

Why do Economists Disagree so much on Aid Effectiveness?

Aid Works (in mysterious ways)

Oliver Morrissey

CREDIT and School of Economics, University of Nottingham

Initial Draft paper for presentation at the *IMF-CFD Conference on Financing for Development*, Geneva 15-17 April 2015

Abstract

Although the issue of aid effectiveness, the impact of aid on growth, has attracted considerable research attention there is no consensus view. It is not unusual for different researchers using the same data and econometric approach to arrive at dramatically different conclusions by altering a specific modeling choice. The economics literature is characterized by disagreement because econometric results are not robust. This paper argues that three specific limitations of cross-country growth regressions in the aid context explain the inconclusive nature of the literature: measuring aid so as to capture the ways in which it can affect growth; challenges in addressing simultaneity and endogeneity; and heterogeneity characterizes the data. These issues are discussed for studies of aid and growth, and for studies of aid and taxation, with a focus on heterogeneity – the delivery and effects of aid vary significantly across countries. The conclusion discusses implications for aid policy.

1 Introduction

Aid effectiveness, interpreted as the effect of aid on growth, has attracted considerable research attention that appears to have generated more heat than light. There is no consensus view: it is not unusual for different researchers using essentially the same data and econometric approach to arrive at dramatically different conclusions by altering a specific modeling choice (typically in a context where there is no evident reason to prefer one modeling choice over the other). The economics literature on aid effectiveness is characterized by disagreement, not simply because econometric results are not robust but because often they are conflicting. This is at least unfortunate because aid is an important issue politically and in terms of economic research. The current state of the literature is unsatisfactory because it is not capable of shifting priors: those who believe aid is ineffective can cite studies in support of that view while those who want to claim that aid is effective can also point to supporting evidence.

This paper argues that three particular features of how aid effectiveness has been addressed help to explain why cross-country growth regressions are inconclusive. The first is that it is not clear how to measure aid to capture how it may affect growth. The amount donors record as allocated to a recipient is greater than the amount the recipient records as received (especially for the amount delivered through the government). The amount itself may not be the most important factor as how the aid is implemented and associated policy reforms (conditionality and the donor-recipient relationship) are significant determinants of the effects of aid. The second is inherent endogeneity: economic performance influences aid received, aid influences performance and any (two-way) effects are likely to be persistent; identification is extremely difficult if not strictly impossible. The countries that are major recipients are by definition poor and have low growth for a variety of reasons that are inter-related with the impact of aid and donor-related policy reforms. The third is that heterogeneity is fundamental and pervasive; the recipient environments into which aid is injected are very different and multiple political and economic, internal and external, factors are interacting in complex dynamic ways. The way in which aid is implemented also varies considerably as donors differ in the types of aid and procedures employed (and recipients receive many forms of aid from numerous donors).

The interest here is not in differences in econometric results per se, nor in the tendency for alternative approaches to addressing an empirical question to yield contrasting results. Rather, the issue of concern is that comparable studies (using the same or very similar

samples, specifications and methods) can yield qualitatively and quantitatively different conclusions. In the aid effectiveness literature, the basic rationale for cross-country regressions, that of identifying broad patterns or effects that appear to hold on average, has not been achieved. Although the simple question of what is the effect of aid on growth on average is valid (and we may like to have an answer), given the three problematic features identified above cross-country growth regressions have not yielded a meaningful average effect. This does not mean that research, even cross-country regressions, can't be informative but does imply a need for caution in drawing inferences.

This paper elaborates on the three problematic features with a focus on selected studies chosen to highlight issues of concern; there is no claim to provide a comprehensive review of a large literature (on the contrary, we are intentionally selective of opposing studies). Section 2 considers the cross-country aid effectiveness literature by focusing primarily on three pairs of papers that highlight the sharpest disagreements. Section 3 shows that similar issues arise in examining the relationship between aid and tax revenue, as one dimension of the fiscal effects of aid that aim to capture the effects of aid on government fiscal behavior (chosen because cross-country regressions have been used). Section 4 concludes by discussing the relevance of aid modalities, whether the way in which aid is delivered (such as budget support or donor projects) relates to how aid can affect growth.

