17/9-2006

# Aid effectiveness on growth: A meta study

Hristos Doucouliagos, Department of Economics, Deakin University, Melbourne, Australia

Martin Paldam Department of Economics, University of Aarhus, Aarhus, Denmark

**Abstract:**

*The aid effectiveness literature contains about 100 papers that see aid as a treatment given to poor countries to generate development. 68 of these papers provide a total of 543 comparable estimates of the effect of aid on growth, which are the data of our meta-analysis. We consider two questions: (Q1) Are the estimates converging to a clear result over time as aid agencies gain experience, models become better and data accumulates? We find that the results do have a positive average, but it is small, insignificant and falling. (Q2) Can we identify the main factors that explain the large differences in the results? We find that much of the variation between studies can be attributed to publication outlet, institutional affiliation, data and specification differences. However, some of the difference between studies is real. In particular, the aid-growth effect is stronger for Asian countries. The meta-analysis indicates also the existence of indirect channels, which need to be further explored.*

JEL: B2, F35, O35

Keywords: Aid effectiveness, meta study, economic growth

**Acknowledgement:** Our cooperation was supported by the Aarhus University Research Foundation. Pia Wichmann Christensen has been a very competent research assistant. Tryggvi Thor Herbertsson, Peter Sandholt Jensen and Allan Würtz are thanked for discussions inspiring this paper. Useful comments were received from seminar participants at the University of Aarhus, Deakin University and the Kiel Institute of World Economics, as well as from the Public Choice Society 2006 Meetings in New Orleans and Turku.

## 1. Introduction: Two questions to a literature

At present the annual flow of development aid to poor countries amounts to more than US$ 60 billion, and the average recipient country receives about 7½% of its GDP in aid.[1] The basic (alleged) reason aid is given is to generate development, so that poverty can be reduced.

Thus, aid is a *treatment* given to poor countries in order to *generate development*. Many have asked the AE-question: *Is aid effective?* For almost 40 years economists have tried to answer the AE-question by estimating macro relations containing an aid effectiveness term, such as the aid-savings term, the aid-investment term and the aid-growth term. Our focus in this paper is on the aid-growth term, $\mu$.

The aid-growth term turns the AE-question into a question about a coefficient: Is $\mu$ positive and statistically significant? It has been known since the early 1970s, from models that concentrate on the aid-growth term, that aid shares and growth rates have a correlation of about zero.

Many researchers and policy makers have found such negative answers unacceptable and unreasonable, and this has generated the **AEL**, **A**id **E**ffectiveness **L**iterature, which may be seen as a quest to *overcome* the zero correlation result by adding control variables to the model, by using estimators correcting for various biases, etc. The present paper is a quantitative study of the aid-growth subset of this literature using the tools of meta-analysis.[2] The meta-analysis is used to analyze two questions:

**(Q1)** Has the AEL established that aid has an impact on economic growth? And, if it has, how large is the impact? This is analyzed as a question of *convergence*, where we study if the estimates of $\mu$ converge to a result, $\mu(N,t) \rightarrow \hat{\mu} \neq 0,$ when the number of observations, $N$, and time, $t$, goes up. (Q1) is analyzed in Section 3.

When the AEL started just before 1970, large scale development aid was new. Since then there has been almost 40 years of learning by doing in the aid industry, so aid effectiveness should have risen. In 1970 there were hardly any data, but since then they have accumulated steadily. In 1970 econometrics was a new subject at most economics departments. Since then econometrics has been vastly developed. Consequently, we expected that *the estimated* $\mu(N,t)$ *published should increase in size and significance.* This very reasonable expectation

1. The data and the basic correlations are presented as a background note in Paldam (2005).
2. Till now the reviews of the AEL have been qualitative and confined to only a portion of the available evidence, either because they are old as White (1992) or Jepma (1995) or partial as Tsikaka (1998) or Hansen and Tarp (2000). Ours is the first to survey the entire body of evidence, and it is the first quantitative assessment of it.

failed: The published aid-growth effects actually decrease in size and statistical significance. This is an 'unreasonable' result that demands an explanation.

From reading the theoretical, empirical and policy literature, considering the issues and introspection, we believe that many researchers see aid as a hope for a better world without poverty. Hence, they would like to find positive aid-growth effect, i.e. $\mu > 0$. Consequently, researchers are reluctant to publish negative values of $\mu$. We show that this *reluctancy hypothesis* may explain the unreasonable trend in the results.

**(Q2)** We use meta-regression analysis to investigate the source of heterogeneity/variation in the published estimates of $\mu$. Q2 is analyzed in Section 4.

In principle heterogeneity in reported aid-growth effects may have four somewhat overlapping sources: (a) Real differences, due to the countries and periods considered in the studies. (b) Differences due to models and estimation techniques. (c) Differences due to the quality of the research. (d) Differences due to priors of researchers or the publication process. We have included variables trying to account for all these possibilities.

Thus our project may be termed "forensic" economics. It is borne by the hope that the market for research enforces a trend towards truth. That is, truth is revealed by the process of *innovation* (of models and techniques) and *independent replication*: which is replication by other authors on *other* data sets.[3] Independent replication causes innovations to be either rejected or confirmed. In the latter case they become gradually more credible; so that it gives convergence $\mu(t) \to \hat{\mu}$. We try to identify the innovations that generate the results.

This paper thus has a simple structure: Section 2 classifies the AEL and explores some of the puzzles in the literature. Section 2 also explains how we have chosen the 68 aid-growth studies, and discusses the theories behind the AEL. Section 3 presents the data generated from the 68 studies and analyze Q1: Has the literature reached a result that differs from the zero correlation result, which started the quest. Section 4 studies the pattern in the results trying to identify factors that have proven to influence the estimated aid effectiveness. The paper is concluded in section 5. Appendix 1 lists the 68 papers that are the data of our study, and Appendix 2 is an introduction to the techniques of meta-analysis.

3. The criterion of *independent replication* is standard practice in medicine, physics, etc. In macroeconomics such as the AEL it may take some years to obtain independent data sets, but see e.g. Easterly, Levine and Roodman (2004) and Jensen and Paldam (2006). One sometimes sees that studies in some research areas even fail to replicate on the authors own data set. We have not encountered this problem in the AEL.

## 2. The AEL, some puzzles, theory and models

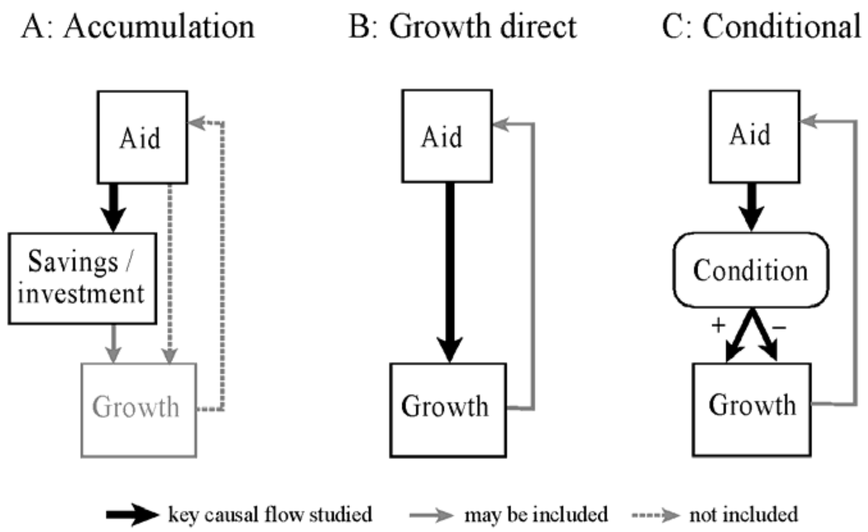
After a thorough search we found 97 empirical macroeconomic studies of the effect of development aid. [4] Papers available after 1st of January 2005 are not included. The entire list for the 97 papers can be found in Doucouliagos and Paldam (2006). In this paper we concentrate on 68 of these studies that contain reduced form models of the type given in Table 1. The basic model is thus simple and though the AEL frequently refers to theory the link from that theory to the estimation model is rarely very explicit.

Table 1. The studies included are based on models of the following type

| Model | (1) $g_{it} = \alpha + \mu h_{it} + \gamma_j \boldsymbol{x}'_{jit} + u_{it}$ | $\mu$ | aid-growth effect |
|---|---|---|---|
| $_{it}$ | index to countries and time [a)] | $\boldsymbol{x}_{itj}$ | vector of $j$ controls |
| $g_{it}$ | real growth rate | $\alpha$, $\gamma_j$ | coefficients to be estimated |
| $h_{it}$ | aid as share of GDP, GNI | $u_{it}$ | residuals |

Note: a) The time unit is normally 3-5 years to reduce variation.

Figure 1. The causal structure in the three families of AEL models




### 2.1 *The three families of studies in the AEL*

The 97 AEL studies use three families of models characterized by their causal structure, as shown in figure 1. About half of the studies contain estimates of models of more than one type. With the resulting overlapping we have reached the following amounts of studies:

4. Extensive searches of Econlit, Proquest, Web of Science and Google were undertaken, from which we could track citations backward.

**A:** 43 papers contain *accumulation estimates* of the impact of aid. These models are covered in Doucouliagos and Paldam (2006), who show that aid has a positive but insignificant effect on savings and investment.

**B:** 68 papers contain a total of 543 *direct estimates* of aid-growth effects from models defined in Table 1. These papers are listed in Appendix 1. They provide the data of the present paper.

**C:** 31 papers contain *conditional estimates*, where the effect of aid on growth depends upon a third factor *z*, so that if *z* is favorable, the result is growth, and vice versa if *z* is unfavorable. Formally the models contain the terms $g = \alpha + \mu h + \omega zh + \ldots$.[5] They are covered in Doucouliagos and Paldam (2005a).

Our data are thus the 543 estimates in the studies of family B. They are all derived from models which all explain growth with an aid-growth term, $\mu h$, giving a comparable estimate of $\mu$. The models differ as regards the *controls* (the *x*'es), the data *sample* and the *estimator*. The meta-regression analysis of section 4 analyze if these differences matter: For that purpose each $\mu$ is provided with a vector of 51 variables characterizing the estimate, almost as a check list, The variables cover all controls which are met in most papers, variables characterizing the sample and estimator, as well as the quality of the research and the possible priors of authors and journals.

The meta-analysis thus analyze if it matters for the key result, $\mu$, if the study is old or new, if the study is published in a top journal or in a working paper, if the model is controlled for simultaneity, if the paper is written by a female researcher etc.

Note that most of the conditionality studies, of type C, report both aid-growth effects, $\mu h$, and second order terms, $\omega zh$. In this paper we treat the second order term as any other *x*.

### *2.2 Casual evidence and the zero correlation result*

The AEL has been fuelled by the fact that the evidence is weak and contradictory, and by the fact that the data seems ideal. Notably three pieces of evidence do not mesh:

(P1) *Contrasting country stories*: South Korea received substantial aid for a decade just before its take off-into high growth. Tanzania has been a main recipient for 40 years, and has

5. The 10 conditions are: 1 good policy; 2 non-linear aid effects; 3 income; 4 democracy; 5 external vulnerability; 6 quality of institutions; 7 trade openness; 8 economic freedom; 9 political stability; and 10 GDP. See Doucouliagos and Paldam (2005a) for details. Only the first two have been subjected to independent replication, which have found them not to be robust.

had little growth. Considerable aid for 30 years to Zambia goes together with an unusually poor economic performance.[6]

(P2) *Micro-macro result*: Studies summarizing project evaluations typically find that about half of all projects succeed while the other half fails, but hardly any harms development. Thus, by aggregation the micro evidence should show that aid increases growth.[7]

(P3) *The zero correlation result*, $r(g,h) \approx 0$, mentioned in the introduction. (P3) contradicts (P2). It has been known as the *micro-macro paradox* since Mosely (1986).

