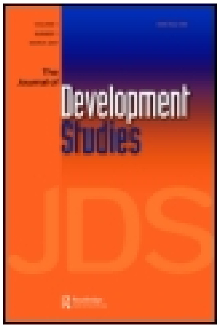


Journal of Development Studies
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/fjds20

On Aid, Growth and Good Policies

C-J. Dalgaard & H. Hansen

Version of record first published: 29 Mar 2010

On Aid, Growth and Good Policies

CARL-JOHAN DALGAARD
and HENRIK HANSEN

I. INTRODUCTION

The World Bank policy research report, *Assessing Aid* [*World Bank, 1998*], provides a careful, and rather self-critical, evaluation of the Bank's recent experience with foreign aid. A large part of the report can be read as advocating policy dialogue, beneficiary involvement, and local ownership instead of policy conditionality and enforced additionality of aid financed projects. There is also a clear recognition of the need for 'conditionality' in the design and choice of aid instruments in the sense that the type of aid to a given country should be conditional on the stage of development. In the analysis of the importance of the stage of development the Bank has chosen to concentrate, almost exclusively, on government institutions and macroeconomic policy though it is noted that other factors such as civil liberties are also important for the impact of foreign aid.

Nevertheless, in the discussions following the report most of the attention has focused on the first chapter in which the Bank seems to opt for policy based selectivity in future aid allocations. Specifically, the Overview states: 'Financial aid works in a good policy environment [and therefore] financial assistance must be targeted more effectively to low-income countries with sound economic management' [*World Bank*, *1998: 2, 4*]. The unambiguous policy message has provoked quite a few development economists and resulted in a new wave of studies of the link between aid and growth. Not surprisingly, many of the new studies are critical to the policy selectivity results and some even question the robustness of the empirical support underlying the recommendations in the report.[1]

The primary background paper to Chapter 1 in the report is Burnside and Dollar [*1997*], which appears in a slightly modified form in the American Economic Review [*Burnside and Dollar, 2000*]. By means of cross-country regressions Burnside and Dollar show that foreign aid has no impact on growth in countries with poor macroeconomic policies while it leads to faster growth in countries with good policies. In *Assessing Aid* the World

Carl-Johan Dalgaard and Henrik Hansen are both at the Institute of Economics, University of Copenhagen. The authors are grateful for the helpful comments of an anonymous referee.

Bank takes the consequence of this finding in stating that donors should direct foreign aid to countries with good economic policy (since aid is otherwise wasted when the sole purpose of aid is to foster economic growth). The empirical results reached by Burnside and Dollar have been questioned in recent studies by Hansen and Tarp [*2000, 2001*] and Lensink and White (this collection). Yet due to differences in data definitions Burnside and Dollar, according to Beynon [*1999*], state that their results have never been challenged.

In the present study we reconsider the foundations for the policy selectivity recommendation both theoretically and empirically. In section II we develop a simple neo-classical growth model in which firms face a risk of having part of the production destroyed because of social unrest, riots, thievery, and the like. A novel feature of the model is that this 'risk of destruction' is inversely related to the level of consumption, which in turn can be increased by foreign aid transfers. The main predictions from the model are: (i) aid has a positive impact on income per capita, (ii) there are diminishing returns to aid in raising long run (steady state) income per capita, and (iii) good policies – which are beneficial to growth by themselves – are likely to *reduce* the effectiveness of aid. Hence, the model motivates a further investigation of the aid-growth nexus.

The empirical investigation of the aid-growth results – which is based on the Burnside and Dollar data set – starts in section III with a comparison of different aid measures. In section IV the central growth regression in Burnside and Dollar [*2000*] is analysed using standard regression diagnostics. The main finding is that the policy selectivity result is very fragile, as it is extremely data dependent. It appears that five observations, which are excluded in Burnside and Dollar's preferred regressions, have a critical influence on the parameter of main interest. In a simple counter example it is shown that one may, on an equally valid statistical basis, claim that aid spurs growth – unconditionally. In section V it is demonstrated that a model with decreasing marginal effect of aid on growth is preferred to the policy selectivity model when the choice is based on statistical significance of parameters. This result was also established in Hansen and Tarp [*2000*], so the novelty in this paper lies in the use of the exact Burnside and Dollar data. A new result is that in the model with decreasing returns to aid there is a significant difference between least squares estimates and instrumental variable estimates. This points towards endogeneity of aid in the growth regressions and it highlights the importance of the choice of instruments. Finally, section VI concludes.

II. THEORETICAL MODELS OF AID AND GROWTH

In view of the importance of foreign aid, both for donors and recipients, and the voluminous aid literature there is a surprising scarcity of theoretical models linking aid and growth. Neo-classical models became highly influential in development economics from the beginning of the 1980s, but the prototype model used in assessing the macro-effectiveness of foreign aid continued to be the Harrod-Domar model and the two-gap model by Chenery and Strout [*1966*] even until the mid 1990s. Boone [*1994*, *1996*] was one of the first to analyse the macroeconomic impact of aid in a neoclassical growth model. Boone looked at fungibility issues in a standard growth model with productive public expenditure as in Barro [*1990*]. He found no effect of aid in the long run because aid is consumed instead of invested.[2]

Burnside and Dollar [*2000*] do not consider productive government spending but discuss government consumption and tax distortions. The economic model they consider is based on an aggregate production function of the form $Y = BK^{\theta}$, where Y is production and K is capital. Assuming that aid can only affect output through capital accumulation, effectiveness of aid can be approximated by

$$\frac{dY}{Y} = \theta \frac{Y}{K} \frac{\partial K}{\partial A} \frac{dA}{Y},$$

where A is real aid and $\theta \frac{Y}{K}$ is the marginal productivity of capital which, in the absence of credit rationing, equals the rate of return on capital.

Burnside and Dollar interpret the estimated derivative of growth with respect to aid as an estimate of $\theta \frac{Y}{K} \frac{\partial K}{\partial A}$, the product of the marginal productivity of capital and the marginal propensity to invest aid. (The latter reflects the fungibility problem.) This interpretation is no different from many of the aid-growth regressions from the 1970s and 1980s following the approach used by Papanek [*1973*]. However, while the expression was taken to be (roughly) constant in the early studies, Burnside and Dollar assume that it varies with economic policy. But, with this interpretation any variable that changes the marginal productivity of capital must be included in an interaction with aid. A cursory reading of the recent growth literature suggests an overwhelming number of additional variables, such as the Adelman-Morris index of socio-economic development [*Adelman and Morris, 1967; Temple and Johnson, 1998*], income inequality (Alesina and Rodrik [*1994*], among others), and of course human capital, just to mention a few.[3]

It is important, however, to realise that even within a neo-classical framework it is possible to derive predictions that directly contrast Boone and Burnside and Dollar. We will, therefore, show in this section that one can formulate growth models in which (i) fungibility is not the main problem for aid effectiveness, (ii) the marginal effect of foreign aid is not equal to the return to capital, and (iii) good policies (ones that are themselves important for growth) may *reduce* the marginal impact of foreign aid on growth.

Our growth model, which is capable of illustrating these points, is based on the Ramsey-Cass-Koopmans framework. It must be stressed from the outset that this model only serves as an example. We do not consider this to be the only way to model the relationship between aid and growth. But we do believe the model captures some important aspects of aid and growth in developing countries.

Starting with the standard assumptions, we consider a closed economy with competitive factor markets and perfect credit markets. For simplicity there is no exogenous technical progress and the population is constant.