2 Aid and Growth: The Nature of the Disagreements

The empirical cross-country evidence is based on a rather strict criterion of aid effectiveness, namely that the amount of aid received has an effect on economic growth within a relatively short period (typically about five years). This is problematic for a number of reasons. First, it is difficult to identify the determinants of growth; many variables could potentially affect growth and the set of relevant variables will differ across countries and within a country over time. Although it is not essential to include all potential determinants it is necessary to consider important factors that may be related to the impact of aid to ensure that correct inferences are drawn for the conditional effectiveness of aid. Second, as more aid tends to go to poor countries that suffer growth-retarding characteristics there are simultaneity and endogeneity concerns. Third, not all aid is spent in a way that would be expected to have a medium-term observable impact on growth. Aid that finances physical capital investment should promote growth, depending on the productivity of investment (itself determined by the recipient environment), in the medium term. Aid for human capital, such as health and

education services, or technical assistance would only affect growth in the long-term. As no more than a third of aid finances physical investment, the measure of aid used in most studies over-states the volume of aid that could influence growth in the medium term. Studies that attempt to construct a measure capturing aid for investment tend to find an effect on growth (Gomanee *et al*, 2005; Clemens *et al*, 2012). As the policies associated with aid (conditionality and the donor-recipient relationship) affect the recipient environment, the amount of aid alone is a weak proxy for the potential of aid to affect growth. These observations taken together imply that cross-country regressions are biased against finding that aid is effective (as defined), i.e. a reasonable prior (or null hypothesis) is that the coefficient on aid is zero.

This section is not an attempt to review the large literature on aid effectiveness. Rather, we focus on a few studies, in particular pairs that reach conflicting results with similar approaches and data, to highlight problems inherent in cross-country regressions. Burnside & Dollar (2000) is rightly considered the core paper initiating the recent literature with the principal conclusion that aid was only effective in recipients with good policies as captured in an index combining three policy indicators – the Sachs-Warner (SW) measure of openness, the budget surplus (deficit) as a share of GDP and inflation. An important implicit assumption was that aid (or donors) did not affect the three policy indicators. This is likely to be true of SW up to about 1990, but not thereafter. As the majority of recipients followed donor advice in implementing trade, exchange rate and economic liberalization since the 1980s, there is a donor policy effect on the openness measure (tending to reduce tariffs, black market premia and state control of tradable sectors). The other two indicators (which capture outcomes rather than policy inputs) are likely to be affected by aid and by donors. To the extent that aid finances government spending and donor advice supported improved fiscal management (see Section 3), more aid will tend to be associated with lower deficits net of aid.¹ As aid reduces the need for seignorage and donor advice supported improved macroeconomic management, more aid will tend to be associated with lower inflation *ceteris paribus*. Thus, all three components of the policy index associated with growth can be improved by aid. It is difficult to conclude that Burnside & Dollar (2000) actually showed that aid is not effective (which is distinct from not finding any evidence that aid is effective).

Hansen & Tarp (2001) was the first paper to demonstrate that even using essentially the same data (they had difficulty reconstructing exactly the same data) and specification, the

¹ This is not the precise measure used; although recognizing that (some) aid is included in revenue they assume incorrectly that ‘aid-financed projects [are] included in expenditures, so that there is no necessary relationship between aid and this measure’ (Burnside & Dollar 2000: 849-50). As discussed in Morrissey (2015), aid is incompletely captured in revenue and expenditure so this may not be correct.

results were not robust to apparently innocuous adjustments, such as the treatment of outliers; to including squared aid and policy terms; and to using lagged values instead of a first stage instrumental variables (IV) regression to address endogeneity. The interesting feature was not that they found the results of Burnside & Dollar (2000) not to be robust but that the various changes always seemed to yield evidence of aid effectiveness in Hansen & Tarp (2001), i.e. the two studies were comparable in approach with the same sample but conflicting in results.

Interest in the coefficient on an aid-squared term also arose at this time. This was interpreted as capturing diminishing marginal returns to aid, albeit without any agreement (within remarkably wide ranges) as to where this may set in. The focus on aid instability in Lensink and Morrissey (2000) offers an alternative interpretation as aid volatility tends to have a negative impact on growth; beneficial effects of (public) investment are captured by aid whereas negative effects of volatility are captured by aid². The adverse effects may be because unpredictable aid flows undermine the ability of the government to manage the budget (and aid within the budget), leading to fiscal instability (and lower public investment). It may also arise if external shocks elicit an increase in aid, so adverse growth episodes are associated with unanticipated increases (and subsequent reductions – reversion to trend) in aid, captured by greater volatility.

Many subsequent studies showed that results tend not to be robust to changes in the sample or specification (Roodman, 2007). Studies following the specification and IV strategy of Burnside & Dollar (2000) were less likely to find evidence of aid effectiveness, whereas those altering the specification (especially with alternative interaction terms) and/or using lagged values (or GMM) rather than an IV strategy were more likely to find aid to be effective. The way in which endogeneity was addressed does appear to be very important. Rajan and Subramanian (2008) posit that aid impacts growth by adding to (physical) capital (they assume no effect on productivity) and recognize that not all aid goes to investment so the anticipated effect if there is one will be modest. Their major innovation is the IV strategy: rather than dynamic panel methods (given persistence and weak instruments problems) they use an estimated supply of aid from donors to instrument for aid in a regression on ‘long-run’ average growth. For alternative periods and specifications (including a dynamic panel approach) they find no evidence that aid has an effect on growth. Arndt *et al* (2010) argue that some supply-side variables in the IV regression may be correlated with growth, or with omitted variables that are correlated with growth, and experiment with alternative instruments (but in the spirit of the same IV strategy); colonizer dummies transpire to be poor instruments (colonial legacy can have effects on subsequent growth). Arndt *et al* (2010) apply their IV

regression to estimate the regressions of Rajan and Subramanian (2008) and find that aid is effective with a significant coefficient of about 0.2 in most regressions (interestingly, this is often the value in studies that find a positive effect of aid), although their preferred estimate is 0.1 (which happens to be the null hypothesis suggested by Rajan and Subramanian (2008)).