These contradictions raise obvious research questions, and the data seems perfect for providing answers: The (*h,g*)-data cover about 150 countries over 45 years – even when there are gaps there is still about 5000 annual observations. The average aid share is 7.4%. This is substantial relative to other variables that are found to effect growth, i.e., it is more than half the share of net investments, and larger than either the share of the education budget or the share of the health budget in most LDCs. Furthermore, the standard deviation of the aid share is 9.4, so aid data have considerable variation.

### *2.3 Three reflections on the zero correlation result*

First, we note that (P3) is indirectly confirmed by the studies trying to find the factors that give a robust contribution explaining growth and the (lack of) convergence of the poor countries, discussed in Section 2.6. Aid has never made it to the list of robust factors, see Barro and Sala-i-Martin (2004; 541-59) and Sturm and Haan (2005). The 1800 pages of the *Handbook of Economic Growth* (Aghion and Durlauf 2005) that tries to cover all relevant subjects for growth and convergence, does not even mention aid.

Second, the macro result (P3) means that countries receiving much aid do not grow faster than other countries receiving little aid. But, strictly speaking, we want to know how countries that receive aid grow compared with how *they would have grown without the aid.* Thus we need a model providing counterfactuals for all countries receiving aid. The logical counterfactual is to include a set of variables controlling for country heterogeneity.[8]

Third, (P3) implies that any result where $\mu$ is non-zero, based on these data, must be due to the *imposition of structure on the data.* Researchers have done this in four ways:

---

6. Many country studies of the aid-growth connection exist, see e.g., White (1998). Below we only include country studies that contain econometric estimates, of $\mu$.

7. The micro result appears to be uncontroversial, see e.g. Cassen (1986, 1994) and IBRD-OED (annual).

8. The literature on absolute and conditional convergence (see 2.6) use two methods to control for country heterogeneity: (i) fixed effects for countries and (ii) the Barro-set from Barro (1991 and later). Neither of the two sets overcomes the zero-correlation result in the AEL. Consequently, researchers have looked for other controls, often coming up with control sets that seem *ad hoc*.

(S1) By including a *control set*, $\boldsymbol{x}_j$, of *j* other variables in the aid-growth relation. Most such sets control for country differences as already discussed.

(S2) By limiting the data to a particular subset.

(S3) By correcting the (*h,g*)-relation for various biases, notably simultaneity biases, by using the right *estimator*, for each possible bias.

(S4) By interacting the aid effectiveness with a *conditioning factor, z* – see the C group of studies above.

Consequently, many possibilities exist for applying structure to the data. This has allowed researchers to reach all four possible answers to the AEL question. Namely, the effect of aid on growth is: positive, insignificant, negative, or (in the studies of family C) it depends on *another variable, z.*

The second purpose (Q2) of our study is to analyze the pattern in the answers. Why do some authors find a positive aid-growth effect, others no effect and still others a negative effect? Is the heterogeneity in the estimates an artifact of the way studies are conducted (e.g. specification, estimation and data differences), or does it reflect a real phenomenon associated with the existence of a distribution of aid-growth effects (e.g. regional differences and changes in the aid-growth association over time)?

The question lies at the heart of controversy over the impact of foreign aid. If differences in results are created by the application of structure to the data, then the information available to taxpayers and policy makers maybe distorted, and we need to quantify the impact of specification differences on aid-growth effects. If the variation in results is due to an underlying distribution aid-growth effects, so that aid increases growth in certain circumstances and decreases it in others, then it is vital to identify the conditions and realign aid to take advantage of them.[9]

### *2.4 The starting point: Politics, IS-LM theory and the Harrod-Domar legacy*

When the first papers were made in the late 1960s few aid data were available, so the older papers spent a lot of space in the papers discussing politics/policy and economic theory.

9. A parallel body of literature deals with the reverse causality. It explains aid flows (see Alesina and Dollar (2000). The main causal factors appear to be human rights, humanitarian concerns, commercial and security interests and bureaucratic inertia. This literature does not point to growth in the recipient country as an important causal factor for giving aid. Hence, while reverse causality may give a bias in the AEL-relation, we expect that the bias is small.

Several of the early papers were rather explicitly politically, with some researchers belonging to the *new left* – notably Griffin (1970) and Weisskopf (1972) – while others were explicitly *libertarian* – notably Friedman (1958) and Bauer (1971). Interestingly, authors from both schools believed that aid could easily become harmful to the recipient countries, though for different reasons.[10] In our terminology these authors may be thought of as less reluctant to report negative results.

Economic theory may enter as a frame of reference for the variables authors choose for controls and conditioning. That choice also depends upon the availability of the data. We have found more or less explicit references to four types of economic theory in the AEL. Of which the two first are discussed further in Doucouliagos and Paldam (2006).

***Type 1: IS-LM macro theory***. The AEL deals with the activity, $\Delta Y$, which is caused in the longer run, by an amount $H$ (= $hY$) of aid entering a country. The early AEL spent some effort on classifying effects. Two main problems were found: (1) Aid is *fungible*, so what aid actually finances is often different from the marginal outcome. The AEL tries to bypass the fungibility complication by using *reduced form* estimates between aid and "final outcome" variables. (2) The short-run *activity* effect should be separated out from the longer-run *capacity* effect, which is the key purpose of aid. The (IS, LM)-framework suggests that there is both an activity and a capacity effect. This distinction is amazingly rare in the AEL.

***Type 2: Two-gap models.*** For the first two decades of aid, its macro-economic rationale was *Harrod-Domar* type models, which were made explicitly to dynamize the IS-LM model. The policy implication was that the main constraint to development was the *savings* necessary to finance investments. The original Harrod-Domar model is a closed private sector model, so savings are constrained by domestic savings behavior. The introduction of a public sector budget balance gives the first *gap*. When the model is opened, the balance of payment provides a *second gap*,[11] where savings can be provided as transfers from the DC world. This is the approach in the accumulation family of models. This model disappeared from the theory of economic growth during the 1960s, but lingered on in development economics due to its operationality and the clear policy prescriptions it generated.

Most of the AEL-papers from before 1990 have some references to the Harrod-Domar framework. However, gradually the Harrod-Domar framework was replaced by the more flexible *neo-classical framework*, even in development economics. It implies a richer set of

---

10. Also, some studies have been published showing how excessive aid may distort the economy of a country and create a low growth economy; see e.g. Paldam (1997).

11. The best known model of this type was Chenery and Strout (1966). Chapter 2 in Easterly (2001) gives the sad story of the savings gap in development.

channels from aid to growth. However, the regression equations used are formally surprisingly similar to the later models, and we have found no structural breaks in the results reported, due to the shift in the underlying theory.

### *2.5 The Barro-framework and more eclectic theory*

The new papers devote gradually less space to discussing theory. The problem is defined, and there is a set of papers at hand to start from. Many just start from a published paper and suggest that it may have a problem, which can be solved by a change of one or two terms in the relation, by using a newer estimator etc. The reference to theory is hence often vague. However, the main inspiration appears to be:

***Type 3: Barro-type growth regressions***. Since the early 1990s the best known empirical tool in modern growth theory has been the Barro-type growth regression, which has some relation to the neoclassical growth model.[12] The connection is general, so that the theory suggests where to seek for explanatory variables, but the actual choice is (strongly) influenced by data availability, and does have a large element of *ad hoc* choice. The Barro-model was set up to study convergence, $\beta$, which is the coefficient in the *convergence term*, $\beta \log y_{it}$, in the following estimating relation:[13]

(2) $g_{it} = \alpha + \beta \log y_{it} + \boldsymbol{\gamma}_j \boldsymbol{x'}_{jit} + u_{it}$, see table 1 for definitions of variables.

Many such models have been estimated – in robustness experiments even millions.

Each model (2) is converted into a model of type (1), by replacing or supplementing the convergence term with the aid efficiency term. It is a problem for this conversion that the estimate of $\mu$ may be biased, if the controls used depend on aid. Some of the variables in the original Barro-set of controls have this problem. For example, the standard Barro estimates have variables for education, health and the share of public consumption, $c_{it}$, in the $\boldsymbol{x}$-set. If $c$ is singled out the model looks like this:

(3) $g_{it} = \alpha + \mu h_{it} + \nu c_{it} + \boldsymbol{\gamma}_j \boldsymbol{x'}_{jit} + u_{it}$

---

12. It is published in various versions since 1991. The newest is Barro and Sala-i-Martin (2004; cpt. 12).

13. The convergence literature distinguishes between *absolute* convergence, which disregards country heterogeneity, and *conditional* convergence, which controls for country heterogeneity. The standard result is that absolute convergence is rejected, while conditional convergence is accepted. If this terminology is used about the AEL, it is uncontroversial that absolute aid effectiveness is rejected, while the discussion deals with conditional aid effectiveness. Unfortunately the term *conditional* is used in the AEL, for models where aid enters multiplied with another term. In the text we adopt the AEL terminology.

One of Barro's results, which is also found in the AEL, is that public consumption is harmful for growth, $\nu < 0$. The typical result in the accumulation family (A) of the AEL is that considerable crowding out of the capacity effect of aid does occur, so that aid increases $c$. Then (3) biases the estimate of $\mu$ upwards.[14]

***Type 4: Political economy model.*** A new approach notes that aid is a new external rent entering the economy. Here it may influence domestic political stability, and consequently growth via the *stability-growth channel*: If political stability changes, the investment climate improves, investment changes and so does growth.[15] This may give a new family of models in the future, but till now political economy variables, have just been added as other control variables tried.

The theories referred to in the AEL are thus a well-known set of theories, which has changed in line with the general development of economic theory. However, irrespective of the development of theory the models estimated have till now remained formally similar.

### *2.6 Estimators: Have they changed results?*

The models of type (1) have been run as cross-country models with no $t$-index, as single country time-series models with only the $t$-index or as panel models with both indices. Also, biases have been removed by doing TSIV-estimates using some of the controls in the first stage, or as GMM-estimates using lags systematically, etc. The rapid development in econometrics has caused the standard econometric packages to be constantly expanded with new tools. These tools are also used in the AEL. When a new tool appears, it moves the frontier of best practice outward. Consequently the new tool has to be applied for articles to increase the chances of acceptance.

However, only one example (Giles 1994) has been reported in any of the papers in the AEL where the use of more advanced econometrics has changed the substantial results (from just below to above significance).[16] We have also seen examples where two-stage techniques were used in the working paper, and replaced with OLS in the final version or vice versa, with no obvious effect. We test for the effect of estimation technique below, and find that it does not appear to matter.

---

14. That result is neatly reproduced below. When we test for the effect of including various variables in the estimating equation for $\mu$ we find that the inclusion of $c$ does indeed increase the size of $\mu$.
15. An extra rent can in principle have two effects: It may turn some internal distributional fights from a negative sum game into a positive sum game, and hence be politically stabilizing and *increase* growth. Alternatively, it is a new contestable rent that increases distributional fights, and *reduces* political stability and growth as resources are often found to do, see e.g. Collier and Hoeffler (2000). This gives two possible outcomes which become the subject for empirical analysis.
16. The AEL thus seems to support the critique of econometrics in chapter 6 of McCloskey (1998).

## 3. Three data sets and Q1: Do the estimated aid-growth effects converge?

The data we wish to submit to our analysis is the *population* of the 68 aid-growth studies listed in Appendix 1.[17] We claim that they are virtually every study that contains a reduced form macro estimate of the aid-growth effect, $\mu$.

### *3.1 The three data sets for estimated aid-growth effect*

Our measure of size of the effect is the partial correlation between foreign aid and economic growth. While most studies provide enough information from which to calculate the elasticity of growth with respect to foreign aid, many do not.[18] Hence, we use partial correlations.[19]

We derive three data sets for the aid-growth effect. The ***best-set*** consists of 68 observations, one from each of the 68 papers, using the key regression from each paper. Unfortunately, it is not always clear what is the authors' preferred estimate, so we have sometimes had to assess. The ***all-set*** consists of all 543 regressions reported in the 68 papers, greatly increasing the data available for tests, but it gives some interdependence between data points. We construct also the ***average-set*** by taking an average of all regressions reported by each study. This produces 68 observations. We use the average-set as a useful robustness test for the best-set, as there is no subjectivity involved in the construction of the average-set.