The model deviates from the standard neo-classical growth model in two respects. First, we include foreign aid by considering pure income transfers, which enter the budget of the representative consumer as in Obstfeld [*1999*]. Second, following Barro and Sala-i-Martin [*1995: 159–60*] we let producers face a risk of expropriation, or some similar loss of the return to capital. We assume that with some probability $(1 - p)$ output is destroyed as a consequence of social unrest, riots and the like. The return probability, p, is taken as given at the individual level but will be endogenous at the aggregate level. While Barro and Sala-i-Martin model the return probability as a function of government expenditure, Alesina and Perotti [*1996*] consider the distribution of income and the standard of living. Combining the two types of models we can formalise the return probability as

$$p = p(G(t)/Y(t), c(t), \Delta^2),$$

where $G(t)/Y(t)$ is the relative size of government expenditure (policy), $c(t)$ is per capita consumption which proxies the standard of living, and Δ^2 is a measure of income inequality. For tractability, we assume that it is the flow of government expenditure that affects the return probability. One could argue that current income ought to enter the p-function on its own. It is worth noting, however, that by allowing the level of consumption to enter the expression we are implicitly assuming that permanent income is what matters for the incentive to engage in disruptive activities. Furthermore, we abstract from income inequality altogether in what follows. Consequently, it is dropped from the p-function. For government expenditure we make the standard assumption of a balanced, tax financed, budget at all times:

$$G(t) = \tau Y(t),$$

where τ is the constant proportional tax on production.

Based on the above p is increasing in policy and consumption ($p_\tau > 0$, $p_c > 0$). We follow Barro and Sala-i-Martin [*1995*] in imposing diminishing returns to expenditure ($p_{\tau\tau} < 0$). In addition, we assume diminishing returns to consumption per capita ($p_{cc} < 0$). Hence, an incremental increase in consumption will, on the margin, reduce the frequency of public disturbances more in a poor country compared to a richer one.

Producers employ capital and labor until the expected after tax marginal productivity equals the price on each factor. When the production technology is given as a Cobb-Douglas function with constant returns to scale and allowing for capital depreciation at the rate δ, it holds at all points in time that[4]

$$\begin{aligned} R(t) &= p(\tau, c(t))(1-\tau)\alpha k(t)^{\alpha-1} - \delta, \\ w(t) &= p(\tau, c(t))(1-\tau)(1-\alpha)k(t)^{\alpha}, \end{aligned} \tag{1}$$

where $R(t)$ is the expected after tax return to capital, $w(t)$ is the real wage, and $k(t)$ is the capital–labour ratio.

Consumers maximise the discounted utility from consumption. In the present model this is formalised as[5]

$$\max_{(c(t))_{t=0}^{\infty}} U(0) = \int_0^{\infty} \frac{c(t)^{1-\sigma}}{1-\sigma} e^{-\rho t} dt,$$

subject to

$$\begin{aligned} &\dot{k}(t) = R(t)k(t) + w(t) + a - c(t), \quad k(0) \text{ given} \\ &k(t) \geq 0 \quad \text{for all } t. \end{aligned}$$

Foreign aid, a, is included in the model as a pure income transfer to the representative consumer. For simplicity the aid inflow is assumed to be constant.

The usual computations lead to the Keynes-Ramsey rule and by using equation (1) it follows that

$$\frac{c(t)}{c(t)} = \frac{1}{\sigma}[p(\tau, c(t))(1-\tau)\alpha k(t)^{\alpha-1} - \delta - \rho]. \tag{2}$$

Hence, consumption will be growing if the return on capital investments exceeds the rate of time preferences, ρ. Additionally, the equation that governs the accumulation of capital is

$$\dot{k}(t) = p(\tau, c(t))(1-\tau)k(t)^{\alpha} + a - c(t) - \delta k(t). \tag{3}$$

The solution of the model is given by the two differential equations (2) and (3). In order to ensure existence of a (saddle point stable) steady state we assume

$$\frac{\rho+(1-\alpha)\delta}{\alpha}\frac{k}{c}\eta_{p,c}<1-\alpha$$

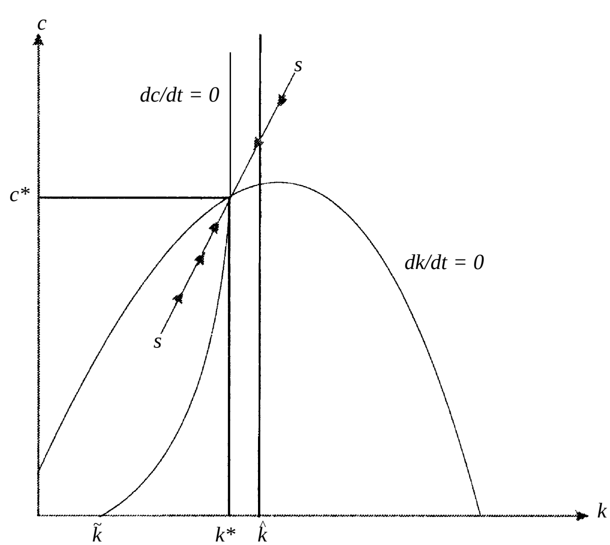
to hold for all steady state values of k, c, where $n_{p,c}$ is the elasticity of the return probability with respect to consumption.[6]

At any given point in time the economy will be traveling along the saddle-path, the SS-curve in Figure 1. The main difference to the standard neoclassical growth model is that the curve is upward sloping, being only asymptotically vertical. This is because a higher level of consumption increases the return on capital. This change of the standard growth model results in quite interesting predictions about aid, policies and growth. Some of these are discussed below.

First of all, foreign aid has an impact on the long-run level of income as long as the return probability is less than unity. We expect p to be low in poor countries as they have low levels of per capita consumption. In these

FIGURE 1

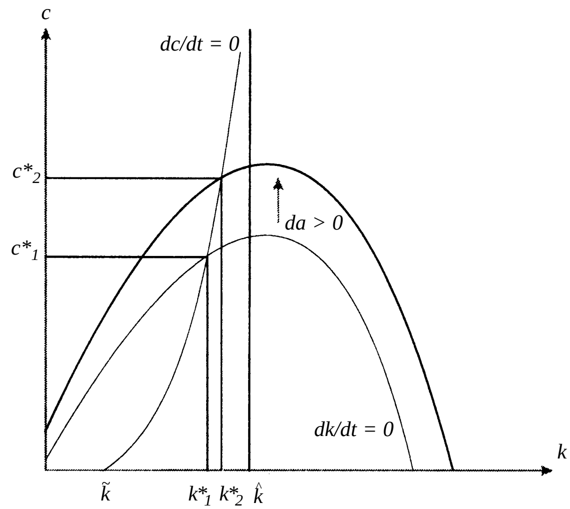
PHASE-DIAGRAM



$$\tilde{k} = \bar{p}^{\frac{1}{1-\alpha}}\left(\frac{\alpha(1-\tau)}{\rho+\delta}\right)^{\frac{1}{1-\alpha}},\ \hat{k} = \left(\frac{\alpha(1-\tau)}{\rho+\delta}\right)^{\frac{1}{1-\alpha}}.$$

FIGURE 2

THE EFFECT OF AN UNANTICIPATED PERMANENT INCREASE IN FOREIGN AID



countries foreign aid leads to an increase in consumption, which in turn increases the expected return on investment for a given level of government expenditure. This result is illustrated in Figure 2. Moreover, it can readily be shown that the model allows for diminishing returns to aid in the sense that $\partial^2 k^*/\partial a^2 < 0$, when the level of aid becomes sufficiently large. Figure 2 illustrates this result as well. Increases in foreign aid moves the $\dot{k} = 0$ curve upwards leading to changes in the steady state levels along the $\dot{c} = 0$ curve. As the slope of this curve is changing ($p_{cc} < 0$) the result follows for sufficiently large increases in aid compared to the initial steady state level of consumption.