There is no single correct method to account for endogeneity: methods based on lagged values are subject to problems of persistence and validity of internal instruments, whilst it is very difficult to find truly exogenous valid external instruments (and results are sensitive to the instrument set). There is no consensus that aid affects growth, or that aid does not affect growth. Although cross-country regressions have failed to yield a consensus average effect, might there be a meaningful average of averages, i.e. a preponderance of a particular effect across the many studies? Doucouliagos & Paldam (2008) argue that meta-regression analysis (MRA) is an appropriate method to summarize the results from the many existing studies of aid effectiveness, and conclude that the evidence for aid having an impact on growth is absent. Mekasha & Tarp (2013) apply the same MRA methods to the same set of studies and find that aid is effective. In this case the reversal of the result is due to a particular modelling choice, between random effects and fixed effects. This transpires to be very important because it encapsulates priors about heterogeneity. If, in line with Doucouliagos & Paldam (2008), one assumes there is one 'true' coefficient for the effect of aid on growth and that all studies share a specification that captures this one coefficient, a fixed effects model is appropriate. If one accepts the argument of Mekasha & Tarp (2013) that the actual effect of aid on growth can be heterogeneous across studies a random effects model is appropriate to allow for differences between studies. As samples and specifications vary, often considerably, across the almost 70 studies included in the MRA the random effects approach seems sensible, or at the very least equally valid. As the two approaches yield diametrically opposed estimates, even the statistical summary technique is inconclusive.

Mekasha & Tarp (2013) also note that MRA results are sensitive to how one treats studies with interaction terms. Fielding and Knowles (2011) show that great caution should be deployed in interpreting interaction terms in this context, what they term conditional aid effectiveness, and that the signs can be reversed but significance maintained by replacing aid/GDP with aid per capita. Heterogeneity arises again as although either aid/GDP or aid per capita can be incorporated into a theoretical growth model (albeit with different implications for interpretation) the cross recipient distribution of the two aid measures will be quite different. For example, small (island) middle-income countries tend to have high aid per capita but relatively low aid/GDP, so any estimated effect of aid on growth conditional on a

chosen interaction term (such as policy) may be quite different. The measure of aid matters but, more importantly, interaction terms exacerbate heterogeneity; intuitively this is because they are attempts to condition on factors that mediate the effect of aid.

Herzer & Morrissey (2013) demonstrate the prevalence of heterogeneity in a different way. They posit that effect of aid on GDP depends on a country specific combination of two effects. The first is the direct effect through financing investment, which is expected to be positive although it may not be significant and may be modest as it depends on the productivity of investment. The aid effectiveness literature essentially only considers this direct effect. The second is an indirect effect through aggregate productivity; the argument is that aid is mediated by, may affect and interacts with various factors that represent the recipient's economic environment, such as determinants of the productivity of investment. The indirect effect can be negative if aid is given to countries that exhibit severe growth-retarding factors, such as poor governance or macroeconomic instability (that may reflect poor policy). The indirect effect can be positive if the economic environment is relatively good. There is no presumption that aid affects the environment, but policies associated with aid/donors can do so (the estimation strategy does not capture this); rather, the presumption is that the environment affects aid.

Herzer & Morrissey (2013) use data for 59 developing countries over 1971-2003 to derive country specific time series estimates of the long run association between aid and output (the indirect effect), and between investment and output (the direct effect). The aid-output coefficients are, on average, negative: for 25 countries the negative coefficient is significant; in about a third of cases the coefficient is positive (but only significant in nine cases); and for 25 countries the estimate is insignificant. The aid-output coefficients are generally smaller than the positive investment-output coefficients (positive in 44 cases, 26 of which are significant, and only negative and significant in four countries). Insofar as aid is used to finance investment, the overall effect on output is likely to be positive. The results may not be robust (many country estimates are insignificant) but the purpose is to demonstrate differences in the country-specific environment that influences aid effectiveness. Country differences in factors influencing aggregate productivity influence the effect of aid in the country; of 20 variables considered, the most important are found to be law and order (for quality of institutions), religious tensions (political stability) and government size.

