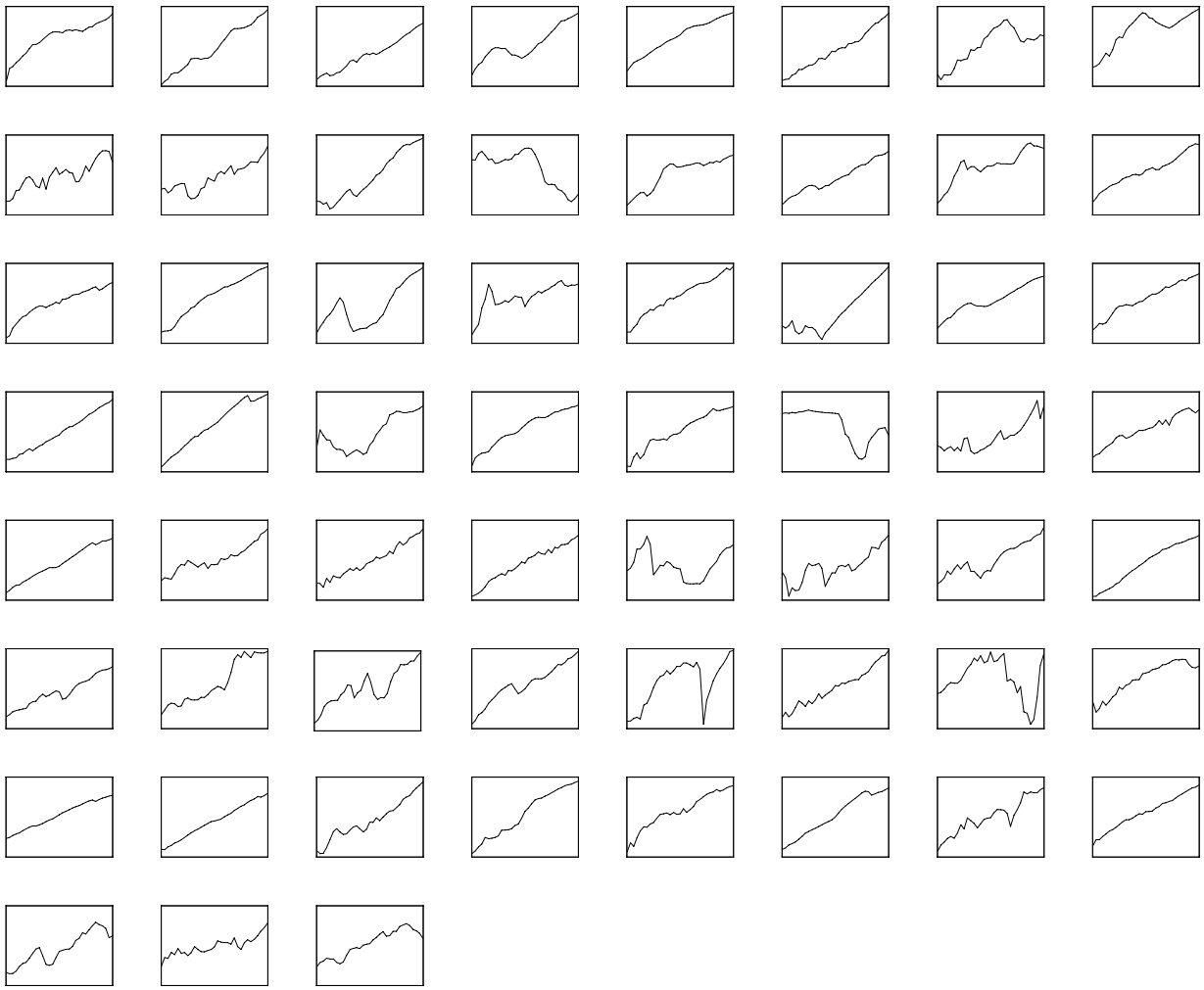
Finally, the long-run effect of aid on GDP is likely to depend on the 'absorptive capacity' of a country (Chauvet and Guillaumont, 2004). Absorptive capacity is here measured by the secondary school enrolment rate from World Development Indicators.