The order of magnitude of the impact of foreign aid on the long-run levels of consumption, capital and income per capita reveals that empirical results must be interpreted with great care. Letting $v = (\rho + (1 - \alpha)\delta)/\alpha$, and $z^* = k^*/c^*$, where the asterisk denotes steady state values, we find the following impact:

$$\frac{dc^*}{da} = \frac{1-\alpha}{1-\alpha - v z^* \eta_{p,c}} > 1, \tag{4}$$

$$\frac{dk^*}{da} = \frac{z^* \eta_{c,p}}{1-\alpha - v z^* \eta_{p,c}} > 0, \tag{5}$$

$$\frac{dy^*}{da} = \frac{1+\alpha\eta_{k,p}}{1-\tau} \frac{(\nu-\rho)z^*\eta_{p,c}}{1-\alpha-\nu z^*\eta_{p,c}} > 0, \tag{6}$$

As seen, the long-run impact on consumption exceeds one while all we can say about the impact on capital is that it is positive. Yet, this implies that aid effectiveness cannot be tested by system regressions of consumption and investment on the aid inflow as argued in, for example, Boone [*1996*] because there is no adding up constraint. As for the impact on long run income we note that it is quite different from both the expected private return and the expected social return on capital investment. Hence, there is no theoretical reason for comparing partial effects of investment and aid in growth regressions.

Turning to the interplay between aid and good policy we first need to be specific about the notion of good policy. In most models with tax financed productive government expenditure there is an inverse U-shaped relation between the relative size of the public sector ($G/Y = \tau$) and growth. The present model also has this property. Therefore we will start by assuming that government policies designed to ensure private property rights are initially below the growth maximising level.[7] This makes an increase in τ a good policy.

Good policy has an impact on the transmission of aid to long-run growth if it changes the marginal impact of consumption on the return probability, p_c. To see why, consider the slope of the $\dot{c} = 0$ curve

$$\frac{dc}{dk} = (1-\alpha)\frac{p(\tau,c)}{p_c}\frac{1}{k}, \quad \text{or} \quad \frac{dc}{dk}\frac{k}{c} = \frac{1-\alpha}{\eta_{p,c}}.$$

One possibility would be to argue that, say, more police will dampen the incentive to participate in, for example, riots, which suggests that $p_{c\tau} < 0$. If this is the case, the $\dot{c} = 0$ curve becomes steeper following a once and for all increase in τ. This makes the effect of increasing foreign aid smaller. Hence, good policies can be beneficial for growth and at the same time *reduce* the effectiveness of foreign aid. This is so because good policies and aid are 'substitutes' in this case.

If government expenditure is initially above the growth maximising level then reducing τ is good policy. (The security provided is too costly.) In this case good policy will increase the effectiveness of aid. Overall, the result is that aid can replace government expenditure when $p_{c\tau} < 0$, while government expenditure is only able to replace aid up to a certain point.

Of course one may also assume the opposite relation between consumption and government expenditure, $p_{c\tau} > 0$. In this case government services complement the aid inflow in ensuring social stability. This reverses the above results. Growth stimulating policies enhance the effect of foreign aid when expenditures are initially too low, while they lower the effect when expenditures are initially too high.

In sum, the link between aid and good policies in the growth process is ambiguous. This makes empirical work all the more important and it stresses the need for careful testing of new empirical regularities before wide-ranging policy changes are initiated. In the sections to follow we will therefore test Burnside and Dollar's empirical result of increased effectiveness of aid in a good policy environment.

III. REAL EFFECTIVE AID VERSUS NOMINAL OFFICIAL AID

Burnside and Dollar are the first to use a new database on foreign aid compiled by Chang *et al.* [*1998*] for the World Bank. The main difference between the new aid measure (effective development assistance, EDA) and the measure used by other authors (official development assistance, ODA) is that EDA is the sum of grants and the grant equivalents of official loans whereas ODA includes both the direct grants and concessional loans for which the grant component is above 25 per cent.

Furthermore, Burnside and Dollar refrain from the standard practice of relating the aid flows in current dollars to GDP in current dollars. Instead they construct real aid, measured in constant 1985 dollars, using the unit-value of imports price index from the IFS. Real effective development assistance is subsequently divided by real GDP from the Penn World Tables, Mark 5.6 [*Summers and Heston, 1991*].

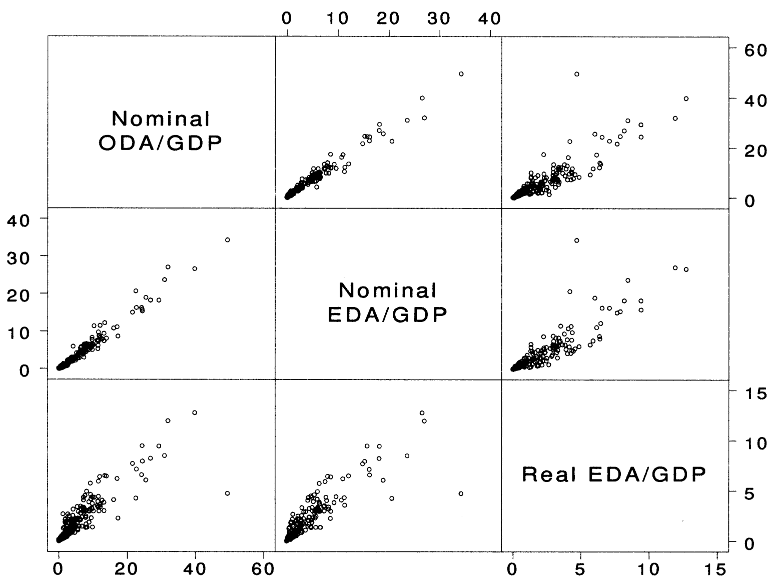
While the EDA measure in all likelihood provides a better picture of actual resource flows compared to ODA it seems odd to name the flow *effective* development assistance when the effect of procurement tying of aid by bilateral donors is not even touched upon. But in relation to the growth regressions in Burnside and Dollar it is more interesting to discuss the distinction between real (PPP-adjusted) and nominal flows. According to Beynon [*1999*], Burnside and Dollar argue that nominal aid to nominal GDP is vulnerable to suggesting spurious changes in aid levels in response to rapid changes in the exchange rate. The example used in Beynon is that 'a 50% devaluation that effectively halves the $ denominated level of GDP would imply an instant but erroneous (assuming the bulk of aid dollars to be spent on foreign currency items) doubling in aid' [*Beynon, 1999: Annex 2, 20*].

However, long discussions of theoretical consistency and spurious changes in aid flows seem immaterial once we look at the data. Figure 3 shows cross plots of three aid measures that may all enter the growth regression. The top row (left column) in the matrix plot has the standard, nominal ODA to nominal GDP on the vertical (horizontal) axes. The center row (column) has nominal EDA to nominal GDP on the vertical (horizontal) axes. Finally, the bottom row (right column) has real EDA to real GDP, as defined by Burnside and Dollar, on the vertical (horizontal) axes.

As seen, ODA is somewhat higher than EDA, but it is quite easy to draw a straight line through most of the points. Hence, despite the valuable effort by Chang *et al.* [*1998*] in their construction of a better measure of aid flows, the difference between EDA and ODA seems to be a simple transformation. This is confirmed by the correlations between nominal ODA and EDA given in Table 1. In the table, standard Pearson correlations are given above the diagonal while Spearman's rank correlations are reported below the diagonal. The correlation between the two nominal measures is 0.98 using either formula.

FIGURE 3

CROSS PLOTS OF AID MEASURES



Vertical axes:
Top: nominal ODA to nominal GDP. Center: nominal EDA to nominal GDP. Bottom: real EDA to real GDP.

Horizontal axes:
Left: nominal ODA to nominal GDP. Center: nominal EDA to nominal GDP. Right: real EDA to real GDP.

TABLE 1

CORRELATIONS BETWEEN AID MEASURES

	Nominal ODA/GDP	Nominal EDA/GDP	Real EDA/GDP
Nominal ODA/GDP		0.98 (0.98)	0.89 (0.94)
Nominal EDA/GDP	0.98		0.88 (0.93)
Real EDA/GDP	0.95	0.95	

Note: Pearson (Spearman) correlations above (below) diagonal. Pearson correlations when Somalia 78–81 is excluded are in parentheses.