The basic argument is that endogeneity and heterogeneity combine to seriously limit how informative cross-country regressions can be because there are multiple factors that

affect and may be affected by growth and by aid in different ways in different countries. Consider corruption, especially where the measure is based on survey responses by private sector investors, to illustrate the argument. Corruption would generally be accepted as a good indicator of the recipient economic environment such that higher corruption is expected to be associated with lower productivity of investment, hence lower aid effectiveness via investment (and perhaps also via government spending). However, whether aid affects corruption is a separate question, and one that is not easy to address. As availability of money is likely to attract corrupt activity, one may posit an association between aid and corruption (separate from the tendency of poor countries to attract more aid and have higher corruption). However, it is reasonable to posit that corrupt individuals are most concerned with easy money, i.e. less transparent transactions are more attractive because the corrupt wish to minimise the possibility of being observed. This suggests that they would rank the attractiveness of money flows according to transparency, suggesting a ranking of private capital (including FDI), government spending (and procurement) and finally aid. On this basis aid is the least attractive simply because it is subject to greater monitoring by donors (the only claim is less, not no, corruption); corruption involving donor funds is more likely to be observed and investigated. To the extent that donor technical support and institutional interventions have improved public sector management (of tax and spending), public financial flows are likely to be monitored more openly than private flows. Aid may be subject to corruption but less so than private capital flows that are less transparent. To the extent that donors are effective in improving the monitoring of aid and public expenditure, more of the corrupt activities may be diverted to private flows. If corruption is measured by surveying private investors it is possible that reductions in corruption due to donors are associated with an increase in measured corruption (pertaining to private flows). Cross-country regressions of effects of aid on corruption or including aid-corruption interaction are likely to be misleading or uninformative.

2 Aid, Government and the Public Sector

Arguments that aid is not, or has not been, effective in contributing to economic growth often locate the problem in giving aid to governments that do not make the best use of the aid. Most aid either goes to the government to finance public spending or supports the provision of public goods and services (through donor projects), so may encourage too large a public sector, rather than promoting the private sector (or effectively ensuring that the government

supports private sector development). This effect may be exacerbated to the extent that aid alleviates the need for government to impose fiscal and budgetary discipline: the availability of aid may encourage corruption (but see above) and allows governments to continue with 'bad' or inferior policies. There is often a belief that governments misuse aid because it is fungible (it may not be used for the purposes intended by donors). Underpinning these arguments is a belief that conditionality does not work, i.e. donor/aid leverage does not ensure that governments implement good policies, or even the policies that donors promote. Conditionality failure is a particular concern to those who accept the Burnside & Dollar (2000) argument that aid makes a positive contribution to growth only in those countries with high values for the policy indicator. In this case, if aid/donors can't improve policy, aid can't improve growth.

The strength of this critique is moderated to the extent that aid is effective independently of policy (which some studies find) and/or that conditionality influences policy. The latter seems quite likely as most developing countries have, to a greater or lesser extent, implemented the types of policy reforms supported by donors over the last 25 years. It is appropriate to take a long-term view because sustained policy reform is a slow process: the more (politically) challenging the reform the longer it is likely to take. For this reason, and because external shocks undermine the implementation and effect of policy reform, it can be very difficult to identify an effect of reform on performance. Indeed, at least in a cross-country regression context, it may be futile to attempt to quantify effects of conditionality on reform and performance. Nevertheless, taking Africa as an example, it is evident that in most countries trade and macroeconomic policy is better now than in the 1980s, and much of this attributable to donors working with recipients.

While the argument that the allocation of aid to government is the source of ineffectiveness has some historical merit it is not supported by evidence. Donor influence and aid conditions do influence policy. Since the 1980s, for example, aid recipients have implemented policies liberalising exchange rates, reducing tax distortions (especially tariffs), removing price controls in agriculture, and establishing more stable macroeconomic policy. These have improved the quality of the economic environment so that aid financed investment can be more productive, noting that in the poorest countries public investment is necessary to establish the physical and human capital infrastructure to crowd in private investment.

Even if aid is fungible the evidence suggests it has supported increases in spending on the provision of social (public) services, and increasing social sector spending is a clear intention of donors (and perhaps the most important reason why less than half of aid is allocated to capital investment). Aid recipients face a situation where domestic revenue is too low to meet expenditure needs to provide the level of public goods required for growth and development. The low domestic revenue does not imply that tax effort is weak, i.e. that tax/GDP ratios are particularly low; given the tax base and structure of such economies domestic revenue mobilisation may be as high as is feasible or even desirable, given that poor countries tend to rely on highly distorting taxes on trade. The evidence on tax reform is promising: significant reforms to tax structure have been implemented in many countries, increasing efficiency (of the tax system and collection) and reducing distortions, especially given the reduced dependence on taxes on trade.

The possibility that aid is fungible relates to two issues: is the aid all allocated to the particular area of expenditure targeted by donors and does expenditure on that area increase by the amount of the aid (is the aid additional). Even if all the aid is allocated as intended, domestic resources previously allocated may be redirected so aid is not fully additional. The general argument is that recipients divert aid to government consumption spending rather than using it to finance growth-promoting investment. The evidence for this is weak as consumption (recurrent) spending includes wages and equipment, for example for health and education (necessary for human capital), and costs of maintaining infrastructure, so aid is often intended to support recurrent spending and such spending can contribute to growth.