*Sample of countries used in the analysis of the determinants of aid effectiveness*

Algeria, Botswana, Burkina Faso, Cameroon, Chile, Congo, Rep., Cote d'Ivoire, Dominican Republic, Ecuador, Egypt, El Salvador, Gabon, Gambia, Ghana, Guatemala, Honduras, India, Indonesia, Jamaica, Kenya, Liberia, Madagascar, Malawi, Malaysia, Mali, Morocco, Nicaragua, Niger, Nigeria, Pakistan, Panama, Papua New Guinea, Peru, Philippines, Senegal, Sierra Leone, South Korea, Sri Lanka, Syria, Thailand, Togo, Tunisia, Uruguay, Zambia, Zimbabwe

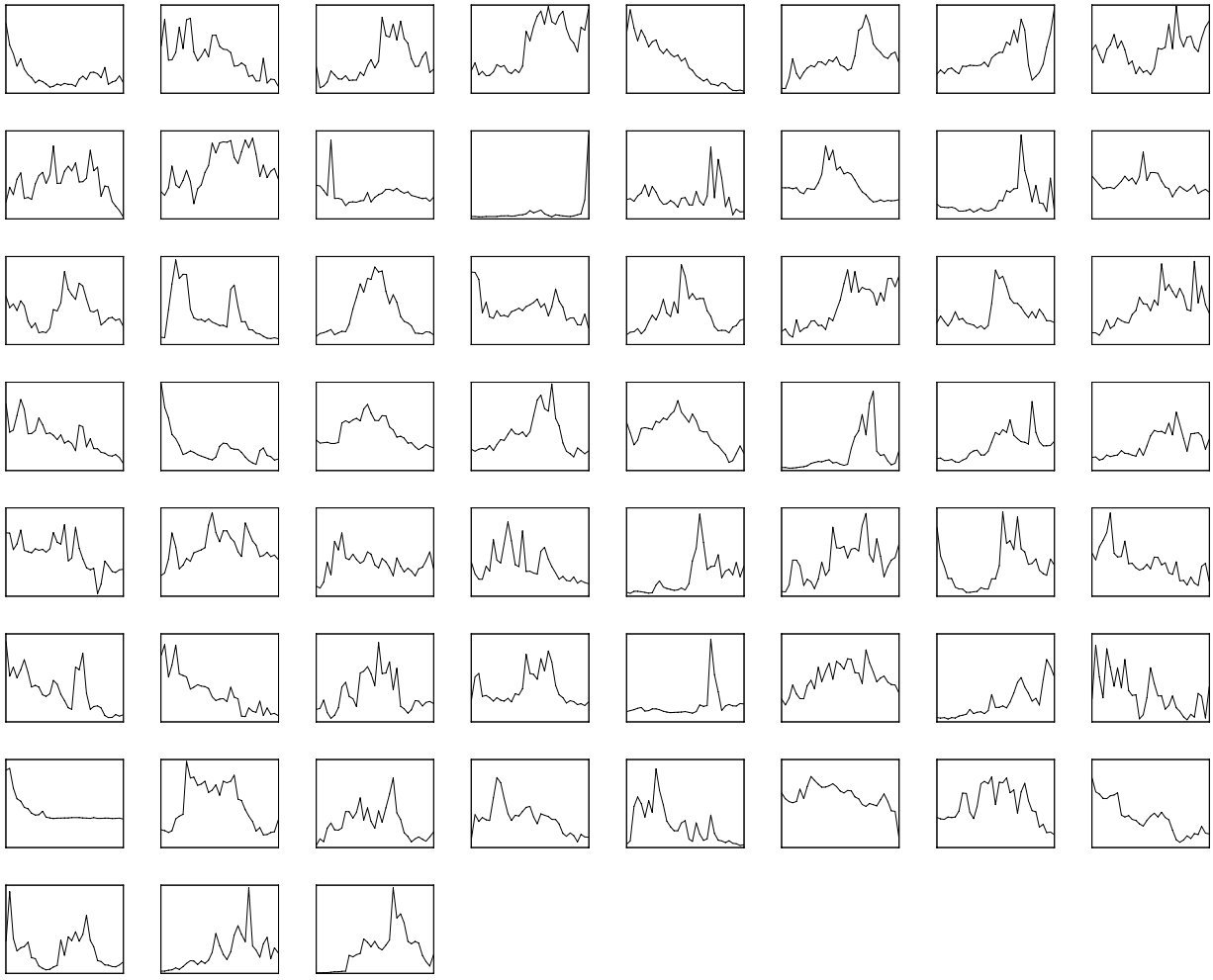
[12] The democratic accountability index from ICRG has been excluded, since we have chosen to use the democracy index from Polity IV, as in Kosack (2003). Moreover, we do not include the government stability index from ICRG, since we measure political and thus government instability by the number of the number of revolutions and coups.