Turning to the relation between the nominal and real measures we find a higher dispersion in the cross plots. However, much is caused by a single outlier. The outlier in the plots is Somalia 78–81. The nominal EDA/GDP ratio for Somalia is 32 per cent while the real ratio is six per cent. The reason for this discrepancy can be found in the ratio of the PPP-adjusted GDP from PWT to the constant dollar GDP from the World Bank. The PPP-adjusted GDP is more than six times the GDP from the World Bank for Somalia in all periods for which we have data. Hence, this is not a rapid change in the exchange rate. Calculating correlations between the real and nominal aid measures with and without the observation for Somalia results in some differences in the Pearson correlations but no change in the Spearman correlations. Yet, even including Somalia 78–81 we find a Pearson correlation of 0.89 between the new measure used by Burnside and Dollar and the standard aid measure. This correlation increases to 0.94 when the single Somalia observation is left out.

In conclusion, with respect to discussions of the proper aid measure Burnside and Dollar may or may not be right in the way they have chosen to measure aid flows in terms of a world price metric. But this certainly depends on the underlying theoretical model. Based purely on simple statistical properties of the different aid measures it seems as if the aid effectiveness results obtained by Burnside and Dollar, using real effective development assistance, are comparable to studies using nominal official development assistance. Thus, the cause of the divergence must be sought elsewhere.

IV. INFLUENTIAL OBSERVATIONS AND OUTLIERS

In this section we will take a close look at Burnside and Dollar's preferred growth regressions. The focus will be on detecting influential observations and outliers in the data set. The reason for this special interest is that while the important interaction between aid and policy is insignificant in the full sample of 56 developing countries and in a sub-sample of 40 low-income countries, Burnside and Dollar show that once five 'big outliers' are excluded from the regressions the interaction is significant. As shown below, there are many other influential observations in the data set, which may deserve special attention.

In the analysis we make use of standard regression diagnostics for influential observations and outliers (see Belsley, Kuh, and Welsch [*1980*] among many others). This means that all diagnostics are based on ordinary least squares regressions with no account for possible heteroskedasticity. Some results will change if heteroskedasticity consistent standard errors are used. But this is at the expense of the simple relations between diagnostics, residuals, and regressor influence. As the regression diagnostics have been developed as tools in informal analyses, that is, we do not use formal hypothesis testing, and for the sake of easy replicability, we have chosen to apply the standard tools.[8]

Table 2 reports results of reestimation of Burnside and Dollar's preferred growth equations. Regressions (1)–(3) include all 56 countries while regressions (4)–(6) only include the 40 lower-income countries as defined in Burnside and Dollar [*2000*]. For the sake of clarity, we will briefly browse through the list of regressors, even though the specification and the data are identical to the Burnside and Dollar study.

The dependent variable is the average growth rate in real GDP per capita over six four year periods, starting with 1970–73 and ending with 1990–93. The GDP variable is from Penn World Tables (Mark 5.6).

The first regressor, Initial GDP, is the logarithm of GDP per capita in the last year preceding the period for which the growth rate is calculated. The variable is expected to have a significant negative influence on the growth rate, capturing the conditional convergence effect.

The following three regressors, Ethnic fractionalisation, Assassinations, and the product of the two, are included in growth regressions to capture political instability. The number of assassinations varies over time while ethnic fractionalisation is time constant (based on data from 1960). The two variables are expected to have a negative influence on growth.

The two regressors, Institutional quality and M2/GDP, are included as proxies for the quality of institutions and the financial markets. The first variable is an index based on evaluations of five different institutional indicators made by the private international investment risk service, International Country Risk Guide. The five indicators are: Quality of the bureaucracy, Corruption in Government, Rule of Law, Expropriation Risk, and Repudiation of Contracts by Government [*Knack and Keefer, 1995*]. It is worth noting that the, time constant, institutional quality variable is based on evaluations in 1982 or later, which is roughly in the middle of the sample. Hence, there is a strong assumption of constancy and exogeneity of institutions as measured by the five indicators. The proxy for the development of financial markets is broad money (M2) relative to income (GDP). In the growth regression predetermined observations are used in order to avoid simultaneity problems.

Turning to policy, Burnside and Dollar create an index covering aspects of fiscal, monetary, and trade policies. Fiscal policy is measured by the budget surplus. The level of inflation measures the success or failure of monetary policy, while trade policy is proxied by a binary (0/1) openness indicator, constructed by Sachs and Warner [*1995*]. To avoid collinearity problems Burnside and Dollar create an index using a weighted average of the three measures:[9]

$$\textit{Policy} = 1.28 + 6.85\ \textit{Budget surplus} + 1.40\ \textit{Inflation} + 2.16\ \textit{Trade openess}.$$

As seen, the construction of the index is such that good policy, in terms of a budget surplus, low inflation and an open economy, leads to a high value of the index. Hence, the effect on growth is expected to be positive.

Finally aid is included in the growth regression. Aid is real EDA to real GDP as discussed in section III. In the preferred regressions aid enters as a regressor on its own and multiplied by the policy index. This latter regressor is denoted the interaction effect.

In addition to the regressors listed in Table 2 the model also includes time dummies and dummies for sub-Saharan Africa and East Asia. The dummies are found significant in almost all empirical growth studies. However, in terms of ‘explaining’ differences in growth rates they have little to offer and they are, therefore, not reported in the following. Instead, two goodness-of-fit measures are reported for all regressions. The first is the standard R^2 measure, while the second is the partial R^2 for the model conditional on time dummies and regional dummies.

Regressions (1) and (4) in Table 2 give results for the preferred specifications estimated on the two full samples. The regressions show that only two variables are significant at a five per cent level: Institutional quality and the policy index. Most importantly, aid has no significant impact on growth in these regressions.

Moving from regressions (1) and (4) to (2) and (5) shows the changes in the parameters when five observations are excluded from the samples. The five observations are Gambia 1986–89, 1990–93, Guyana 1990–93, and Nicaragua 1986–89, 1990–93. While the coefficient to the interaction effect is small and highly insignificant in (1) and (4) it increases more than ten-fold in the large sample and even fifty-fold in the lower income country sample. In addition, in (2) and (5) the interaction parameters are significant at the five per cent level.

The reason for excluding the five observations is apparently that these observations have a very big influence on the coefficient to the aid-policy

TABLE 2

GROWTH REGRESSIONS WITH INTERACTION EFFECTS BETWEEN AID AND POLICY

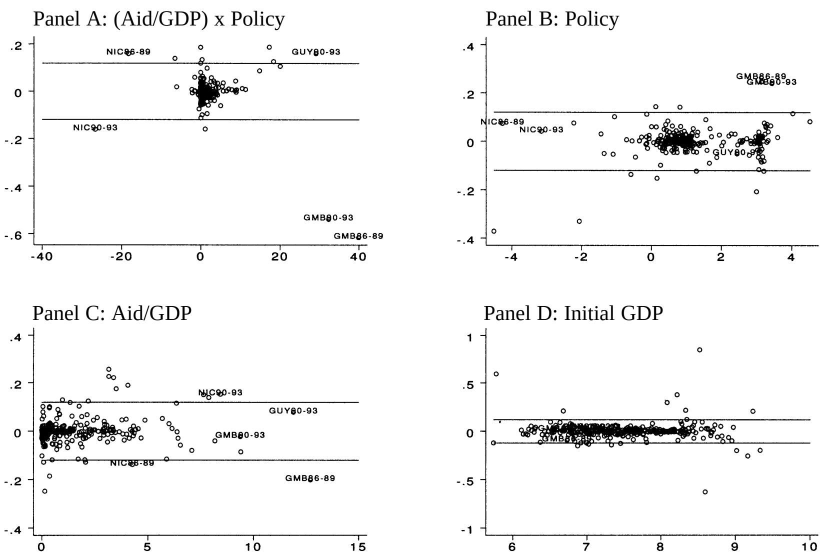
	All 56 countries			40 lower income countries		
Regression	(1)	(2)	(3)	(4)	(5)	(6)
Initial GDP	–0.618	–0.595	–0.451	–0.730	–0.697	–0.358
	(0.581)	(0.586)	(0.575)	(0.833)	(0.847)	(0.794)
Ethnic fractionalization	–0.564	–0.431	–0.498	–0.784	–0.587	–0.766
	(0.744)	(0.747)	(0.732)	(0.849)	(0.840)	(0.835)
Assassinations	–0.441	–0.449*	–0.425	–0.748	–0.787*	–0.670
	(0.271)	(0.268)	(0.265)	(0.478)	(0.458)	(0.482)
Ethnic frac. x Assassin.	0.807*	0.794*	0.824*	0.926	0.678	1.108
	(0.457)	(0.455)	(0.449)	(0.943)	(0.962)	(0.927)
Institutional quality	0.645**	0.693**	0.704**	0.784**	0.856**	0.887**
	(0.177)	(0.177)	(0.169)	(0.205)	(0.206)	(0.192)
M2/GDP (lagged)	0.013	0.012	0.004	0.027	0.023	0.015
	(0.014)	(0.015)	(0.013)	(0.017)	(0.018)	(0.015)
Policy Index	0.956**	0.705**	1.041**	1.107**	0.541*	1.168**
	(0.190)	(0.195)	(0.142)	(0.323)	(0.320)	(0.189)
Aid/GDP	0.017	–0.016	0.249**	–0.036	–0.173	0.214*
	(0.125)	(0.165)	(0.124)	(0.135)	(0.175)	(0.126)
(Aid/GDP) x policy	0.013	0.184**		0.005	0.265**	
	(0.050)	(0.071)		(0.062)	(0.089)	
Observations	275	270	270	189	184	184
R^2	0.39	0.39	0.42	0.46	0.47	0.51
Partial $R^{2(a)}$	0.22	0.26	0.25	0.29	0.41	0.35