### *3.2 A first look at the data – and three observations*

Figure 2 is a funnel plot of the partial correlations, showing their association to the sample size, for the all-set and best-set, respectively. A majority of the estimates are positive, but most are statistically insignificant. The weighted average partial correlation is +0.08 for the all-set and is +0.09 for the best-set, with sample size used as weights.[20] Funnel plots should be symmetrical around the true population effect, if there is only one, and the sampling error should be larger for the smaller studies so we expect that they are more spread out than the larger studies.

17. The focus of this paper is solely on the analysis of the estimates of the impact of aid on growth; see however, Doucouliagos and Paldam (2005a and 2006) for studies of the remaining effects.
18. These are typically studies that do not measure aid as the percentage of GDP or where the scaling of the variables was not clear.
19. We have reproduced the meta-analysis for separate groups of studies, but the conclusions remain robust. For example, ignoring the early group of studies (those published in the 1970s) does not change any of the conclusions. There is no reason for ignoring the contributions made and information contained in the early literature.
20. Standard practice in meta-analysis requires that effect sizes (such as correlations) are weighted by the sample size of the study assuming that larger studies will, *ceteris paribus*, be more accurate (Hunter and Schmidt 2004).

Figure 2. Funnel plot for aid-growth effects, $\mu$. Looking for: $\mu(N) \rightarrow \hat{\mu}$

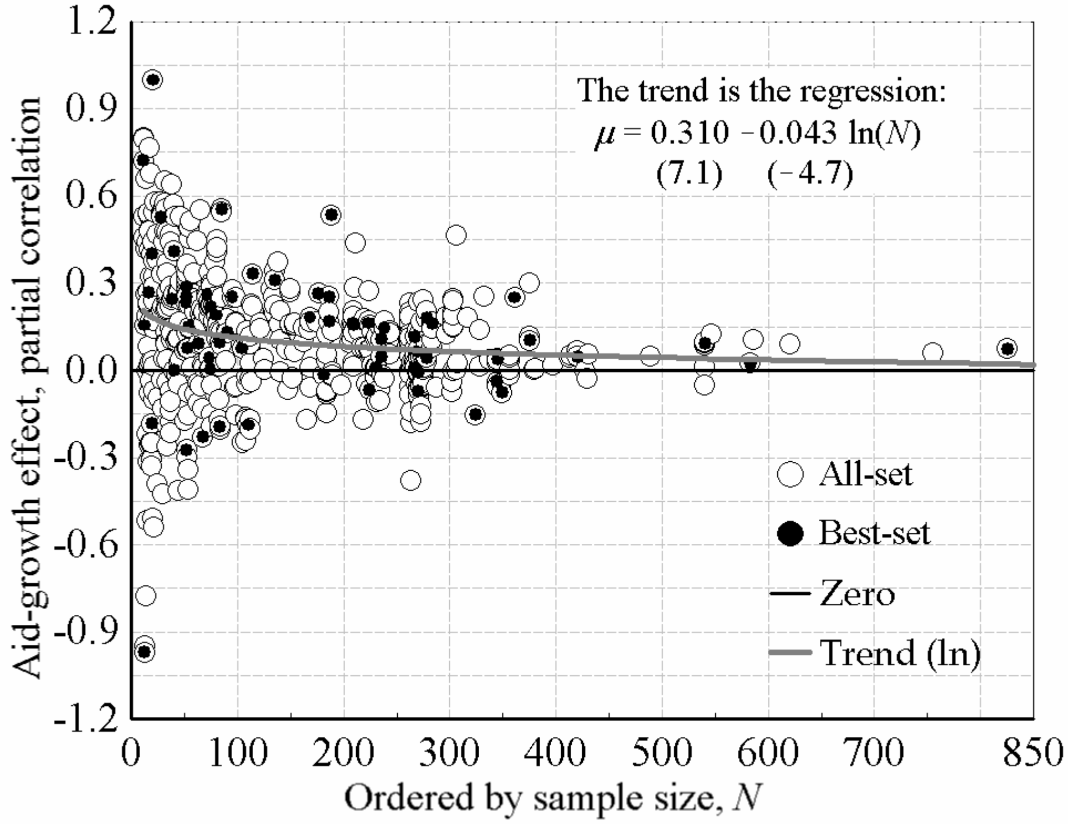



Figure 3. Time series graph of aid-growth effects: Looking for $\mu(t) \rightarrow \hat{\mu}$

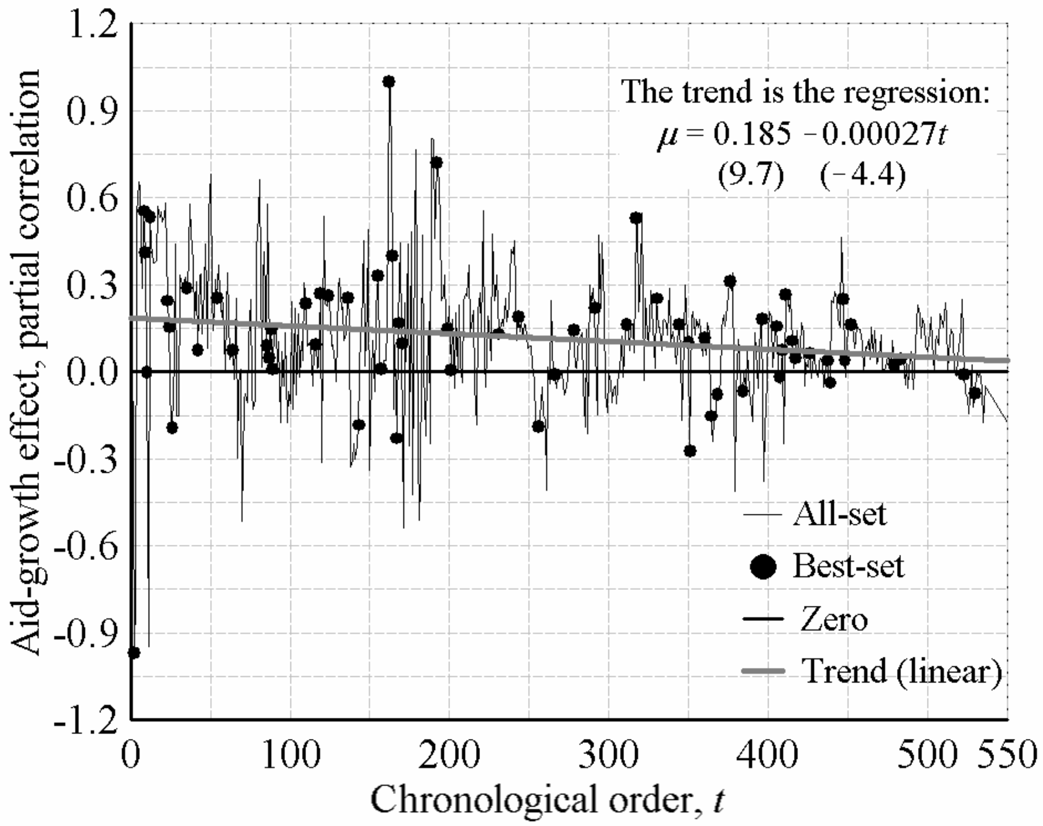


Figure 3 is a time series graph of the empirical findings presenting the same partial correlations as in Figure 2, but this time in chronological order. It is clear that although most results are above zero, they fluctuate around the zero line, and they have a significant downward time trend. Figures 2 and 3 give rise to three observations:

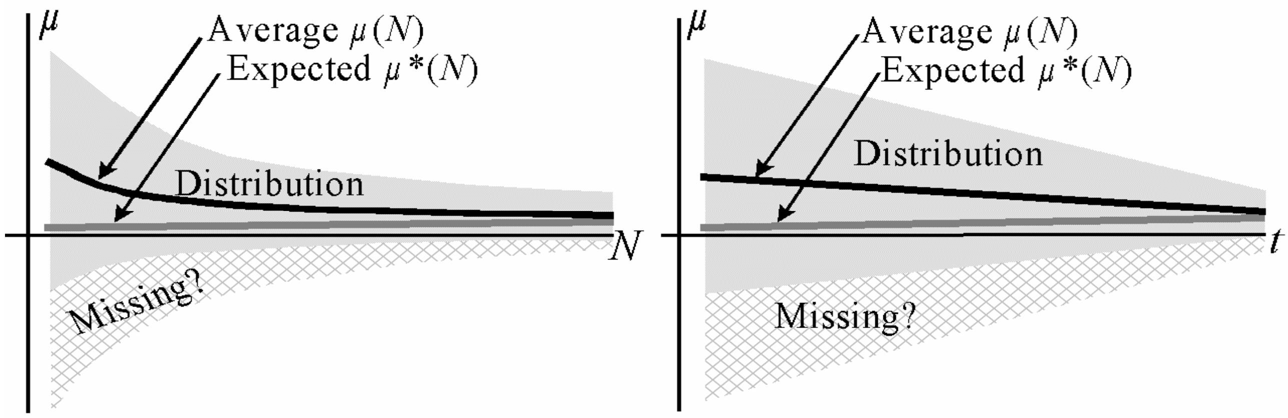
(O1) ***Convergence:*** Contrary to expectations, the trends on the figures have negative slopes. Also, it is clear that if $\mu$ converges it is to a small $\hat{\mu}$. See the MST tests of Table 3.

(O2) ***Asymmetry:*** If we accept that $\hat{\mu}$ is small then there are too few negative estimates at the left hand part of the two figures. See the FAT tests of Table 3.

(O3) There are no structural breaks, when new ***models*** or ***estimates*** are introduced. See Section 4.

Figure 4. Our explanation of the falling trends on Figures 2 and 3

Figure 4a. Stylized version of Figure 2 Figure 4b. Stylized version of Figure 3




Note: The gray areas give the distribution of the reported coefficient estimates, and the checkered gray area are the estimates which may have been hit by reluctance. While the $\mu = \mu^*(t)$ lines are the expected ones, which we believe are closer to the truth, the literature has revealed the $\mu = \mu(t)$ lines.

The introduction presented the *reluctancy hypothesis* that aid effectiveness is a field where researchers are reluctant to publish negative coefficients. We take (O2) to corroborate the hypothesis,[21] and the missing negative values may then explain the unreasonable result (O1), as sketched on Figure 4.

21. We have tried to imagine alternative explanations, based on the change in models over time, but we have not managed to construct anything that appears plausible.

The standard tests of meta-analysis are made precisely to prove if suggestions such as (O1) to (O3) can be rejected. Here (O1) and (O2) are relevant for Q1 and will be tested in the rest of the section. Appendix 1 gives a brief introduction to these tests.

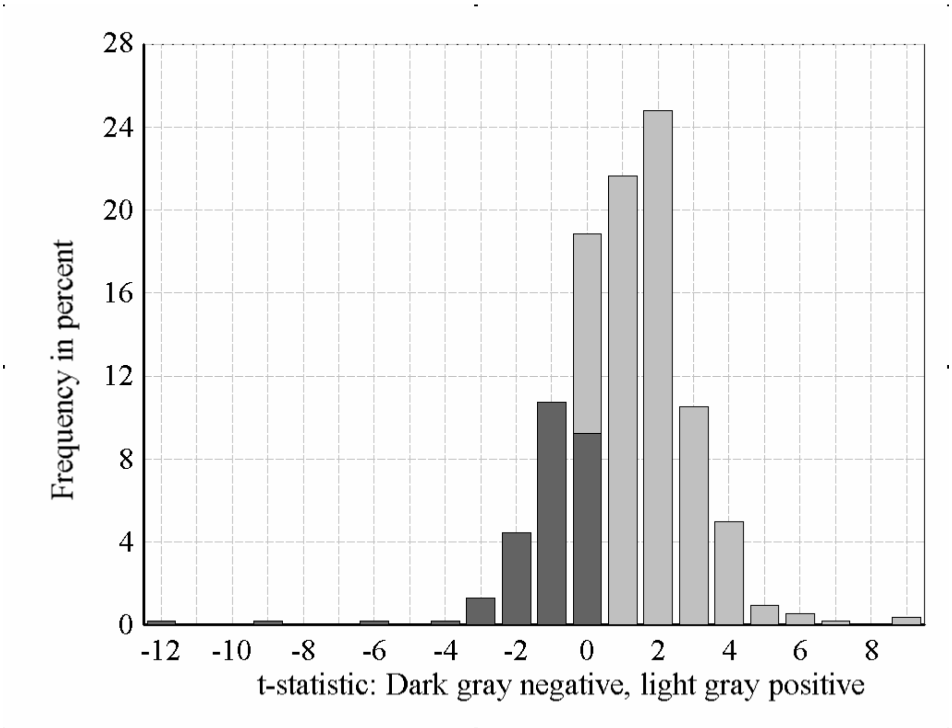
### *3.3 Vote counting*

Table 2 classifies the empirical results by sign and statistical significance. As explained in Appendix 1, this is almost an extreme bounds analysis, but it is not a reliable way to summarize the results of a literature in the presence of model polishing and asymmetry. However, it offers a first overview of what the literature has established.