**Figure A1 Log of GDP by country over the period 1971-2003, $Y_{it}$**



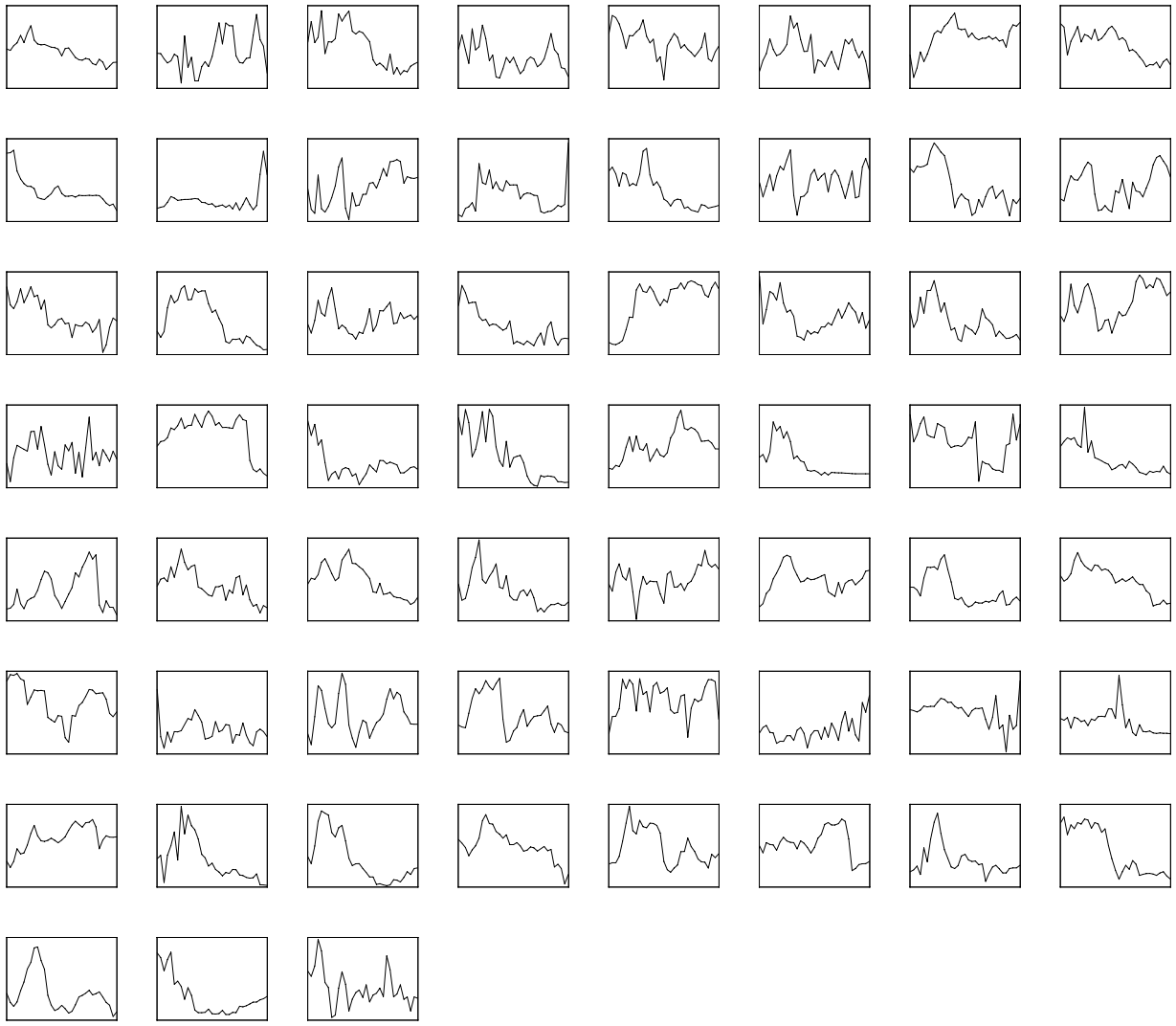
*Notes:* The countries from the left to the right are: Algeria, Belize, Benin, Bolivia, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Chile, Democratic Republic of Congo, Republic of Congo, Costa Rica, Cote d'Ivoire, Dominican Republic, Ecuador, Egypt, El Salvador, Gabon, Gambia, Ghana, Guatemala, Honduras, India, Indonesia, Jamaica, Kenya, Lesotho, Liberia, Madagascar, Malawi, Malaysia, Mali, Mauritania, Morocco, Nicaragua, Niger, Nigeria, Pakistan, Panama, Papua New Guinea, Peru, Philippines, Rwanda, Senegal, Sierra Leone, Solomon Islands, South Korea, Sri Lanka, Sudan, Swaziland, Syria, Thailand, Togo, Tunisia, Uruguay, Zambia, Zimbabwe