Note: The dependent variable is real per capita GDP growth. All regressions include time dummies for each period in the sample and dummies for sub-Saharan Africa and East Asia. White's heteroskedasticity consistent standard errors in parentheses. [(a)]R^2 when the effect of time and regional dummies is partialled out. *Significant at the 10 per cent level. **Significant at the 5 per cent level.

interaction term. Each observation's influence on the parameter estimates can be investigated simply by regressing the model on a sample in which the single observation is excluded. Figure 4 shows cross-plots of the scaled changes in the estimated coefficients for four of the regressors in the model when observations are excluded one-by-one. The changes in the estimated coefficients are plotted against the excluded observation.[10] Burnside and Dollar use a slightly different measure, as they do not scale the change in the estimated coefficient by the estimated standard error. While there are different views on whether or not the changes should be scaled, in the present context the scaled measures have the advantage that as the unit of measurement is in terms of (approximate) standard errors it is possible to compare the magnitude of changes across parameters. Moreover, for the scaled measure there are simple rules of thumb for changes worth investigating. Belsley *et al.* suggest using $\pm 2/\sqrt{n}$ as a cut-off point, where n

FIGURE 4

INFLUENTIAL DATA POINTS FOR SOME PARAMETERS OF INTEREST



is the number of observations in the regression. Others, for example, Velleman and Welsch [*1981*], have suggested less stringent cut-off values such as ±1 or ±2 based on the *t*-like diagnostics. In Figure 4 horizontal lines indicate the cut-off points $\pm 2/\sqrt{n}$ are indicated by horizontal lines.

Starting with Panel A in Figure 4 it is obvious that all of the five excluded observations have a critical influence on the estimated coefficient to the interaction term. Especially the two observations for Gambia move the estimate towards zero. However, notice that the five observations are not the only ones having a critical influence on the estimate; there are six other observations outside the more strict cut-off value while none of the scaled changes exceed one in absolute value. Hence, without more information some investigators would not find any cause for action, such as deleting observations from the sample.

Moving to Panel B, we find nine possibly critical observations for the policy coefficient, of which only Gambia is in the Burnside and Dollar exclusion set. In Panel C it is possible to pinpoint no less than 12 influential observations for the aid coefficient when the strict cut-off value is applied. Among these we find the two observations for Nicaragua and one of the observations for Gambia. The most interesting plot, however, is for the initial GDP. In this plot (Panel D) there are no less than 19 influential observations, none of which are in the excluded set. Notice that some of the scaled changes in the coefficient to initial GDP are larger in absolute value than the scaled changes in the coefficient to the aid-policy interaction. Thus, using this metric there are other observations more liable as candidates for deletion.

Figure 4 reveals that the five excluded observations are the extreme values of the aid policy interaction regressor. Such points are not considered as outliers in classical regression analysis; they are (possible) leverage points. Of course they may still be deleted if the information they convey is considered to be different from the rest of the observations. But this deletion rule is clearly *ad hoc* and it is rather odd to limit the variation in the central regressor in this way.

In order to provide a more comprehensive picture of the influential observations in the sample we have listed influence measures for 23 observations in Table 3. The criteria for inclusion in the Table is that at least

TABLE 3

POTENTIALLY INFLUENTIAL OBSERVATIONS

Country	Period	Outliers	Leverage	Influence on (Aid/GDP) x Policy	Policy	Aid/GDP	Initial GDP	Fit
Argentina	74–77		0.315		–0.137			0.673
Bolivia	82–85			0.139	–0.374			
Botswana	78–81			0.187				0.557
Brazil	86–89			0.186	–0.331			
Chile	82–85			0.134	–0.209			
Cameroon	78–81	3.67					–0.151	0.717
	90–93	–3.05				0.124		–0.661
Ecuador	70–73			–0.160	0.246			
Egypt, Arab Rep.	82–85	2.17					–0.129	0.558
Ethiopia	82–85	–3.92				0.226	0.593	–0.913
Gabon	70–73					0.221	0.376	0.548
	74–77	4.46					0.846	1.214
	78–81	–3.41			–0.123		–0.631	–0.912
Gambia, The	86–89*		0.295	–0.617	0.263	–0.202		–0.989
	90–93*		0.186	–0.539	0.240			–0.778
Guyana	90–93*		0.299	0.160				
Nicaragua	78–81	–3.52				–0.129		–0.702
	86–89*			0.159		–0.139		
	90–93*		0.406	–0.162		0.152		
Nigeria	70–73	2.11				–0.187		0.507
Philippines	82–85				0.141			–0.512
Syrian, Arab Rep.	74–77	2.59				0.256	0.297	0.601
	78–81					0.174	0.215	

Note: Observations are included in the table if they exceed at least two cut-off values. The cut-off values are: |DFITS| > 0.5, |DFBETAS| > 0.12. The studentised residuals (Outliers) are only reported if they exceed 2 in absolute value. The leverage points are only reported if they exceed 0.18. See Appendix for definitions of the applied influence measures. *Outlier in the Burnside and Dollar study.

two influence measures must exceed a preselected cut-off value. Of the many possible influence diagnostics we have, arbitrarily, chosen the scaled change in the coefficients for the four variables shown in Figure 4 and the scaled change in the overall fit. In addition to the scaled changes in the parameters and the overall fit we also report the studentised residuals and a leverage measure (for the diagonal of the hat matrix, see the Appendix).

Table 3 makes clear that the five deleted observations are not outliers in the sense of having extreme studentised residuals. None of the studentised residuals for these observations exceed 2 in absolute value. However, four of the five observations are leverage points, meaning that they have an above-average influence on the fitted values. But, there are other observations in the sample with even higher leverage values.

The danger of deleting observations from the sample based on high influence on one or a few special parameters is revealed in regressions (3) and (6) in Table 2. In these regressions we have omitted the interaction term and searched for a sample, of the same size as the Burnside and Dollar sample, in which the coefficient to aid by itself is positive and significant. By excluding five observations, Gambia 1986–89, 1990–93, Nigeria 1970–73, 1990–93, and Nicaragua 1978–81, we obtain the result we are looking for: Aid has a significant impact on growth. In the 56-country sample the parameter is highly significant, while it is only significant at the nine per cent level in the sample excluding middle-income countries.[11] Notice that the two most influential observations in the Burnside and Dollar regressions are also excluded in (3) and (6). As such it seems extremely difficult to reject regressions (3) and (6) and at the same time accept (2) and (5). Yet, the former model *does not* have a policy selectivity rule.