The counter argument is that if one analyses the dynamics of expenditure within the context of the evolution of fiscal aggregates (including taxes and borrowing), it is apparent that over time spending increases in the areas targeted by donors and often total spending increases by more than the value of aid (McGillivray and Morrissey, 2004). Given the lack of good data over time for many aid recipients on government spending and its composition, there are few major cross-country regression studies of the impact of aid on spending, and no replications so too few studies to consider disagreements. There are a number of studies addressing if aid is fungible; here the evidence is mixed and inconclusive, largely resolving around how off-budget aid is treated (Morrissey, 2015). There are also a number of country time series of the fiscal effects of aid: these demonstrate heterogeneity in the sense that quantitative results differ across countries, although the general pattern is that aid increases aggregate government spending while poor countries have limited ability to alter tax revenue in the short to medium term (Morrissey, 2015).

There is disagreement in cross-country regressions on the effect of aid on tax effort (measured as the tax/GDP ratio). Gupta et al (2004) argue that aid in the form of grants can be expected to reduce tax effort because recipients get ‘free’ money (no obligation to repay in the future) and so have less incentive to exert effort to collect taxes. Aid loans, in contrast, encourage tax effort because recipients know they have to be repaid. They find a negative coefficient on aid grants, but a positive coefficient on loans, in a tax revenue regression and infer this shows that grants reduce tax effort. Benedek et al (2012) confirm these results using updated data. Clist and Morrissey (2011) replicate and extend the analysis of Gupta et al (2004) and show that the results are not robust; indeed, considering the period since the mid-1980s they obtain results that suggest a positive effect of aid on tax revenue. Carter (2013) shows that the results in Benedek et al (2012) are sensitive to the econometric method used and are not robust. Morrissey and Torrance (2015) find no robust relationship between aid and tax revenue, whether using total aid or disaggregating into grants and loans. General claims that aid reduces or increases tax effort are not robust: there are associations between aid and tax in the data, largely due to structural characteristics of the economy whereby determinants of low tax ratios are associated with high aid receipts. Essentially, poor countries get more aid (in the form of grants) and have lower tax ratios; there is no justification for claims that aid has a causal effect on tax effort.

As in the previous section, specific econometric estimates are not of concern: the point is that results are sensitive to specification, estimation method, data and sample. The three basic reasons for inconclusive aid effectiveness results arise again: data problems, especially in measuring tax as distinct from non-tax revenues; endogeneity and heterogeneity. It is also the case that donor-supported policy reforms and technical assistance, such as in tax administration and public finance management, may have a greater effect on tax effort than the actual amount of aid.

Clist and Morrissey (2011) focus on the sensitivity of results to the way in which endogeneity is addressed (see also Carter, 2013). The amount of aid received and the composition between grants and loans will be affected by and affect the expenditure and domestic financing decisions of recipients. An explicit intention of aid is to finance the provision of public goods and services, recognizing that domestic revenue in poor countries is not sufficient to finance the levels of social spending and public investment required for development. There is a natural negative correlation (simultaneity) between aid and tax: poor countries that receive more aid exhibit characteristics such as large agricultural and informal sectors associated with a low tax base, hence low tax ratios; aid flows are typically higher to

countries that have more difficulty in raising domestic revenues because such countries tend to be poorer. Clist and Morrissey (2011) argue that long lags for aid in the revenue regression are required to distinguish behaviour due to aid from the natural simultaneity. While the use of lags alone does not identify the effect of aid they find that the significance of the aid coefficient disappears with longer lags.

Morrissey and Torrance (2015), employing a new source of annual tax revenue data (the Government Revenue Dataset from the International Centre for Taxation and Development), replicate the results of Gupta et al (2004) for pooled OLS. However, the results are not robust for other estimators or including alternative control variables; generally the coefficients on aid variables become insignificant (and when significant are more often positive). The coefficient estimates on aid variables in cross-country tax regressions are no more robust than for growth regressions, and there is no convincing evidence that aid has a causal effect on tax effort.

Heterogeneity is a major reason for the absence of robust results. In simple terms, the tax/GDP ratio is determined by the tax rates applied to components of the tax base, given the effectiveness of tax administration and collection rates. Tax ratio equations use very poor proxies for the tax base, such as the share of trade, agriculture and industry in the economy. More importantly, they rarely include tax rates or capture tax policy reforms. For example, two countries with the same imports/GDP ratio may collect very different amounts of tax from imports, because of differences in tariff rates, exemptions, evasion and policies (one may have many preferential trade agreements with lower tariffs). An inherent limitation of cross-country tax ratio regressions is that the important country-level determinants of revenue are not captured by available data. The results are indicative on how tax ratios vary broadly according to changes in economic structure and income levels. However, it is not obvious that aid is even relevant in such a broad structural context (the reduced form approach does not allow for ways in which aid can affect tax bases or rates or collection efficiency).