**Figure A2 Aid/GDP ratio by country over the period 1971-2003, $A_{it}$**



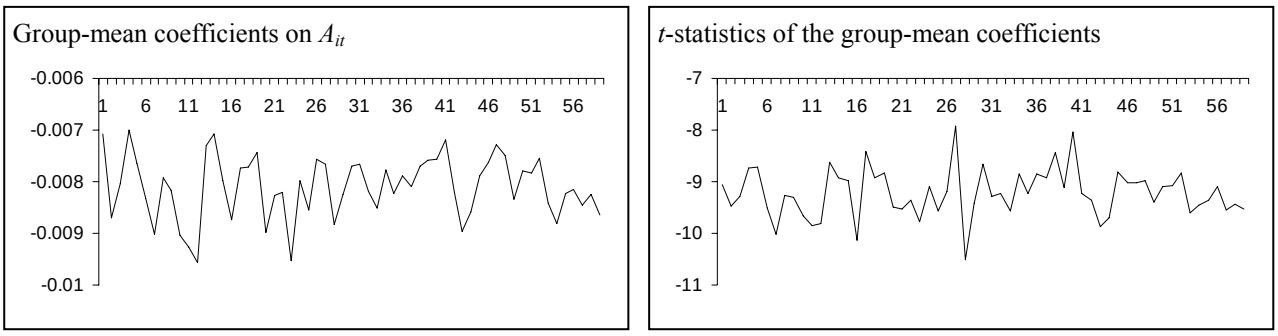
*Notes:* The countries from the left to the right are: Algeria, Belize, Benin, Bolivia, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Chile, Democratic Republic of Congo, Republic of Congo, Costa Rica, Cote d'Ivoire, Dominican Republic, Ecuador, Egypt, El Salvador, Gabon, Gambia, Ghana, Guatemala, Honduras, India, Indonesia, Jamaica, Kenya, Lesotho, Liberia, Madagascar, Malawi, Malaysia, Mali, Mauritania, Morocco, Nicaragua, Niger, Nigeria, Pakistan, Panama, Papua New Guinea, Peru, Philippines, Rwanda, Senegal, Sierra Leone, Solomon Islands, South Korea, Sri Lanka, Sudan, Swaziland, Syria, Thailand, Togo, Tunisia, Uruguay, Zambia, Zimbabwe

**Figure A3 Investment/GDP ratio by country over the period 1971-2003, $I_{it}$**



*Notes:* The countries from the left to the right are: Algeria, Belize, Benin, Bolivia, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Chile, Democratic Republic of Congo, Republic of Congo, Costa Rica, Cote d'Ivoire, Dominican Republic, Ecuador, Egypt, El Salvador, Gabon, Gambia, Ghana, Guatemala, Honduras, India, Indonesia, Jamaica, Kenya, Lesotho, Liberia, Madagascar, Malawi, Malaysia, Mali, Mauritania, Morocco, Nicaragua, Niger, Nigeria, Pakistan, Panama, Papua New Guinea, Peru, Philippines, Rwanda, Senegal, Sierra Leone, Solomon Islands, South Korea, Sri Lanka, Sudan, Swaziland, Syria, Thailand, Togo, Tunisia, Uruguay, Zambia, Zimbabwe

**Figure A4 DOLS estimation with single country excluded from the sample**



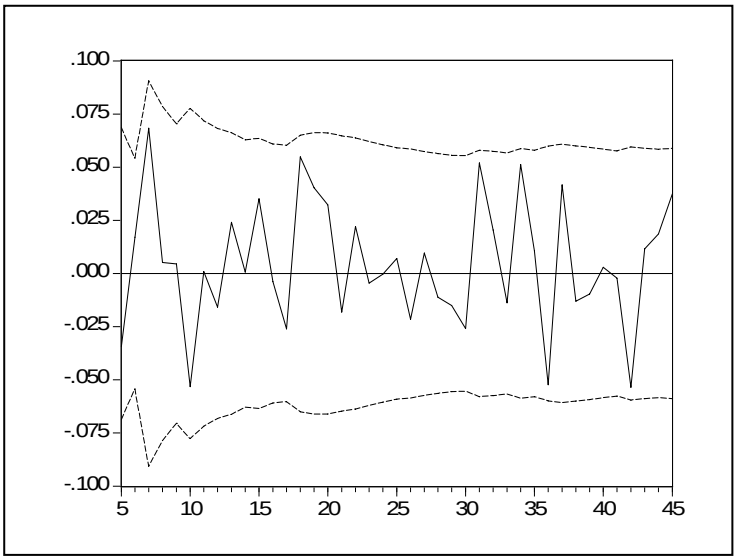


**Figure A5 Recursive residuals**