Table 2. Meta-extreme bounds analysis of the published estimates of aid-growth effects, $\mu$

| | | Positive | | Negative | |
|---|---|---|---|---|---|
| | | Significant | Insignificant | Insignificant | Significant |
| Best-set | 68 Studies | 31 (46%) | 22 (32%) | 9 (13%) | 6 (9%) |
| All-set | 543 Estimates | 207 (38%) | 198 (36%) | 106 (20%) | 32 (6%) |

Figure 5. Histogram of t-statistics for the estimates of aid-growth effects, $\mu$




As shown by Figures 2 and 3, the best-set concentrates on the more significant results. Almost half, 46%, of the 68 studies found a significantly positive aid-growth effect, while 54% are either not statistically significant or negative. The 543 estimates of the all-set give a weaker version of the same pattern. It will come as no surprise that authors prefer the more significant results. This can also be seen in Figure 5, which is a histogram of the 543 t-statistics reported in the AEL. The average t-statistic is 1.05, the median is 1.15, and the weighted average t-statistic is 1.22 (using sample size as weights).

Table 3. Meta-significance and funnel asymmetry tests, aid-growth effects

(Dependent Variable = natural logarithm of t-statistics for MST and reported t-statistic for FAT)

| Variable | (1) MST All-set | (2) MST Best-set | (3) MST Average-set | (4) FAT All-set | (5) FAT Best-set | (6) FAT Average-set |
|---|---|---|---|---|---|---|
| | | | *All Studies/ All Years* | | | |
| Constant | 0.27 | 1.05 | 0.28 | **0.80** | **1.02** | **0.64** |
| | (1.14) | (1.20) | (0.38) | (5.03)*** | (2.24)** | (1.25) |
| ln(*df*) | **-0.04** | **-0.17** | **-0.05** | - | - | - |
| | (-0.77) | (-0.98) | (-0.30) | | | |
| 1/SE | - | - | - | 0.02 | 0.01 | 0.03 |
| | | | | (1.76)* | (0.34) | (0.66) |
| Adjusted $R^2$ | 0.00 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 |
| F-statistic | 0.76 | 1.26 | 0.15 | 2.65 | 0.04 | 0.35 |
| N | 543 | 68 | 68 | 539 | 67 | 67 |
| Average *r* | +0.08 | +0.09 | +0.08 | +0.08 | +0.09 | +0.08 |
| Average *t* | +1.22 | +1.54 | +1.28 | +1.22 | +1.54 | +1.28 |
| | | | *1970-1999* | | | |
| Constant | 0.26 | 1.17 | -0.14 | **0.73** | **0.45** | **0.33** |
| | (0.96) | (1.06) | (-0.19) | (3.51)*** | (0.96) | (0.57) |
| ln(*df*) | **-0.02** | **-0.20** | **0.11** | - | - | - |
| | (-0.27) | (-0.87) | (0.74) | | | |
| 1/SE | - | - | - | 0.04 | 0.06 | 0.08 |
| | | | | (1.51) | (1.05) | (1.68) |
| N | 259 | 35 | 35 | 257 | 34 | 34 |
| Average *r* | +0.11 | +0.12 | +0.15 | +0.11 | +0.12 | +0.15 |
| Average *t* | +1.51 | +2.28 | +2.26 | +1.51 | +2.28 | +2.26 |
| | | | *2000- 2005* | | | |
| Constant | 0.61 | 0.88 | 0.64 | **0.50** | **2.76** | **0.61** |
| | (1.20) | (1.02) | (0.73) | (1.46) | (3.48)*** | (0.67) |
| ln(*df*) | **-0.11** | **-0.13** | **-0.15** | - | - | - |
| | (-1.11) | (-0.75) | (-0.82) | | | |
| 1/SE | - | - | - | 0.04 | -0.10 | 0.02 |
| | | | | (1.46) | (-1.70)* | (0.27) |
| N | 284 | 33 | 33 | 282 | 33 | 33 |
| Average *r* | +0.07 | +0.07 | +0.05 | +0.07 | +0.07 | +0.05 |
| Average *t* | +1.11 | +1.09 | +0.87 | +1.11 | +1.09 | +0.87 |

Explanation: The bolded coefficients are the key statistics in the table. If aid has an effect on growth, ln(*df*) in the MST should have a positive and statistically significant coefficient. This fails in all 9 cases. If the literature is free of publication bias, the constant in the FAT should not be statistically significant. It is significant in 4 cases. Note: *, **, *** statistically significant at the 10%, 5% and 1% level, respectively. t-statistics in brackets, using robust standard errors. The bootstrap was used for the All-Set. Average *r* is the weighted average partial correlation between aid and economic growth, using sample size as the weight. Average t is the weighted average t-statistic. Some observations are lost due to missing data in some cases.

### 3.4 *Tests for* (O1) *convergence of results and* (O2) *asymmetry*

Table 3 presents meta-analytical test, which are designed to analyze precisely the questions raised by observations (O1) and (O2): The MST-tests of columns (1) to (3) in the table tests if the results of the 68 aid-growth studies converge to a non-zero result. The FAT-tests of columns (4) to (6) in the table tests if the literature suffers from *polishing* effects, in the following two senses: Are the results too good in small samples? Are the results asymmetrical around the results generated in the largest samples, as suggested by the reluctancy hypothesis.

The top panel of Table 3 presents the tests for all studies combined together. For sensitivity purposes, the second panel looks at the earlier literature (studies published between 1970 and 1999) and the last panel explores the more recent literature. The average partial correlation between aid and growth, *r*, and the average t-statistic is also reported. The average *r* is positive, but we need to test whether it is statistically significantly different from zero. To do this, it is not sufficient to simply look at the value of the average t-statistic, it is necessary to explore the association between t-statistics and sample size (the MST).

The MST regresses the natural logarithm of degrees of freedom on the natural logarithm of the associated t-statistics. The desired outcome for a genuine effect (the MST) between aid and growth is that the coefficient to degrees of freedom (ln(*df*)) should be significantly positive (Stanley 2001 and 2005).[22] That is, t-statistics should rise as sample size and hence estimation accuracy rises. This is not the case in any of the 9 tests reported. There is no evidence that development aid has a direct effect on economic growth.[23]

The FAT explores the association between reported standardized aid-growth effects and their associated standard errors. If authors using smaller samples publish too many statistically significant results, then this will be revealed in a positive association between aid-growth effects and the standard errors of the effects. However, instead of regressing the standard error against aid-growth effects, it is standard practice to remove heteroscedasticity by dividing this regression by the standard error.[24] Hence the FAT for polishing is that the constant is non-zero as is the case for most estimates in the all-set and best-set in the table. This is the typical finding in meta-analysis (Roberts and Stanley 2005).[25] The sign on the

22. See Appendix 2 for discussion on the MST. We estimated also several versions of the MSTMRA tests, which are the multivariate version of the MST. For example, we included control variables for country composition, year dummies, estimation technique and type of data used (EDA and panel data). The results are robust and always similar to those presented in Table 3. The full set of results is available from the authors.

23. Note that even though the average t-statistic for the best-set and the average-set for the 1970-99 group of studies is 2.28 and 2.26, respectively, the MST is not passed.

24. See Appendix 2 for discussion on the FAT.

25. Indeed, it is a poignant fact that the majority of studies have found that polishing is standard practice in empirical analysis. Examples in economics include Card and Krueger (1995), Ashenfelter *et al.* (1999), Görg

constant informs on the direction of the polishing effect. A positive constant means that the literature has, on average, published too many positive aid-growth effects.

Consequently, we have confirmed the two observations (O1) and (O2) in Section 3, and the reluctancy hypothesis. However, there are some signs that the estimated results are heterogeneous. We therefore turn to our second question.

## 4. Q2: Explaining the pattern in the results

In the previous section it was shown that there is a positive, but statistically insignificant association between development aid and economic growth. In this section we use meta-regression analysis (MRA) to take a closer look at the variation in the empirical aid-growth results and identify some of the sources of these differences. Meta-regression analysis is gaining widespread appeal among economists, and the AEL is fertile ground for its application. Recent examples include Gorg and Strobl (2001), Doucouliagos and Laroche (2003), Jarrell and Stanley (2004) and Dobson *et al.* (2006).

### *4.1 The explanatory variables*

The variables used in our meta-regression analysis are defined in table 4. We consider six classes of explanatory variables: The first two are sets of dummy variables for (C1) 7 journals[26] and (C2) 5 author characteristics. As regards (C2) we are influenced by our findings on the aid-conditionality literature, which found that two prolific groups produced significantly different results: One was affiliated with the World Bank and the other with Danida (Danish Development Aid), see Doucouliagos and Paldam (2005a).[27] The variable *influence* tests whether "friends" acknowledged in the paper may influence results.[28]

---

and Strobl (2001), Roberts and Stanley (2005) and Monkerjee (2006). The situation is the same in other disciplines (see Sutton *et al.* 2000).

26. Previous meta-analyses have identified important differences in research findings between journals in labor economics (see Doucouliagos and Laroche 2003) and in the economic growth literature (see Doucouliagos and Ulubasoglu 2006).

27. The World Bank group papers are: Burnside and Dollar (2000; 2004); Collier and Dehn (2001); Collier and Dollar (2002); Collier and Hoeffler (2004); and Svensson (1999). The Danida group papers are: Dalgaard and Hansen (2001); Dalgaard, Hansen and Tarp (2004); and Hansen and Tarp (2000; 2001). It is only fair to mention that Dalgaard is only weakly associated with Danida.

28. This variable is constructed from information contained in acknowledgements made in the actual papers. Influence does not mean affiliation. Rather, the variable captures the effect other authors have on research results. As such, it may be an indication of research quality, if feedback enhances the research process.