The effect of aid on government fiscal behaviour (including taxation) depends more on technical assistance and policy reforms associated with aid than on the amount of aid. The typical argument for why the amount of aid may reduce tax effort is that because people do not like paying taxes, governments have little incentive to exert administrative and political effort in collecting unpopular taxes. This is one reason why poor countries have often relied on trade taxes, especially on imports. The border is one place where transactions are recorded so administratively is the easiest place to collect taxes. Whilst importers have incentives to evade tariffs, such taxes may not be unpopular: import-competing producers lobby for tariff

protection while the public can be persuaded (incorrectly) that foreigners rather than they pay the tariff. Since the 1980s trade liberalisation and especially tariff reductions has been a major donor-promoted policy (often an aid condition); in some cases increases in aid have been intended to compensate for the revenue loss of reducing tariffs (assumed to be temporary while other taxes are increased to substitute for tariffs, although in practice it can be difficult to make up the lost revenue). In this example, allowing for conditionality, one could observe aid increasing while tax revenue declines.

The argument that aid reduces tax effort assumes that governments are concerned about political costs of taxes, but do not off-set these against political costs of aid. This is not consistent with concerns expressed about aid dependence. Governments may be most amenable to aid when they view it as a resource to which they have (personal) access (arguments such as this lie at the root of theoretical papers on aid conditionality failure). Thus, when aid supports corrupt autocratic regimes one would not be surprised if tax collection was relatively low. This is a problem of the particular donor-recipient relationship rather than the effect of an amount of aid on tax effort (in another country the same amount of aid could be used effectively). As noted above, over at least the past 20 years donors have placed more emphasis on improving public finance management and encouraging anti-corruption measures. Even if these initiatives have had only limited success (Andrews, 2013), aid is now more transparent and monitored more carefully than was the case in the past. Governments (politicians and officials) are less likely now to view aid as an easily accessible resource. Furthermore, governments may face greater constraints on how they can use aid than on how they can misuse domestic revenue. Altincekic and Bearce (2014) argue that there is no political aid curse: because aid is subject to conditionality and less fungible than resource (oil) revenues governments are less able to use aid to fund repression or appeasement (and so more likely to use domestic resource revenues).

These changes provide a reason to believe that governments face a different type of trade-off between aid and tax because both involve political costs. A government that wants to be less aid dependent will choose more tax effort. This choice will depend on the nature of the aid relationship and the tax options. Countries with access to resource rents have less incentive to increase taxes whatever their willingness to accept aid. Aid is likely to imply the highest costs of accountability; more effort has to be made to account (to donors) for how aid is used than to account to taxpayers for how taxes are raised or spent. The latter will depend on the nature of the democracy, which is typically weak, at least for accountability, in poor countries. To the extent that resource revenues are the least transparent they incur the lowest

accountability costs. As donor agencies have to account to their own government and parliament, and even public, donors place increasing emphasis on measures to monitor the use of aid, minimise fungible use and tackle corruption. The policy reform conditions and technical assistance associated with aid include public financial management and tax administration. Whilst in principle this can improve accountability to taxpayers it also increases the political costs of aid as some control and influence is ceded to donors.

The requirements of accounting to donors and negotiating on conditionality reduce the autonomy of aid recipients by limiting their policy discretion; even limited conditionality is a constraint on policy action (if only because effort has to be expended to avoid or circumvent the conditions). Governments are not likely to want to be dependent on aid if this implies ceding policy influence to donors; they would prefer greater autonomy (being seen as subject to foreign influences can undermine their domestic popularity). A desire for greater autonomy would encourage governments to increase tax effort to reduce aid dependence. Increasing taxes may impose a cost of being more accountable to domestic constituencies but this is unlikely to offset the benefits of autonomy. The accountability costs of aid are likely to be higher than for taxation.

In addition to the political costs of accountability and limited autonomy, one should also consider bureaucratic costs of tax administration. These are likely to have been reduced as many low income countries implemented tax administration and fiscal reforms (Moore, 2014) over the past two decades, often with donor support. To the extent that reforms improve the efficiency of tax collection, revenue can be increased without increasing tax rates. There are also bureaucratic costs of aid, such as the costs of numerous officials meeting with and reporting to donors. As recipients face many donors with more and changing requirements on monitoring aid delivered in different ways for multiple purposes, aid costs remain high. To the extent that donors coordinate, for example by meeting with officials as a donor group, the costs may not be increasing, but donor proliferation suggests they are hardly decreasing significantly either. The bureaucratic costs of aid are high while those of taxes are declining; whether the balance has shifted in favour of taxes will vary from recipient to recipient.