Finally, it should be noted that if outliers are detected and down weighted in a mechanical way, using a robust regression method, the main change in the result compared to (1) and (4) is a significant negative coefficient to initial GDP, indicating conditional convergence.[12] Aid and the interaction term are still insignificant in the robust regressions. It must be stressed, though, that the robust regression does not restrict influence from outlying points in the regressor space unless they lead to large residuals. Therefore, the robust regression results can only be used to show that the lack of significance of the interaction term cannot be attributed to big residual outliers.

V. FUNCTIONAL FORM AND ENDOGENEITY OF AID

The lack of robustness of the Burnside and Dollar specification may be due to model misspecification. In particular, the theoretical results in section II and the empirical studies by Hadjimichael *et al.* [*1995*]; Durbarry *et al.*

[*1998*]; Hansen and Tarp [*2000, 2001*], and Lensink and White (this collection) show that modelling decreasing returns to aid may be important.

Whether one should prefer a policy selectivity model with interaction between aid and policy or a diminishing returns model with a polynomial effect of aid – or even a combination – is a simple testable hypothesis. In Hansen and Tarp [*2000, 2001*] it is argued that a full model must include five aid-policy terms: aid, policy, aid squared, policy squared, and aid interacted with policy. The argument is quite simple; these five terms define a complete, second order, polynomial response surface in the growth-aid-policy space. To be precise, consider the stylised growth regression equation

$$g_t = \gamma' Z_t + \beta_1 P_t + \beta_2 A_t + \beta_3 P^2{}_t + \beta_4 A^2{}_t + \beta_5 P_t A_t + \varepsilon_t ,$$

where g_t is the growth rate, Z_t is a set of controls, P_t is a policy index, A_t is aid and γ, β denotes parameters. In this model Burnside and Dollar set $\beta_3 = \beta_4 = 0$ while Hadjimichael *et al.*, Durbarry *et al.*, and Lensink and White set $\beta_3 = \beta_5 = 0$, all without testing the hypothesis. Lensink and White (this collection) test $\beta_5 = 0$ conditional on $\beta_3 = 0$. They find the interaction term to be highly insignificant.

When the specifications are tested within the full model Hansen and Tarp [*2001, 2000*] find statistical support for diminishing returns ($\beta_3 = \beta_5 = 0$, $\beta_4 \neq 0$). According to Beynon [*1999*], Burnside and Dollar reconcile this finding by stressing that they are using different data compared to the other studies. Therefore we end the analysis of the Burnside and Dollar data by showing that when the set of instruments is chosen to achieve a good (time series) fit of the endogenous aid variables the diminishing returns model is preferred to a policy selectivity model.

Table 4 presents results of instrumental variable regressions of real growth in GDP per capita in which all regressors that are functions of aid are modeled as endogenous variables. The Table is organised as Table 2 in that the first three columns (regressions 7–9) give results for the full sample of 56 countries while the last three columns (regressions 10–12) give results for the sample of 40 low-income countries. In Table 4 there is no exclusion of observations due to outliers but the first estimation period (1970–73) is unavailable because lagged observations of all aid regressors are used as instruments. This is why there are only 223 and 153 observations in the two samples instead of 275 and 189.

Regressions (7) and (10) reveals that aid and aid squared are both significant while the aid–policy interaction and policy squared are both insignificant, at any conventional level of significance, when all four terms are included in the growth regression. This result is in contrast to Burnside and Dollar [*2000*] but in agreement with Hansen and Tarp [*2001*].

TABLE 4

INSTRUMENTAL VARIABLE GROWTH REGRESSIONS WITH ENDOGENOUS AID

	All 56 countries			40 lower income countries		
Regression	(7)	(8)	(9)	(10)	(11)	(12)
Initial GDP	–0.007	0.012	–0.372	–0.011	–0.021	–0.346
	(0.811)	(0.771)	(0.747)	(1.181)	(1.037)	(1.031)
Ethnic fractionalization	0.549	0.575	–0.176	0.309	0.302	–0.350
	(0.992)	(0.969)	(0.862)	(1.120)	(1.067)	(0.986)
Assassinations	–0.455*	–0.453*	–0.414	–1.019*	–1.018**	–0.823
	(0.268)	(0.267)	(0.275)	(0.438)	(0.431)	(0.495)
Ethnic frac. x Assassin.	0.887*	0.882*	0.779	1.583	1.589	1.334
	(0.466)	(0.460)	(0.474)	(0.991)	(0.989)	(1.039)
Institutional quality	0.862**	0.865**	0.698**	0.933**	0.933**	0.878**
	(0.223)	(0.224)	(0.200)	(0.250)	(0.246)	(0.228)
M2/GDP (lagged)	0.009	0.009	0.012	0.026	0.026	0.024
	(0.020)	(0.019)	(0.015)	(0.020)	(0.020)	(0.017)
Policy Index	0.927**	0.958**	1.056**	1.127**	1.133**	1.273**
	(0.267)	(0.153)	(0.227)	(0.415)	(0.206)	(0.416)
Aid/GDP	1.327**	1.352**	0.229	1.031*	1.027**	0.166
	(0.549)	(0.530)	(0.211)	(0.546)	(0.514)	(0.212)
(Aid/GDP) squared	–0.126**	–0.127**		–0.095**	–0.095**	
	(0.046)	(0.049)		(0.044)	(0.045)	
Policy squared	0.012			–0.002		
	(0.064)			(0.076)		
(Aid/GDP) x Policy	0.006		–0.052	0.002		–0.062
	(0.065)		(0.071)	(0.081)		(0.083)
Effect of aid at mean	0.931**	0.946**	0.167	0.633*	0.629*	0.092
	(0.390)	(0.385)	(0.180)	(0.362)	(0.339)	(0.175)
Observations	223	223	223	153	153	153
R^2	0.36	0.36	0.39	0.45	0.45	0.48
Partial R^2 [(a)]	0.30	0.27	0.22	0.40	0.35	0.32
Wald test[(b)]		0.981	0.019		0.999	0.093
Sargan test[(c)]	0.942	0.942	0.494	0.998	0.998	0.959
DWH test[(d)]	0.016	0.004	0.121	0.046	0.014	0.221
Partial R^2 in reduced form regressions[(e)]						
Aid/GDP	0.71	0.68	0.67	0.75	0.78	0.74
Aid/GDP squared	0.34	0.32		0.43	0.46	
Aid/GDP x Policy	0.48		0.45	0.51		0.46

Note: The dependent variable is real per capita GDP growth. All regressions include time dummies for each period in the sample and dummies for sub-Saharan Africa and East Asia. White's heteroskedasticity consistent standard errors in parentheses. [(a)] R^2 when the effect of time and regional dummies is partialled out. [(b)] The p-value of a Wald type test of the imposed restrictions. [(c)] The p-value of a Sargan type test of over-identifying restrictions. (d) The p-value of a Durbin-Wu-Hausman test of equality of OLS and IV estimates. [(e)] R^2 in the reduced form regressions when the effects of the exogenous regressors in the growth regression are partialled out. *Significant at the 10 per cent level. **Significant at the 5 per cent level.

Instruments: See Table 5. (Aid/GDP) squared is not used as instrument in (9) and (12) as it leads to rejection of a test of over–identifying restrictions.

Moving to regressions (8) and (11) it is clear that omitting the two statistically insignificant variables, aid-policy and policy squared, leaves the significant parameters virtually unchanged. This result is substantiated by the Wald type test of the joint exclusion of the two variables. As seen from the table the *p*-values of the restriction are out of the ordinary, making it difficult to maintain the assumption of important aid–policy interaction effects in the growth equation. In contrast, regressions (9) and (12) reveal that exclusion of aid squared and policy squared leading to Burnside and Dollar's preferred specification, is rejected quite strongly in the 56 country sample and marginally in the 40-country sample. Moreover, when the two variables are excluded, the effect of aid on growth becomes insignificant. Overall, these results underline that the insignificance of the aid–policy interaction is not caused by collinearity problems between the three aid regressors.