Table 4. Definition of variables and their size and effect, when used alone (All-Set)

| Nr | Variable | BD: binary dummy that is 1 if condition holds, otherwise 0 | Mean | SD | Eff. | Prob |
|---|---|---|---|---|---|---|
| | *Dependent* | The partial correlation of aid and economic growth | 0.11 | 0.22 | - | - |
| | | *(C1) Publication Outlet* | | | | |
| 1 | *WorkPap* | BD for unpublished paper | 0.26 | 0.44 | -0.05 | 0.003 |
| 2 | *Cato* | BD for Cato Journal | 0.03 | 0.16 | 0.08 | 0.04 |
| 3 | *JDS* | BD for Journal of Development Studies | 0.08 | 0.27 | 0.01 | 0.71 |
| 4 | *JID* | BD for Journal of International Development | 0.05 | 0.22 | -0.01 | 0.80 |
| 5 | *EDCC* | BD for Economic Development and Cultural Change | 0.06 | 0.23 | -0.16 | 0.00 |
| 6 | *AER* | BD for American Economic Review | 0.04 | 0.20 | -0.11 | 0.00 |
| 7 | *AE* | BD for Applied Economics | 0.11 | 0.31 | 0.04 | 0.38 |
| | | *(C2) Author Details* | | | | |
| 8 | *Danida* | BD for author(s) affiliated with the Danida group | 0.09 | 0.29 | 0.01 | 0.50 |
| 9 | *WorldBk* | BD for author(s) affiliated with the World Bank | 0.09 | 0.29 | -0.15 | 0.00 |
| 10 | *Gender* | BD if at least one of the authors is female | 0.12 | 0.33 | 0.03 | 0.39 |
| 11 | *Expectations* | BD for author with realized expectations about aid-growth | 0.10 | 0.30 | -0.05 | 0.02 |
| 12 | *Influence* | BD for authors acknowledging other authors in the AEL | 0.22 | 0.42 | 0.03 | 0.15 |
| | | *(C3) Data* | | | | |
| 13 | *Panel* | BD for use of panel data | 0.67 | 0.47 | -0.06 | 0.01 |
| 14 | *NoCount* | Number of countries included in the sample | 48 | 28 | -0.09 | 0.02 |
| 15 | *NoYears* | Number of years covered in the analysis | 20 | 8 | -0.01 | 0.00 |
| 16 | *Africa* | BD for countries from Africa included | 0.83 | 0.37 | -0.13 | 0.00 |
| 17 | *Asia* | BD if countries from Asia included | 0.75 | 0.43 | -0.02 | 0.49 |
| 18 | *Latin* | BD if countries from Latin America included | 0.76 | 0.43 | -0.09 | 0.01 |
| 19 | *Single* | BD if data from a single country | 0.05 | 0.22 | 0.16 | 0.08 |
| 20 | *Y1960s* | BD if data for the 1960s | 0.21 | 0.40 | 0.02 | 0.53 |
| 21 | *Y1970s* | BD if data for the 1970s | 0.81 | 0.39 | -0.11 | 0.00 |
| 22 | *Y1980s* | BD if data for the 1980s | 0.84 | 0.37 | -0.10 | 0.00 |
| 23 | *Y1990s* | BD if data for the 1990s | 0.60 | 0.49 | -0.03 | 0.11 |
| 24 | *SubSam* | BD if data relate to a sub-sample of countries | 0.28 | 0.45 | 0.03 | 0.18 |
| 24 | *LowInc* | BD if data relate to a sub-sample of low income countries | 0.09 | 0.28 | -0.03 | 0.38 |
| 26 | *EDA* | BD for use of EDA data | 0.24 | 0.43 | -0.10 | 0.00 |
| 27 | *Outlier* | BD if outliers were removed from the sample | 0.13 | 0.34 | -0.05 | 0.00 |
| | | *(C4) Conditionality* | | | | |
| 28 | *Nonlinear* | BD for aid squared added | 0.18 | 0.38 | 0.03 | 0.03 |
| 29 | *Aid*Policy* | BD for aid interacted with policy | 0.26 | 0.44 | -0.10 | 0.00 |
| 30 | *Aid*Institut* | BD for other aid interacted terms (mainly institutions) | 0.26 | 0.44 | -0.21 | 0.00 |
| | | *(C5) Specification and Control* | | | | |
| 31 | *Capital* | BD for control for domestic savings or investment | 0.47 | 0.50 | 0.14 | 0.00 |
| 32 | *FDI* | BD for control for foreign capital inflows (other than aid) | 0.27 | 0.44 | 0.10 | 0.00 |
| 33 | *GapModel* | BD for two-gap model | 0.28 | 0.45 | 0.07 | 0.01 |
| 34 | *Theory* | BD for paper developing a theory | 0.30 | 0.46 | -0.05 | 0.02 |
| 35 | *Average* | Number of years involved in data averaging | 6.7 | 5.4 | 0.01 | 0.58 |
| 36 | *LagUsed* | BD for use of lagged value of aid | 0.11 | 0.31 | 0.06 | 0.13 |
| 37 | *Inflation* | BD for control for inflation | 0.14 | 0.35 | -0.03 | 0.05 |
| 38 | *Instability* | BD for control for political instability | 0.29 | 0.46 | -0.08 | 0.00 |
| 39 | *Fiscal* | BD for control for fiscal stance | 0.12 | 0.32 | 0.00 | 0.96 |
| 40 | *GovSize* | BD for control for size of government | 0.12 | 0.32 | 0.05 | 0.01 |
| 41 | *FinDev.* | BD for control for financial development | 0.32 | 0.47 | -0.04 | 0.03 |
| 42 | *Ethno* | BD for control for ethnographic fractionalization | 0.32 | 0.47 | -0.09 | 0.00 |
| 43 | *Region* | BD for regional dummies | 0.36 | 0.48 | -0.08 | 0.00 |
| 44 | *HumCap* | BD for control for human capital | 0.25 | 0.43 | -0.02 | 0.42 |
| 45 | *Open* | BD for control for trade openness | 0.29 | 0.46 | 0.06 | 0.00 |
| 46 | *PopSize* | BD for control for population size | 0.29 | 0.46 | 0.06 | 0.01 |
| 47 | *GdpLev* | BD for control for per capita income | 0.63 | 0.48 | -0.11 | 0.00 |
| 48 | *Policies* | BD for control for policies | 0.35 | 0.48 | -0.10 | 0.00 |
| | | *(C6) Estimation* | | | | |
| 49 | *OLS* | BD for use of OLS | 0.68 | 0.47 | 0.04 | 0.002 |
| 50 | *Gr&Aid* | BD eqs. system with *both* a growth and an aid equation | 0.04 | 0.20 | 0.07 | 0.00 |
| 51 | *Gr&Savings* | BD eqs. system with *both* a growth and a savings equation | 0.02 | 0.13 | 0.08 | 0.04 |

The remaining 4 classes of variables contain all main differences between the operational models we have been able to quantify. They are: (C3) 15 data characteristics; (C4) are 3 variables for the inclusion of conditional variables, where aid is supplemented with a second order aid term; (C5) 18 variables for model formulation and controls; and finally (C6) are 3 variables for estimation techniques. Hence, (C3) to (C6) represent the modeling and research design choices of the authors.

The 15 variables (C3) for data differences try to capture the differences in data used by researchers. This includes controls for: type of data (typically panel or cross-sectional); whether the study looks at a group of countries or a single country; sample size (countries and years); region (Africa, Asia, Latin America); time period covered; and whether parts of samples are used (sub-sample, low income sample and removal of outliers). Model specifications (C5) may influence results – we use 18 variables that capture most of the differences in specifications, and are hence able to perform a meta-robustness testing. Through these we are able to explore how robust the aid-growth effects are and can quantify the effects of changing specification, as well as data differences etc.

It is important to note that most of these control or moderator variables are chosen by the authors simply by being included in the models. The large number of variables is due to the variability of the models covered. It is common in meta-analysis in economics to use an extensive set of control variables (see, for example, the papers in Roberts and Stanley 2005). In order to secure as much objectivity as possible we have included as much as we could, and then tested down to reduce the number of variables to the significant ones.

### *4.2 Explaining the reported aid-growth effects*

Our aim in this section is to use MRA to explore the source of the heterogeneity in reported aid-growth effects. We use the all-set as its larger sample size enables a broader analysis of study differences. The MRA results for the all-set are presented in Table 5.

Column 1 presents the results for several alternative groupings of variables. Column 2 presents the OLS results with all potential explanatory variables combined. Many of the observations in the all-set are not strictly statistically independent. Accordingly, Table 5 reports two t-tests, (t) and (tb) the first is calculated from the usual standard errors and the second from bootstrapped standard errors (see Efron and Tibshirani 1993).[29] It is reassuring that the two t-tests are rather close in column (2) reporting both.

29. The bootstrap used 1,000 replications with replacement.

Table 5. Meta-Regression Analysis, Aid-Growth Effects, All-Set
(Dependent variable = partial correlation)

| Variable | (1)<br>OLS (t)<br>Each group | (2)<br>OLS all (t, tb)<br>Fixed Effects | (3)<br>OLS all (tb)<br>Random Effects | (4)<br>Only robust (tb)<br>Fixed Effects |
|---|---|---|---|---|
| *Constant* | various | 0.28 (2.20, 2.04)** | 0.09 (0.92) | 0.21 (4.90)*** |
| *WorkPap* | -0.09 (-4.65)*** | -0.01 (-0.12, -0.11) | 0.04 (1.25) | - |
| *Cato* | -0.13 (-3.22)*** | -0.09 (-0.92, -0.87) | -0.09 (-1.39) | - |
| *JDS* | -0.04 (-1.51) | -0.07 (-1.05, -1.00) | -0.04 (-0.87) | - |
| *JID* | -0.06 (-2.18)** | -0.06 (-0.96, -0.91) | -0.07 (-1.36) | - |
| *EDCC* | -0.20 (-3.41)*** | **-0.57 (-3.17, 2.80)*** ** | **-0.62 (-7.70)*** ** | **-0.58 (-2.94)*** ** |
| *AER* | -0.16 (-7.49)*** | -0.05 (-0.70, 0.63) | 0.01 (0.18) | - |
| *AE* | -0.01 (-0.24) | **-0.12 (-1.74, -1.68)* ** | **-0.13 (-2.55)** ** | **-0.14 (-2.30)** ** |
| *Danida* | -0.01 (-0.43) | **0.11 (2.02, 1.86)** ** | **0.16 (3.91)*** ** | **0.09 (3.88)*** ** |
| *WorldBk* | -0.17 (-7.81)*** | 0.06 (1.09, 1.02) | 0.01 (0.05) | 0.06 (1.88)* |
| *Gender* | 0.02 (0.53) | 0.01 (0.10, 0.09) | 0.05 (1.27) | - |
| *Expectations* | 0.03 (1.23) | 0.01 (0.24, 0.22) | 0.07 (1.26) | - |
| *Influence* | 0.04 (2.19)** | **0.11 (3.67, 3.65)*** ** | **0.10 (3.75)*** ** | **0.12 (5.87)*** ** |
| *Panel* | 0.02 (0.54) | -0.01 (-0.20, -0.18) | 0.02 (0.63) | - |
| *NoCount* | 0.00 (0.07) | 0.00 (0.06, 0.06) | 0.001 (1.66)* | -0.001 (-2.20)** |
| *NoYears* | -0.01 (-1.09) | -0.01 (-1.46, -1.36) | -0.00 (-0.45) | - |
| *Africa* | -0.07 (-1.52) | -0.04 (-0.93, -0.88) | -0.03 (-0.70) | - |
| *Asia* | 0.11 (2.35)** | **0.12 (2.51, 2.45)** ** | **0.11 (2.62)*** ** | **0.09 (2.60)*** ** |
| *Latin* | -0.15 (-3.17)*** | -0.07 (-1.42, -1.31) | -0.04 (-0.97) | - |
| *Single* | 0.08 (0.69) | 0.26 (1.60, 1.50) | 0.56 (6.64)*** | 0.27 (2.83)*** |
| *Y1960s* | -0.09 (-1.83)* | -0.04 (-0.68, -0.64) | -0.06 (-1.60) | - |
| *Y1970s* | -0.11 (-2.93)*** | **-0.12 (-2.92, -2.81)*** ** | **-0.07 (-1.95)* ** | **-0.12 (-3.91)*** ** |
| *Y1980s* | -0.16 (-2.80)*** | **-0.14 (-1.98, -1.91)* ** | **-0.12 (-2.13)** ** | **-0.07 (-1.75)* ** |
| *Y1990s* | 0.08 (2.30)** | 0.15 (2.59, 2.52)** | 0.03 (0.86) | 0.08 (2.01)** |
| *SubSam* | -0.01 (-0.20) | -0.03 (-0.83, -0.79) | 0.01 (0.33) | - |
| *LowInc* | -0.01 (-0.08) | 0.02 (0.40, 0.39) | -0.01 (-0.25) | - |
| *EDA* | -0.06 (-2.97)*** | **-0.06 (-2.51, -2.35)** ** | **-0.05 (-1.89)* ** | **-0.07 (-3.41)*** ** |
| *Outlier* | -0.02 (-1.28) | -0.01 (-0.23, -0.22) | -0.01 (-0.34) | - |
| *Nonlinear* | 0.03 (2.22)** | -0.02 (-0.68, -0.66) | -0.02 (-0.82) | - |
| *Aid·Policy* | -0.08 (-5.67)*** | 0.01 (0.21, 0.20) | 0.02 (0.71) | - |
| *Aid·Institut* | -0.18 (-6.93)*** | **-0.09 (-1.93, -1.73)* ** | **-0.08 (-1.90)** ** | **-0.10 (-3.22)*** ** |
| *Capital* | 0.12 (2.30)** | 0.09 (1.88, 1.74)* | 0.10 (2.39)** | - |
| *FDI* | 0.03 (0.86) | **0.08 (1.87, 1.72)* ** | **0.08 (2.01)** ** | **0.09 (2.79)*** ** |
| *GapModel* | -0.07 (-1.64) | -0.01 (-0.14, -0.13) | 0.07 (1.46) | - |
| *Theory* | 0.03 (1.19) | 0.02 (0.57, 0.55) | -0.02 (-0.91) | - |
| *Average* | 0.01 (0.85) | 0.00 (0.39, 0.39) | 0.00 (0.33) | - |
| *LagUsed* | 0.02 (0.44) | 0.07 (1.43, 1.41) | 0.09 (2.93)** | - |
| *Inflation* | -0.07 (-2.26)** | -0.08 (-1.85, -1.80)* | -0.06 (-1.53) | -0.07 (-2.84)*** |
| *Instability* | 0.01 (0.13) | 0.10 (1.52, 1.31) | 0.07 (1.51) | 0.12 (2.10)** |
| *Fiscal* | 0.05 (1.28) | 0.05 (1.11, 1.41) | 0.08 (1.84)* | 0.05 (2.02)** |
| *GovSize* | 0.08 (3.45)*** | 0.07 (1.75, 1.72)* | 0.06 (1.51) | 0.08 (3.13)*** |
| *FinDev* | 0.04 (1.86)* | 0.01 (0.22, 0.20) | 0.00 (0.13) | - |
| *Ethno* | -0.01 (-0.21) | -0.10 (-1.82, -1.57) | -0.07 (-1.29) | -0.13 (-2.25)** |
| *Region* | -0.04 (-2.31)** | -0.03 (-1.18, -1.12) | -0.02 (-0.80) | - |
| *HumCap* | -0.11 (-3.59)*** | -0.03 (-0.64, -0.59) | 0.02 (0.62) | - |
| *Open* | 0.01 (0.39) | -0.02 (-0.45, -0.41) | -0.06 (-1.95)* | - |
| *PopSize* | 0.04 (1.75)* | -0.04 (-0.76, -0.72) | -0.06 (-2.16)** | - |
| *GdpLev* | -0.01 (-0.13) | 0.07 (1.64, 1.49) | 0.03 (0.63) | - |
| *Policies* | -0.04 (-1.92)* | **-0.12 (-3.47, -3.30)*** ** | **-0.09 (-2.48)** ** | **-0.13 (-6.02)*** ** |