It should come as no surprise that the cross-country regression evidence for the effect of aid on tax ratios is weak and inconclusive. Adequate data are not available to properly specify the determinants of tax effort for a cross-section of countries over time so estimates are based on weak proxy variables for tax bases. The specifications are unable to adequately account for policy changes, whether in tax rates or administration (that may be influenced by

donors). Once endogeneity is allowed, a reasonable prior is a zero coefficient on aid (and that is most often the result). As in the case of cross-country growth regressions, heterogeneity is highly prevalent: when aid works or does not work, this happens in different ways in each recipient.

4 Conclusion: Aid Modalities and Effective Aid

This conclusion considers some implications for aid effectiveness research and aid policy. The broader political context of understanding the effects of aid should be recognized. Donor agencies want to demonstrate that aid, or at least their aid, is effective. They can point to studies finding a positive effect of aid as evidence, supplemented with case studies of good examples of impacts of specific projects or in particular countries. The agencies may operate in a domestic environment where there are opponents of aid, usually vocal politicians or even parties (in the UK, UKIP would like to cut most of the aid budget). These can point to studies finding that aid is ineffective to support their agenda, and it is relatively easy to get anecdotes of particular aid projects that failed. Researchers should aim to go beyond this and try to resolve what the evidence can tell us. One approach is to focus on particular types of interventions, which can be informative but not general enough to influence the broader political debate. This paper has addressed the disagreements at the broadest level – cross-country regressions that aim to find an ‘average’ effect on growth.

The principal conclusion is that any effect of aid on growth varies across countries and within countries over time. The determinants of the growth process differ across countries; it may be true that investment is the major growth driver, but productivity of investment varies greatly within and across countries, as does any way that aid-financed investment can affect growth. Furthermore, not all aid finances capital investment so the potential benefits of much aid, such as for social sector public goods (investment in human capital that only delivers long-run benefits), is missed by a focus on investment and growth. As a result of this fundamental heterogeneity cross-country regressions are not very informative, even if the data and endogeneity issues can be addressed (at some future time – so far it is hard to believe they have been adequately addressed). Similar arguments apply to attempts to find an average effect of aid on tax effort. There is no robust evidence that aid, in general or on average, has a consistent effect on growth or on tax revenue. In respect of the effect on growth we can perhaps be more confident that the association is positive, but one should not draw causal inferences (regarding tax it seems most appropriate to conclude there is no general effect).

Donor agencies are probably well aware of heterogeneity insofar as they know the nature of their relationship differs across recipients, and that the characteristics of recipients influence how aid is delivered. A good example of this is budget support, as donors are only likely to commit aid to the budget (as distinct from specific sectors or projects) if they are confident that the recipient manages the budget and expenditure reasonably well. Clist et al (2012) argue that donors will only give budget support if government spending allocation is aligned with the donor's objectives, otherwise they will retain control through donor projects (corruption would have a similar effect, reflecting donors concerns over fiduciary risk). They find that this does help explain which countries receive budget support from the World Bank or European Commission. Interestingly, whilst it explains eligibility it does not explain the amount of aid. In other words, donors may decide on how much aid to give a recipient based on need (such as high poverty or low domestic revenues) but will adjust how the aid is delivered (the aid modality) according to characteristics such as government effectiveness and corruption.

This suggests one reason why cross-country regressions using an aggregate aid measure yield inconclusive results: donors adapt aid modalities, but not the aid amount, to features of the recipient. In weak recipients (with poor governance, corruption, poor expenditure management), those with poor environments and low productivity of investment, donors will give less aid to the government and more through specific projects, so the link between aid and growth will be weakened. In stronger recipients with a good environment, a greater proportion of aid goes to the government; as these will tend to be better performing recipients anyway, aid is more likely to be associated with growth. When this is all combined in cross-country regressions the differences in how aid is delivered (correlated with underlying country performance) is mixed with recipient heterogeneity. In this context an interaction with a policy indicator may capture the conditional aid effectiveness but does not capture the potential effects of aid in the weak recipients (because this is sector or project specific and may not be observed in medium term growth).

The discussion of aid and taxation is useful to elaborate on this point as where recipient processes are weak the donors focus on strengthening those processes which (at best) will only impact on tax revenue slowly (and may not even impact on growth). If early support by donors improves budgetary and expenditure management methods then aid has beneficial effects, even if cross-country regressions fail to reveal these. There is evidence that aid is improving and becoming more effective, even if much more needs to be done to assist the poorest countries. The allocation of aid improves across countries as more goes to those

countries that are able to use it better (allowing for the fact that much should go to the weakest with the greatest need). Allocation has also improved within countries as aid projects or technical support help improve the economic environment (productivity of investment) and mechanisms for monitoring and accounting for aid and public expenditure.