A new result in Table 4 is the significance of the Durbin-Wu-Hausman (DWH) test for equality of the IV and OLS results. As this test is often interpreted as a test of endogeneity we give quite strong evidence in favor of endogeneity of aid in all four regressions involving aid and aid squared.

Possible endogeneity of aid disbursements has been recognised since the early 1970s. Papanek [*1972*] was the first to argue that a negative correlation between aid and savings may be caused, in part, by a need-based allocation of aid. Yet, until the 1990s Mosley [*1980*] was the only study in which endogeneity of aid flows were taken into account in the econometric analysis. In the 1990s Boone [*1994, 1996*] has had a significant impact on later studies by his emphasis on endogeneity and the choice of instruments. Both Hadjimichael *et al.*, Burnside and Dollar, and Lensink and White discuss endogeneity and Boone inspires Burnside and Dollar in the choice of instruments in their regressions. Interestingly, none of the studies find significant bias in the OLS regressions when they apply DWH type tests.

TABLE 5

INSTRUMENTS IN BURNSIDE AND DOLLAR'S REGRESSIONS AND IN TABLE 4

Specific to Burnside and Dollar	Common instruments	Specific to Table 4
Egypt dummy	Franc Zone dummy	Aid/GDP, lagged
Central America dummy	Policy x (logarithm of Initial GDP)	$(\text{Aid/GDP})^2$, lagged
Arms imports, lagged	Policy x $(\text{logarithm of Initial GDP})^2$	(Policy x Aid/GDP), lagged
Policy x (Arms imports, lagged)	Policy x (logarithm of population)	
Policy x $(\text{logarithm of population})^2$		
Logarithm of population		

While Boone has only one endogenous regressor (aid) in his studies, Burnside and Dollar have two (aid and aid times policy), and there are three in Table 4 (aid, aid times policy, and aid squared). The increase in the number of endogenous regressors gives rise to increasing demands for good instruments in terms of variation and correlation with the endogenous regressors. The increase in demands is reflected in the choice of instruments in the studies. Boone [*1996*] uses three different sets of instruments; (i) the log of population, (ii) Friends of US, Friends of OPEC, and Friends of France, and (iii) aid lagged twice. Burnside and Dollar [*2000*] combine the two first sets of instruments and add interactions with policy as seen from Table 5, while in Table 4 we use one of the political variables (Friends of France denoted Franc Zone following Burnside and Dollar), lagged aid, and some of the interactions with policy.

The most important difference in the choice of instruments between Burnside and Dollar and Table 4 is that and Burnside and Dollar rely much on time constant dummy variables as instruments for aid; a dummy for Egypt (Friend of US), Friends of France, and a dummy for Central America (see Table 5). In addition, the logarithm of population is only changing slowly over time. This means that the relation between aid and the instruments is mainly a cross country correlation leaving the time series variation in aid unexplained. Following Hansen and Tarp [*2000, 2001*] we try to increase the time series variation and the identification of the individual regressors by including lags of the three endogenous regressors. As seen from the bottom part of Table 4 there is a substantial correlation between the endogenous regressors and the instruments even after the variation, which is correlated with the exogenous regressors in the growth regressions, has been removed. Furthermore, the Sargan type test reported in Table 4 does not lead us to reject the validity of the instruments.

In sum, we have shown that combining the specification from Hansen and Tarp [*2001*] with the data from Burnside and Dollar [*2000*] lead us to the same conclusion as reached in Hansen and Tarp.

VI. CONCLUSION

In this study we have reassessed the aid effectiveness results in ‘Aid, Policies, and Growth’ by Burnside and Dollar [*2000*], using the same data set as the original study. We develop a neo-classical growth model in which aid spurs growth even in economies in which aid does not enter the production function directly as investment and we show that in this model the interplay between good policy and aid is ambiguous. If anything, good policy is likely to reduce the growth effect of aid because they act as substitutes in the growth process. This shows that the Burnside and Dollar result is far from obvious on theoretical grounds.

The main outcome of the empirical re-examination is that the finding of a more positive impact of aid on growth in good policy environments is not a robust result. It depends crucially on deletion of a few observations. We show that once we apply a Burnside-Dollar-type sample selection procedure, in which a single parameter of interest determines the estimation sample, it is possible to obtain different results. In particular, we obtain a positive effect of aid on growth in any policy environment.

A related result is that the Burnside and Dollar data is consistent with a non-linear relation between aid and growth in which there is diminishing returns to aid. This result conforms well to our theoretical model as well as other recent empirical aid effectiveness studies.

Based on the above results we find it premature to apply policy selectivity rules in future aid allocations as advocated in chapter one of *Assessing Aid*. This is so even though applying the policy selectivity rule will, almost surely, increase returns to aid when these returns are measured as the correlation between aid and growth in income per capita. But this is because good policy leads to higher growth. None of the recent aid effectiveness studies question the importance of good policy. Yet, what is stressed in many of the papers challenging the Burnside and Dollar result is that aid effectiveness must be evaluated *after* we have conditioned on good policy. Once we condition on policy in the regressions we find that aid spurs growth regardless of the policy environment.

NOTES

1. See Lensink and White [*2000*] for a critique of the calculation of poverty efficient aid allocations and McGillivray and Morrissey [*2000*] for a critique of the fungibility discussion in Assessing Aid.
2. Other authors, for example, Lensink and White (this collection) have modified the neoclassical model with productive public expenditure, in other ways than Boone, leading to new results.
3. See Durlauf and Quah [*1998*] for an extensive list of other variables that have appeared in recent growth regressions.
4. We have normalised the scale parameter in the Cobb-Douglas function to unity: $F(K, L) = K^{\alpha} L^{1-\alpha}$.
5. As the economy is closed total wealth equals the capital stock, *K(t)*. For this reason we have replaced the standard no-Ponzi-game condition with a non-negativity constraint.
6. A technical appendix with all derivations is available from the authors on request.
7. The growth maximising level is given as the solution to

$$p(\tau, c^*(\tau)) = (1-\tau)(p_\tau + p_c \cdot \tfrac{\partial c}{\partial \tau}) \quad .$$

8. All regression diagnostics presented in this section are standard output in statistical programs such as Stata [*StataCorp, 1999*].
9. See Lensink and White [*1999*] for a critical assessment of the policy index and Rodríguez

and Rodrik [*1999*] for an illuminating discussion of the openness index and the relation between trade policies and growth in general.

10. The scaled change in the coefficient reported in Figure 4 is often denoted DFBETAS, see Belsley *et al.* [*1980*] and the Appendix for the precise definitions of the various influence measures used in this section.
11. The same result can be obtained in a slightly more sophisticated way, simply by excluding the two countries Gambia and Nigeria from the sample.
12. The robust regression method is iterative re-weighted least squares using Huber- and bi-weights. The procedure is standard in Stata.

REFERENCES

Adelman, I. and C.T. Morris, 1967, *Society, Politics and Economic Development*, Baltimore, MD: Johns Hopkins University Press.

Alesina, A. and R. Perotti, 1996, 'Income Distribution, Political Instability, and Investment', *European Economic Review*, Vol.40, pp.1203–28.

Alesina, A. and D. Rodrik, 1994, 'Distributive Politics and Economic Growth', *Quarterly Journal of Economics*, Vol.108, pp.465–90.

Barro, R.J., 1990, 'Government Spending in a Simple Model of Endogenous Growth', *Journal of Political Economy*, Vol.95, No.5, pp.S103–25.

Barro, R.J. and X. Sala-i-Martin, 1995, *Economic Growth*, New York: McGraw-Hill.

Belsley, D.A., Kuh, E. and R.E. Welsch, 1980, *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*, New York: John Wiley & Sons.

Beynon, J., 1999, '"Assessing Aid" and the Collier/Dollar Poverty Efficient Aid Allocation: A Critique'. DFID, processed.

Boone, P., 1994, 'The Impact of Foreign Aid on Savings and Growth', mimeo, London School of Economics.