| | | | | |
|---|---|---|---|---|
| *OLS* | 0.05 (3.07)*** | -0.03 (-1.38, -1.26) | -0.02 (-1.07) | - |
| *Gr&Aid* | 0.09 (3.91)*** | -0.06 (-1.31, -1.14) | -0.05 (-0.96) | - |
| *Gr&Savings* | 0.12 (2.53)** | -0.01 (-0.25, -0.20) | -0.01 (-0.12) | - |
| Adjusted $R^2$ | various | 0.32 | *na* | 0.31 |
| F-statistic | various | 5.31 (prob-value=0.00) | 565.72 (prob-value=0.00) | 12.12 (prob-value=0.00) |
| Cor(ob, fit) | various | 0.63 | 0.57 | 0.63 |
| *N* | various | 474 | 474 | 487 |

Notes: Significance at the 10%, 5% and 1% level are *, **, ***, respectively. Brackets report t-statistics using robust standard errors. Columns 2, 3 and 4 use the bootstrap with 1000 replications with replacement. Cor(ob, fit) is the correlation between observed and fitted partial correlations. Each group of alternate shaded variables in column 1 represents separate regressions. Bold figures denote robust findings. *na* means not available. Missing observations for some of the controls reduces the sample size for the MRA.

Column 2 present the results using a fixed effects MRA. A fixed effects meta-analysis model is appropriate when there is a common aid-growth effect that all studies are estimating (see Lipsey and Wilson 2001). If this is the case, then study results will differ because of sampling error and systematic differences due to the research design process. In a random effects meta-analysis model, study differences are assumed to result from both sampling error as well as *random* differences between studies.[30]

The random effects model is appropriate if a sample of empirical studies is used in a meta-analysis (as opposed to the entire population) and if the source of differences between studies cannot be identified. The results from the random effects MRA are presented in column 3.[31] The random effects and fixed effects results are fairly similar. This suggests that the apparent heterogeneity in the reported results is not random, but due to systematic differences in research design.[32]

Not surprisingly, many of the variables are not statistically significant. This is due to some controls having no effect at all on the reported partial correlations and in some cases it is due to multicollinearity, which is often a problem with MRA. Accordingly, Column 4 presents the fixed-effects MRA results after sequentially eliminating any statistically insignificant variables.[33]

30. The terms fixed effects and random effects in meta-analysis differ to those used in panel data analysis. See Appendix 2 for details.

31. To estimate the Random Effects model, we assume that the total variance in the aid-growth effects consists of variance due to sampling error, as well as variance due to other factors that are *randomly* distributed. We used the standard error of each partial correlation to calculate the variance due to sampling error, and we estimate the second variance term using the so-called iterative restricted maximum likelihood method, or REML (see Raudenbush 1994 for details).

32. Indeed, the random effects model shows that the between study variance ($\tau^2$ = 0.0091) is actually small. Further, Q is a test for heterogeneity, exploring whether the remaining variability is greater than would be expected from sampling error. For columns 2 and 4, the Q test confirms that there is no excess variability beyond sampling error. Q = 15.00 (prob-value >0.10) and 16.61 (prob-value >0.10), for columns 2 and 4 respectively.

33. A Wald test accepts removal of the redundant variables. $\chi^2$ = 15.87 with a prob-value of 0.98.

Table 6. Sensitivity of Meta-Regression Analysis, Aid-Growth Effect, All-Set

(Dependent variable = partial correlation)

| Variable | (1)<br>Col. (2)<br>Fixed | (2)<br>From Table 5<br>Col. (3)<br>Random | (3)<br>Col. (4)<br>Specific | (4)<br>Specific<br>Random<br>Effects | (5)<br>Without<br>C(1) &<br>C(2) | (6)<br>WLS<br>Quality | (7)<br>Cluster<br>Corresp.<br>To (1) | (8)<br>Permute<br>Fraction<br>Too high |
|---|---|---|---|---|---|---|---|---|
| | | | | *Robust findings* | | | | |
| Asia | 0.12 | 0.11 | 0.09 | 0.05 | 0.14 | 0.12 | 0.12 | 0.005 |
| | (2.45) | (2.62) | (2.60) | (2.18) | (2.48) | (2.59) | (2.04) | |
| 1970s | -0.12 | -0.07 | -0.12 | -0.08 | -0.11 | -0.12 | -0.12 | 0.034 |
| | (-2.81) | (-1.95) | (-3.91) | (-3.41) | (-2.38) | (-2.83) | (-1.67) | |
| 1980 | -0.14 | -0.12 | -0.07 | -0.09 | -0.16 | -0.14 | -0.14 | 0.007 |
| | (-1.91) | (-2.13) | (-1.75) | (-2.73) | (-2.41) | (-2.09) | (-1.30) | |
| EDA | -0.06 | -0.05 | -0.07 | -0.06 | -0.04 | -0.06 | -0.06 | 0.018 |
| | (-2.35) | (-1.89) | (-3.41) | (-3.02) | (-1.81) | (-2.58) | (-2.11) | |
| Aid·Institutions | -0.09 | -0.08 | -0.10 | -0.11 | -0.06 | -0.08 | -0.09 | 0.015 |
| | (-1.73) | (-1.90) | (-3.22) | (-3.76) | (-1.97) | (-1.85) | (-1.14) | |
| Policies | -0.12 | -0.09 | -0.13 | -0.08 | -0.05 | -0.13 | -0.12 | 0.001 |
| | (-3.30) | (-2.48) | (-6.02) | (-3.97) | (-1.59) | (-3.60) | (-2.67) | |
| Influence | 0.11 | 0.10 | 0.12 | 0.10 | - | 0.11 | 0.11 | 0.000 |
| | (3.65) | (3.75) | (5.87) | (5.33) | | (3.69) | (2.33) | |
| EDCC | -0.57 | -0.62 | -0.58 | -0.67 | - | -0.57 | -0.57 | 0.000 |
| | (-2.80) | (-7.70) | (-2.94) | (-10.04) | | (-3.18) | (-3.33) | |
| Applied Econ. | -0.12 | -0.13 | -0.14 | -0.13 | | -0.12 | -0.12 | 0.000 |
| | (-1.68) | (-2.55) | (-2.30) | (-3.34) | - | (-1.71) | (-2.04) | |
| Danida | 0.11 | 0.16 | 0.09 | 0.06 | | 0.10 | 0.11 | 0.000 |
| | (1.86) | (3.91) | (3.88) | (2.57) | - | (1.99) | (1.40) | |
| FDI | 0.08 | 0.08 | 0.09 | 0.07 | 0.14 | 0.08 | 0.08 | 0.023 |
| | (1.72) | (2.01) | (2.79) | (2.51) | (2.88) | (1.84) | (1.38) | |
| | | | | *Some support* | | | | |
| GovSize | 0.07 | 0.06 | 0.08 | 0.09 | -0.01 | 0.06 | 0.07 | 0.054 |
| | (1.72) | (1.51) | (3.13) | (3.54) | (-0.28) | (1.65) | (1.45) | |
| Single country | 0.26 | 0.56 | 0.27 | 0.47 | 0.22 | 0.26 | 0.26 | 0.000 |
| | (1.50) | (6.64) | (2.83) | (7.70) | (1.35) | (1.58) | (1.35) | |
| Instability | 0.10 | 0.07 | 0.12 | 0.12 | -0.07 | 0.10 | 0.10 | 0.069 |
| | (1.31) | (1.51) | (2.10) | (3.41) | (-1.46) | (1.71) | (1.31) | |
| Ethno | -0.10 | -0.07 | -0.13 | -0.10 | 0.08 | -0.10 | -0.10 | 0.122 |
| | (-1.57) | (-1. 29) | (-2.25) | (-2.75) | (1.61) | (-1.78) | (-1.76) | |
| Inflation | -0.08 | -0.06 | -0.07 | - | -0.09 | -0.08 | -0.08 | 0.072 |
| | (-1.80) | (-1.53) | (-2.84) | | (-2.62) | (-2.01) | (-1.31) | |
| Lag Used | 0.07 | 0.09 | - | 0.11 | 0.07 | 0.07 | 0.07 | 0.000 |
| | (1.41) | (2.93) | | (4.17) | (1.67) | (1.71) | (0.75) | |
| Capital | 0.09 | 0.10 | - | - | 0.15 | 0.09 | 0.09 | 0.004 |
| | (1.74) | (2.39) | | | (2.25) | (2.01) | (1.33) | |

Notes: t-statistics reported in brackets use bootstrapped standard errors, except for columns 6 and 7. Columns 1 to 3 reproduce key findings from Table 5, columns 2 to 4, respectively. Column 5 excludes the publication and author variables. Column 6 uses WLS with journal impact factor as weights. Column 7 uses clustered data analysis. Column 8 reports permutation test results.