Aid does not work in a simple manner; the limitations of work on the effect on growth or the narrower remit of the effect on tax effort demonstrate how little is known about how 'aid works on average'. One could conclude that, on average, aid is not effective but this may simply be asking the question in the wrong way (the right answer may be given to the wrong question, although some question if it even is the right answer in the particular context). Aid may have beneficial effects in all recipients but only some impact on growth in a detectable manner or growth effects are only realised over the long term so the average is completely uninformative. Alternatively, aid may have positive effects in some cases, negative in others (by prolonging the life of regimes imposing a very poor economic environment), and no effect in others so the average is zero. What seems more likely is that specific aid is beneficial in specific cases, some aid is ineffective, and most aid has effects that cannot easily be picked up by looking at growth outcomes. Aid may indeed work in mysterious ways.

References

- Altincekic, C. and D. Bearce (2014), Why there Should be No Political Foreign Aid Curse, *World Development*, **64**, 18-32.
- Andrews, M. (2013), *The Limits of Institutional Reform in Development*, Cambridge, UK: Cambridge University Press.
- Arndt, C., S. Jones, and F. Tarp (2010). 'Aid, Growth and Development: Have We Come Full Circle?', *Journal of Globalization and Development*, **1**(2), art 5.
- Benedek, D., E. Crivelli, S. Gupta and P. Muthooru (2012), Foreign Aid and Revenue: Still a Crowding Out Effect? IMF, Washington DC: *IMF Working Paper WWP/12/86*.
- Burnside, C. and D. Dollar (2000). 'Aid, Policies, and Growth.' *American Economic Review*, **90**, 847-868.
- Carter, P. (2013), Does Foreign Aid Displace Domestic Taxation? *Journal of Globalization and Development*, **4**:1-47.
- Clemens, M. A., Radelet, S., Bhavnani, R. R., & Bazzi, S. (2012), Counting chickens when they hatch: Timing and the effects of aid on growth, *The Economic Journal*, **122**, 590–617
- Clist, P., Isopi, A. & Morrissey, O. (2012). Selectivity on Aid Modality: Determinants of Budget Support from Multilateral Donors. *Review of International Organizations*, **7** , 267-284.
- Clist, P. and O. Morrissey (2011), Aid and Tax Revenue: Signs of a Positive Effect since the 1980s. *Journal of International Development* **23**: 165-180.

- Doucouliaqos, H. and M. Paldam (2008), Aid effectiveness on growth: A meta study. *European Journal of Political Economy* **24**, 1-24.
- Dalgard, C-J, H. Hansen and F. Tarp (2004), 'On the empirics of foreign aid and growth', *The Economic Journal*, **114** (496): F191-F216.
- Fielding, D. and S. Knowles (2011), Dangerous Interactions: Problems in Interpreting Tests of Conditional Aid Effectiveness, *The World Economy* (doi: 10.1111/j.1467-9701.2011.01363.x)
- Gomanee, K., S. Girma, and O. Morrissey (2005), 'Aid and Growth in sub-Saharan Africa: Accounting for Transmission Mechanisms', *Journal of International Development*, **17**, 1055-1076.
- Gupta, S, B Clemens, A Pivovarsky, and E Tiongson. (2004), Foreign Aid and Revenue Response: Does the Composition of Aid Matter? in S. Gupta, B. Clemens and G. Inchauste (eds) *Helping Countries Develop: The Role of Fiscal Policy*, Washington DC: IMF.
- Herzer, D. and O. Morrissey (2013), Aid and Domestic Output in the Long Run, *Review of World Economics*, **149**, 723-748.
- Lensink, R. and O. Morrissey (2000), 'Aid Instability as a Measure of Uncertainty and the Positive Impact of Aid on Growth', *Journal of Development Studies*, **36**, 31-49.
- McGillivray, M., and O. Morrissey (2004), 'Fiscal Effects of Aid' in T. Addison and A. Roe (eds), *Fiscal Policy for Development*, Basingstoke: Palgrave Macmillan/WIDER, pp. 72-96.
- Mekasha, T. and F. Tarp (2013), Aid and growth: What meta-analysis reveals. *Journal of Development Studies*, **49**, 564-583.
- Moore, M. (2014), Revenue Reform and Statebuilding in Anglophone Africa, *World Development*, **60**, 99-112.
- Morrissey, O. (2015), Aid and Government Fiscal Behaviour: What Does the Evidence Say?, *World Development*, **69**, 98-105.
- Morrissey, O. and S. Torrance (2015), Aid and Taxation, Chapter 38 in M. Avin and B. Lew (eds), *Handbook of Foreign Aid*, Edward Elgar (forthcoming)
- Rajan, R & A Subramanian (2008), Aid and growth: what does the cross-country evidence really show? *The Review of Economics and Statistics*, **90**, 643-665
- Roodman, D (2007), The Anarchy of Numbers: Aid, Development and Cross-Country Empirics, *World Bank Economic Review*, **21**, 255-277