Boone, P., 1996, 'Politics and the Effectiveness of Foreign Aid', *European Economic Review*, Vol.40, pp.289–329.

Burnside, C. and D. Dollar, 1997, 'Aid, Policies and Growth', Policy Research Working Paper 1777, The World Bank, Development Research Group, Washington, DC.

Burnside, C. and D. Dollar, 2000, 'Aid, Policies, and Growth', *American Economic Review*, Vol.90, pp.847–68.

Chang, C.C., Fernandez-Arias, E. and L. Servén, 1998, 'Measuring Aid Flows: A New Approach', Working paper, The World Bank, processed.

Chatterjee, S. and A.S. Hadi, 1986, 'Influential Observations, High Leverage Points, and Outliers in Linear Regression', *Statistical Science*, Vol.1, pp.379–416.

Chenery, H.B. and A.M. Strout, 1966, 'Foreign Assistance and Economic Development', *American Economic Review*, Vol.56, pp.679–733.

Cook, R.D. and S. Weisberg, 1982, *Residuals and Influence in Regression*, New York: Chapman & Hall.

Durbarry, R., Gemmell, N. and D. Greenaway, 1998, 'New Evidence on the Impact of Foreign Aid on Economic Growth', CREDIT Research Paper 98/8, University of Nottingham.

Durlauf, S.N. and D.T. Quah, 1998, 'The New Empirics of Economic Growth', NBER Working Paper 6422, National Bureau of Economic Research, Cambridge, MA.

Hadjimichael, M.T., Ghura, D., Mühleisen, M., Nord, R. and E.M. Uçer, 1995, 'Sub-Saharan Africa: Growth, Savings, and Investment, 1986–93', Occasional Paper 118, International Monetary Fund.

Hansen, H. and F. Tarp, 2000, 'Aid Effectiveness Disputed', *Journal of International Development*, Vol.12, pp.375–98.

Hansen, H. and F. Tarp, 2001, 'Aid and Growth Regressions', *Journal of Development Economics*, Vol.64, pp.547–70.

Knack, S. and P. Keefer, 1995, 'Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures', *Economics and Politics*, Vol.7, No.3, pp.207–27.

Lensink, R. and H. White, 1999, 'Assessing Aid: A Manifesto for Aid in the 21st Century? Global Evaluation of Swedish Programme Aid', Working Paper No.15, Sida.

Lensink, R. and H. White, 2000, 'Aid Allocation, Poverty Reduction and the Assessing Aid Report', *Journal of International Development*, Vol.12, pp.399–412.

McGillivray, M. and O. Morrissey, 2000, 'Aid Fungibility in Assessing Aid: Red Herring or True Concern?' *Journal of International Development*, Vol.12, pp.413–28.

Mosley, P., 1980, 'Aid, Savings and Growth Revisited', *Oxford Bulletin of Economics and Statistics*, Vol.42, pp.79–95.

Obstfeld, M., 1999, 'Foreign Resource Inflows, Saving, and Growth', in K. Schmidt-Hebbel and L. Servén (eds.), *The Economics of Saving and Growth*, Cambridge: Cambridge University Press, pp.107–46.

Papanek, G.F., 1972, 'The Effect of Aid and Other Resource Transfers on Savings and Growth in Less Developed Countries', *Economic Journal*, No.82, pp.935–50.

Papanek, G.F., 1973, 'Aid, Foreign Private Investment, Savings, and Growth in Less Developed Countries', *Journal of Political Economy*, No.81, pp.120–30.

Rodríguez, F. and D. Rodrik, 1999, 'Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence', Cambridge, MA: National Bureau of Economic Research.

Sachs, J.D. and A.M. Warner, 1995, 'Economic Reform and the Process of Global Integration', *Brooking Papers on Economic Activity*, Vol.1, pp.1–118.

StataCorp, 1999, *Stata Statistical Software: Release 6.0*, College Station, TX: Stata Corporation.

Summers, R. and A. Heston, 1991, 'The Penn World Tables (Mark 5): An Expanded Set of Internatinal Comparisons, 1950–1988', *Quarterly Journal of Economics*, Vol.106, pp.327–36.

Temple, J. and P.A. Johnson, 1998, 'Social Capability and Economic Growth', *Quarterly Journal of Economics*, No.113, pp.965–90.

Velleman, P.F. and R.E. Welsch, 1981, 'Effcient Computing of Regression Diagnostics', *American Statistician*, Vol.35, pp.234–42.

World Bank, 1998, *Assessing Aid: What Works, What Doesn't, and Why*, World Bank Policy Research Report, Oxford: Oxford University Press.

APPENDIX

In this appendix we list the formulas for the influence statistics used in the main text. The classical reference in econometrics is Belsley *et al.* [*1980*], but see also Cook and Weisberg [*1982*] for a likelihood oriented approach, and Chatterjee and Hadi [*1986*] (with comments) for comparisons of different measures and an enlightening discussion.

We consider the linear regression model

$$y = Xb + e$$

where y is a n-dimensional column vector, X is the $n \times k$ matrix of explanatory variables, b is the estimated k-dimensional coefficient vector, and e is the vector of residuals. The objective is to look at the effect on various quantities of omitting a single row of observations from the regression.

The leverage measure h_i is given as the diagonal elements of the least squares projection matrix. It can be given as

$$h_i = x_i (X'X)^{-1} x_i' = \tilde{x}_i (\tilde{X}'\tilde{X})^{-1} \tilde{x}_i' + \tfrac{1}{n}, \tag{7}$$

where x_i is the i'th row of the matrix of regressors and a tilde denotes centered variables. Following Belsley *et al.* [*1980*] an observation is termed a leverage point if h_i exceeds $2k/n$. However as k is small compared to n in this study we will use the less stringent value $3k/n$ as suggested by Velleman and Welsch [*1981*].

The residuals can be scaled in several ways. In this article we make use of the studentised residuals defined as

$$r_i = \frac{e_i}{s(i)\sqrt{1-h_i}}, \tag{8}$$

where $s(i)$ is the root mean square error based on a regression in which the i'th row is omitted and

h_i is defined in (7). The studentised residuals may be compared to a $t(n-k-1)$ distributed random variable.

The influence on the individual estimated coefficients of omitting the *i*th observation is calculated as the scaled change in the parameter estimate

$$\text{DFBETAS}_{ij} = \frac{b_j - b_j(i)}{s(i)\sqrt{(X'X)^{-1}_{jj}}}, \tag{9}$$

where $b_j(i)$ is the estimated coefficient based on a regression in which the *i*'th row is omitted and $s(i)$ is defined above.

One advantage of using the scaled measure of change is that these are comparable across coefficients. The unit of measurement is (approximate) standard errors of the estimated parameters. Belsley *et al.* [*1980*] suggest to use $\pm 2/\sqrt{n}$ as cut-off values for influential observations but less stringent values (unity) are often used because it is a *t*-like diagnostic (see Velleman and Welsch [*1981*]).

The overall influence on all parameters can be measured by the scaled change in the fitted value

$$\text{DFITS}_i = \frac{\hat{y}_i - \hat{y}_i(i)}{s(i)\sqrt{h_i}}. \tag{10}$$

As $s(i)\sqrt{h_i}$ is the root mean square error of the prediction, this measure is also in terms of standard errors and $\pm 2\sqrt{k/n}$ are typically considered as reasonable cut-off values for influential observations. Again there is an alternative suggestion to look at 'one standard error' changes.

It is important to be aware of the limitations of these influence measures and of the dangers of a mechanical usage. As noted by Belsley *et al.* [*1980: 15*]:

> A word of warning is in order here, for it is obvious that there is room for misuse of the above procedures. High-influence data points could conceivably be removed solely to effect a desired change in a particular estimated coefficient, its *t*-value, or some other regression output. While this danger surely exists, it is an unavoidable consequence of a procedure that successfully highlights such points. It should be obvious that an influential point is legitimately deleted altogether only if, once identified, it can be shown to be uncorrectably in error. Often no action is warranted, and when it is, the appropriate action is usually more subtle than simple deletion.