### *4.3 Sensitivity analysis*

Table 6 presents the results of sensitivity analysis for the key findings from Table 5. To assist comparability, columns 1 to 3 of Table 6 reproduce columns 2 to 4 of Table 5. The specific version of the random effects model is reported in column 4.[34] The fixed effects model was

34. A Wald test accepts removal of the redundant variables. $\chi^2 = 1.21$ with a prob-value of 0.20.

reestimated without the publication outlet (C1) and author (C2) variables. These results are presented in column 5. In column 6 we report the results of using a weighted least squares (WLS) regression using journal Impact Factors as weights (as reported by the Social Science Citation Index). This assigns greater weights to papers published in higher ranking journals.[35]

As an alternative to using the bootstrap to derive standard errors we used also clustered data analysis (see Hox 2002). These results are reported in column 7. An increasingly popular test in meta-analysis is the permutation test (Higgins and Thompson 2004). The permutation tests were carried out by *randomly* reallocating the aid-growth partial correlations to sets of covariates. The MRA was then re-estimated, with the reallocation and re-estimation repeated 1,000 times. We then compared the number of times the test statistic from the random reallocations equals or exceeds our initial test statistics. Column 8 reports these results. For example, the permutation test for the EDA variable was found to equal or exceed the initial test statistics in only 1.8% of the 1,000 estimates. This means that the negative and statistically significant coefficient for EDA is highly unlikely to result from chance (0.018 <0.100). We conclude from column 8 that the key results reported in Table 5 are not spurious.

As can be seen from Table 6, the results are generally robust to estimation and specification. There are no sign reversals and most variables remain statistically significant. Several variables are statistically significant as well as robust to the specification and estimation of the MRA.

### *4.4 Summary of results*

(C1) ***Publication outlet:*** Of the six journal variables, only two are statistically significant. Both *EDCC and Applied Economics* have a negative sign that is always statistically significant. This means that *ceteris paribus* these journals publish findings that have smaller positive or larger negative partial correlations.

Column (6) reports virtually the same result as does column (1). It thus appears that as regards to the substantial results, the quality control of the market for economic papers does not matter at all. This might be due to the particular field considered, but it is worth noting.

(C2) ***Groups/influence:*** The most significant result here is that writers from the *Danida* group consistently find results that are more pro-aid (with about 0.1 points larger aid-growth partial correlations ) than others.

---

35. The sample size is reduced further in this case, as some journals do not have an Impact Factor and these are excluded from the WLS analysis. Similarly, studies published in books and working papers are also excluded, by necessity.

It is also important that the *Influence* variable tends to be significant. Researchers tend to confirm the results of those they are associated with. The influence variable has a robust significant positive coefficient indicating that studies conducted by authors who receive comments/assistance from other authors publishing in the same field tend to report higher partial correlations, than those who do not acknowledge receiving feedback. This does not necessarily mean, however, that these studies (*Influence* and *Danida*) are biased. Indeed, they may be better constructed ones.

The other C2 group variables are not statistically significant. Results are not influenced by the author's gender. Importantly, authors' stated expectations do not have a quantifiable impact on the aid-growth effects. The World Bank authors variable has a positive coefficient in column 5 Table 5, but this is not a robust finding.

(C3) ***Data.*** The four decade variables *Y1960s* to *Y1990s* (with 2000 as the base) show that the aid-growth association reported was weaker in the 1970s and in the 1980s, where many countries borrowed heavily on the commercial market, and thus were less dependent on aid. Also, the 1970s were the period of the Oil Crisis which generated rather strong economic fluctuations that were independent of aid. This is also observationally equivalent with the notion that aid effectiveness has increased in recent years. Reported aid-growth effects were lower in the 1970s and 1980s then they were in the 1990s and in the new century.[36]

The use of EDA as a measure of aid leads to lower aid-growth effects. The relation between EDA and ODA data is: $EDA = a\ ODA + \varepsilon$, where $a \approx 0.42$ and $\varepsilon$ is small, as the correlation between the two aid series is 0.83, where both are available. If decision makers consider $\varepsilon$ to be random, we expect the partial correlations to be exactly the same. However, this is not likely, and the results then depend upon the time horizon of the decision makers. Imagine that the decision makers have a long time horizon. Then they recognize that loans have to be paid back. Hence they react much stronger to EDA data than to the ODA data. Thus the elasticities to the EDA data should be systematically larger. However, if the decision makers are myopic then they react to ODA data and do not care about the size of the gift element, which matters in the longer run only. Thus the elasticities to the ODA data should be systematically larger. We actually find that the elasticities to the ODA data are significantly higher and thus one more indication that decision makers are myopic.[37]

36. Note however that the fixed effects models indicate that the inclusion of data from the 1990s results in larger partial correlations than in the 2000s, but this finding is not robust.

37. This is the standard result in the literature on vote and popularity functions, see Nannestad and Paldam (1994). It is also the only really strong explanation for the phenomenon of debt crises, which seems to appear every 20-30 years in a surprising number of countries.

Regional differences in the aid-growth association are also detected. The inclusion of Asian economies in a sample increases the aid-growth effect. While there does not appear to be a direct aid-growth effect on *average* across nations, it appears that Asian countries experience higher aid-growth effects. This is an important factor that is a real phenomenon. Additionally, there is some evidence that single country studies find larger effects than studies that explore groups of countries. The explanation for this regional and country specific variation is beyond the scope of this paper, but is a promising area for further research.

The number of countries included in a sample does not have the expected positive coefficient. Indeed, it is negative in the specific version of the fixed effects model.

(C4) ***Conditionality***. We study the effect of the inclusion of three second order terms: aid·policy, aid·aid and aid·institutions. All three has the expected sign, but only the aid·institutions term is robust. This result is consistent with Doucouliagos and Paldam (2005a) who found that independent replications rejected aid conditionality on the basis of good policy or aid squared. If the aid·policy term is statistically significant in a conditionality study, then we would expect that it would affect the magnitude of the aid-growth effect. Our results show that the inclusion of aid·policy and aid·aid terms in a growth regression has no effect at all on the magnitude of the aid-growth effects. On the other hand, the aid·institutions variable is robust. The inclusion of aid·institutions conditionality results in smaller aid-growth effects. This appears to be an important indirect effect through which aid contributes positively to growth, and warrants further investigation.

(C5) ***Model formulation and controls***. Many of the variables are significant in some columns, but not all are robust throughout. While the aid·policy term has no effect on aid-growth effects, the inclusion of policy variables in a growth regression results in smaller direct aid-growth effects. This is consistent with aid having an indirect effect on policy. When both the aid-growth effect and policy variables are included, the aid-growth effect measures the direct effect of aid on growth. When policy is excluded, the aid-growth effect measures the total effect of aid. The negative coefficient on the policy variable in the MRA means that the direct effect is lower than the total effect. Hence, the MRA results are consistent with the notion that aid has a positive effect on good policy and that this then has an indirect positive effect on growth. That is, it is not the direct aid·policy conditionality interaction that is important, but the indirect effect of aid on policy. We find similar negative coefficients for inflation. If inflation has a detrimental effect on growth, then our results suggest that aid decreases inflation – another indication that aid assists good policy.[38]

The inclusion of FDI has a positive effect on aid-growth effects. If FDI has a positive effect on growth, then our results are consistent with aid reducing FDI.[39] A similar result holds for capital, although this is not robust. Hence, our results are consistent with the notion that aid may be a substitute for domestic and foreign sources of investment.[40] Positive coefficients are found also for government size and instability. Controlling for the share of government results in larger aid-growth effects, as predicted in section 3.[41] In section 3 we discussed also the political economy model and noted that the effect of aid on political stability was an empirical issue. Our findings of a positive coefficient on political instability suggests that the direct effect of aid on growth when political instability is added is higher than when political instability is ignored. Hence, the indirect effect of aid on growth through the political instability channel must be negative. That is, aid appears to increase political instability.

(C6) ***Estimation techniques.*** As mentioned at the end of section 3 there is no apparent difference in results according to the techniques used, once other study differences are controlled for: the choice of estimator leaves results unchanged. Using lagged values of aid results in similar results as using current levels of aid.

The MRA explains about one third of the variation in results, with a reasonably high degree of correlation between the predicted partial correlations and the observed partial correlations, which is good for this sort of analysis.

## 5. Summary

The AEL (aid effectiveness literature) has accumulated a large pool of evidence on aid-savings, aid-investment, aid-growth and aid-conditionality effects. The majority of the authors seem to agree that aid has a small positive effect on growth. The aim of this paper was to apply the methods of meta-analysis to the entire aid-growth effects literature to see if that

38. There is also some evidence that foreign aid has a positive indirect effect on growth working through an ethnographic fractionalization channel. Perhaps development aid eases some of the tensions created by fractionalization. However, this variable does not pass the permutation test, so it should be interpreted with caution.
39. The total effect is lower than the direct effect when FDI is included in a growth regression. This means that the indirect effect of aid working through FDI is negative.
40. Doucouliagos and Paldam (2006) show that aid does not increase investment significantly.
41. That is, aid increases government size and this has a negative effect on growth. So, the indirect effect of aid on growth is negative.

conclusion is justified. We used the population of 68 aid-growth studies as our data set. Our conclusions are depressing: Taking all available evidence accumulated over 40 years of research into consideration, we have to conclude that the AEL has failed to prove that the effect of development aid on growth is statistically significantly larger than zero. We are forced to conclude that aid has not, on average, achieved its stated aims of generating development.

We also investigated the variation in the available results. Some of this variation was found to be a direct outcome of data and specification differences. We found that journals, friends and institutional affiliation influence reported results. However, we found also that some of the variation was real, e.g. including Asian economies in a sample increases the reported effect of aid on growth.

The MRA results suggest that aid has a favorable effect on government policy (perhaps because donors reward good policy by giving more aid?) and tends to lower inflation. At the same time however, we find that aid displaces FDI, results in larger governments and increases political instability. Further research into these and other indirect channels through which aid may affect growth is clearly warranted.

## References (note also Appendix 1 listing the studies used in the meta-analysis)


Adams, D.C., Gurevitch J., Rosenberg, M.S., 1997. Resampling Tests for Meta-analysis of Ecological Data, *Ecology* 78, 1277-83

Aghion, P., Durlauf, S., eds., 2005. *Handbook of Economic Growth.* North-Holland, Amsterdam

Alesina, A., Dollar, D., 2000. Who gives foreign aid to whom and why? *Journal of Economic Growth* 5, 33-63

Ashenfelter, O., Harmon, C., Oosterbeek H., 1999. A review of estimates of the schooling/earnings relationship, with tests for publication bias. *Labour Economics*, *6*, 453-470

Barro, R.J., 1991. Economic growth in a cross section of countries. *Quarterly Journal of Economics* 106, 407-43

Barro, R.J., Sala-i-Martin, X., 1995, 2004. *Economic Growth*. MIT Press, Cambridge, MA

Bauer, P.T., 1971. *Dissent on development: Studies and debates in development economics*. Weidenfeld and Nicolson, London. Sums up several older papers

Card, D., Krueger A.B., 1995. Time-Series Minimum-Wage Studies: A meta-analysis. *American Economic Review* 85, 238-43

Cassen, R., *et al.*, 1986, 1994. *Does Aid Work?* Clarendon, Oxford

Chang, C. C., Fernandez-Arias, E., Serven, L., 1998. *Measuring aid flows: A new approach*. World Bank. URL: <http://www.worldbank.org/research/growth/ddaid.htm>

Chenery, H.B., Strout, A.M., 1966. Foreign assistance and economic development. *American Economic Review* 56, 679-733

Collier, P., Hoeffler, A., 2000. Greed and Grievance in Civil War. The World Bank, Policy Research WP 2355

Dobson, S., Ramlogan, C., Strobl, E. Why Do Rates of (b)-Convergence Differ? A meta-regression analysis. *Scottish Journal of Political Economy*. 53, 153-73.

Doucouliagos, C.(H.), 2005. Publication Bias In The Economic Freedom And Economic Growth Literature. *Journal of Economic Surveys* 19, 367-89

Doucouliagos, C.(H.), Laroche, P. 2003. What do Unions do to Productivity? A Meta-Analysis. *Industrial Relations* 42, 650-691

Doucouliagos, C.(H.), Ulubasoglu, M. 2006. Economic Freedom and Economic Growth: Does specification make a difference? *European Journal of Political Economy*, 22, 60-81

Doucouliagos, H., Paldam, M., 2005a. Conditional aid effectiveness. A meta study[42]

Doucouliagos, H., Paldam, M., 2005b. Aid effectiveness: The lessons from 40 years of research[42]

Doucouliagos, H., Paldam, M., 2006. Aid effectiveness on accumulation. A meta study. *Kyklos* 59, 227-54

Easterly, W., 2001. *The Elusive quest for growth. Economists' Adventures and Misadventures in the Tropics*. MIT Press, Cambridge MA

Efron, B, Tibshirani, R., 1993. *An Introduction to the Bootstrap*. Chapman & Hall, New York

Friedman, M., 1958. Foreign economic aid: Means and Objectives. *Yale Review* 47, 500-16

Gorg, H., Strobl, E., 2001. Multinational Companies and Productivity: a meta-analysis. *The Economic Journal*, Vol. 111, No. 475, pp. 723-39

Greene, W.D., 2002. *Econometric Analysis* (4th, edn.). Prentice Hall, London


42. Available from WP series, Department of Economics, University of Aarhus or http://www.martin.paldam.dk.

Hedges, L. V., Olkin, I, 1985, *Statistical Methods for Meta-Analysis*. Academic Press, Orlando, FL

Herbertsson, T.T., Paldam, M., 2005. Does development aid help poor countries catch up? An analysis of the basic relations[42]

Higgins, J.P.T., Thompson S.G., 2004. Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine* 23: 2663-82

Hox, J.J. 2000. *Multilevel Analysis: techniques and applications*. Mahwah, NJ: Lawrence Erlbaum

Hunter, J., Schmidt, F., 2004. *Methods of Meta-Analysis: correcting error and bias in research findings*. Sage, London

IBRD-OED. Evaluation Results. (Annual) World Bank Operations Evaluation Department, Washington DC

Jarrell, S., Stanley, T.D., 2004. Declining Bias and Gender Wage Discrimination? A Meta-regression analysis. *Journal of Human Resources*, 39, 828-38

Jepma, C.J., 1995. On the effectiveness of developing aid. Manuscript, DVL/OS, Den Haag

Levine, R., Renelt, D., 1992. A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review* 82, 942-63

Lipsey, M.W., Wilson, D.B., 2001. *Practical meta-analysis*. Applied Social Research Methods Series. Volume 49. London: Sage Publications

Longhi, S., Nijkamp, P., Poot, J., 2005. A Meta-Analytic Assessment of the Effect of Immigration on Wages. *Journal of Economic Surveys*, 19, 451-78.

McCloskey, D.N., 1998. *The Rhetoric of Economics*. University of Wisconsin Press, Madison, WI.

Mookerjee, R. 2006. A meta-analysis of the export growth hypothesis. *Economics Letters* 91, 395-401.

Mosley, P., 1986. Aid effectiveness: the micro- macro paradox. *Institute of Development Studies Bulletin* 17, 214-25.

Nannestad, P., Paldam, M., 1994. The VP-Function. A Survey of the Literature on Vote and Popularity Functions after 25 Years. *Public Choice* 79, 213-45

Paldam, M., 1997. Rent Seeking and Dutch Disease. An Essay on Greenland. *European Journal of Political Economy* 13, 591-614

Paldam, M., 2005. A look at the raw data for aid and growth. Posted at <http://www.martin.paldam.dk>

Raudenbush, S.W. 1994. Random effects models. In *The Handbook of Research Synthesis*. Edited by Harris Cooper and Larry V. Hedges. New York: Russell Sage Foundation

Roberts, C.J., Stanley, T.D. 2005. (eds.) *Meta-Regression Analysis: issues of publication bias in economics.* Basil Blackwell, London

Stanley, T.D., 2001. Wheat From Chaff: meta-analysis as quantitative literature review. *The Journal of Economic Perspectives* 15, 131-50

Stanley, T.D., 2005. Beyond Publication Bias. *Journal of Economic Surveys* 19, 309-45

Stanley, T.D., Jarell, S.B., 1998. Gender Wage Discrimination Bias? a meta-regression analysis. *The Journal of Human Resources* 33, 947-73

Sturm, J.-E., Haan J.d., 2005. Determinants of long-term growth: New results applying robust estimation and extreme bounds analysis. *Empirical Economics* 30, 597–617

Sutton, A.J., Duval, S.J., Tweedie, R.L., Abrams, K.R. and Jones, D.R. 2000. Empirical assessment of effect of publication bias on meta-analyses, *British Medical Journal,* Vol. 320, pp. 1574-1577

Tarp, F., Hjertholm, P., eds., 2000. *Foreign aid and development. Lessons learnt and directions for the future.* Routledge Studies in Development Economics, London

Tsikaka, T.M., 1998. Aid effectiveness: A survey of the recent empirical literature. *IMF paper on Policy Analysis and Assessment*

Weisskopf, T.E., 1972a. An econometric test of alternative constraints on the growth of underdeveloped countries. *Review of Economics and Statistics* 54, 67-78

Weisskopf, T.E., 1972b. The impact of foreign capital inflow on domestic savings in underdeveloped countries. *Journal of International Economics* 2, 25-38

White, H., 1992. What do we know about aid's macroeconomic impact? An overview of the aid effectiveness debate. *Journal of International Development* 4, 121-37

White, H., ed., 1998. *Aid and Macroeconomic Performance: Theory, Empirical Evidence and Four Country Cases.* (International Finance and Development Series) Palgrave Macmillan, London

World Bank, 1998. *Assessing aid: What works, what doesn't, and why*. A Policy Research Report. Oxford UP, New York

## Appendix 1: Aid-Growth Studies Used in Meta-Analysis

Only papers in English available till 1/1 2005 are included.

Amavilah, V.H., 1998. German aid and trade versus Namibian GDP and labour productivity. *Applied Economics* 30, 689-95

Boone, P., 1994. The impact of foreign aid on savings and growth. WP London School of Economics

Bowen, J.L., 1995. Foreign aid and economic growth: An empirical analysis. *Geographical Analysis* 27, 249-61. Estimates also in Bowen, J.L., 1998. *Foreign aid and economic growth: A theoretical and empirical investigation*. Ashgate, Aldershot, UK

Brumm, H.J., 2003. Aid, policies and growth: Bauer was right. *Cato Journal* 23, 167-74

Burnside, C., Dollar, D., 2000. Aid, policies and growth. *American Economic Review* 90, 847-68 (Working paper available from World Bank since 1996)

Burnside, C., Dollar, D., 2004. Aid, policies and growth: Reply. *American Economic Review* 94, 781-84 (reply to Easterly, Levine and Roodman, 2004)

Campbell, R., 1999. Foreign aid, domestic savings and economic growth: Some evidence from the ECCB area. *Savings and Development* 23, 255-78

Chauvet, L., Guillaumont, P., 2004. Aid and growth revisited: Policy, economic vulnerability and political instability. pp 95-109 in Tungodden, B., Stern, N., Kolstad, I., eds, 2004. *Toward Pro-Poor Policies - Aid, Institutions and Globalization*. World Bank/Oxford UP

Collier, P., Dehn, J., 2001. Aid, shocks, and growth. WP 2688 World Bank Policy Research

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

Collier, P., Hoeffler, A., 2004. Aid, policy and growth in post-conflict societies. *European Economic Review* 48, 1125-45

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

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

Dayton-Johnson, J., Hoddinott, J., 2003. Aid, policies and growth, redux. WP Dalhousie University

Denkabe, P., 2004. Policy, aid and growth: A threshold hypothesis. *Journal of African Finance and Economic Development* 6, 1-21 (WP version used)

Dowling, J.M., Hiemenz, U., 1983. Aid, savings, and growth in the Asian region. *The Developing Economies* 21, 4-13

Durbarry, R., Gemmell, N., Greenaway, D., 1998. New evidence on the impact of foreign aid on economic growth. Credit WP Univ. of Nottingham

Easterly, W., 2003. Can foreign aid buy growth? *Journal of Economic Perspectives* 17, 23-48

Easterly, W., Levine, R., Roodman, D., 2004. Aid, policies, and growth: Comment. *American Economic Review* 94, 774-80 (Comment to Burnside and Dollar, 2001)

Economides, G., Kalyvitis, S., Philippopoulos, A., 2004. Does foreign aid distort incentives and hurt growth? Theory and evidence from 75 aid-recipient countries. WP Athens Univ. of Economics and Business

Fayissa, B., El-Kaissy, M., 1999. Foreign aid and the economic growth of developing countries (LDCs): Further evidence. *Studies in Comparative International Development* Fall, 37-50

Giles, J.A., 1994. Another look at the evidence on foreign aid led economic growth. *Applied Economics Letters* 1, 194-99

Gomanee, K., Girma, S., Morrissey, O., 2002. Aid and growth: Accounting for the transmission mechanisms in Sub-Sahara Africa. Credit WP Univ. of Nottingham

Gounder, R., 2001. Aid-growth nexus: Empirical evidence from Fiji. *Applied Economics* 33, 1009-19

Griffin, K.B., Enos, J.L., 1970. Foreign assistance: Objectives and consequences. *Economic Development and Cultural Change* 18, 313-27

Guillaumont, P., Chauvet, L., 2001. Aid and performance: A reassessment. *Journal of Development Studies* 37, 66-92

Gulati U.C., 1976. Foreign aid, savings and growth: Some further evidence. *Indian Economic Journal* 24, 152-60

Gulati U.C., 1978. Effects of capital imports on savings and growth in less developed countries. *Economic Inquiry* 16, 563-69

Gupta, K.L., 1975. Foreign capital inflows, dependency burden, and saving rates in developing countries: A simultaneous equation model. *Kyklos* 28, 358-74

Gupta, K.L., Islam, M.A., 1983. *Foreign Capital, Savings and Growth – an International Cross-Section Study*. Dordrecht, Reidel Publishing Company

Gyimah-Brempong, K., 1992. Aid and economic growth in LDCs: Evidence from Sub-Saharan Africa. *Review of Black Political Economy* 20, 31-52

Hadjimichael, M.T., Ghura, D., Mühleisen, M., Nord, R., Ucer, E.M., 1995. Sub-Saharan Africa: Growth, savings, and investment, 1986-93. IMF Occasional Paper, No. 118

Hansen, H., Tarp, F., 2000. Aid effectiveness disputed. *Journal of International Development* 12, 375-398. Also pp 103-128 in Tarp, F., Hjertholm, P., eds. 2000. *Foreign aid and development. Lessons learnt and directions for the future*. Routledge Studies in Development Economics: London

Hansen, H., Tarp, F., 2001. Aid and growth regressions. *Journal of Development Economics* 64, 547-70

Hudson, J., Mosley, P., 2001. Aid policies and growth: In search of the Holy Grail. *Journal of International Development* 13, 1023-38

Islam, M.A., 1992. Foreign aid and economic growth: An econometric study of Bangladesh. *Applied Economics* 24, 541-44

Jensen, P.S., Paldam, M., 2006. Can the two new aid-growth models be replicated? *Public Choice* 127, 147–75

Kellman, M., 1971. Foreign assistance: Objectives and consequences: Comments (to Griffin and Enos, 1970). *Economic Development and Cultural Change* 20, 142-54

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

Landau, D., 1986. Government and Economic Growth in the less Developed Countries: An empirical study for 1960-1980. *Economic Development and Cultural Change* 35, 35-75

Landau, D., 1990. Public choice and economic aid. *Economic Development and Cultural Change* 38, 559-75

Larson, J.-D., 2001. An updated analysis of Weisskopf's savings-dependency theory. *Review of Development Economics* 5, 157-67

Lensink, R., 1993. Recipient government behavior and the effectiveness of development aid. *De Economist* 141, 543-62

Lensink, R., Morrissey, O., 2000. Aid instability as a measure of uncertainty and the positive impact of aid on growth. *Journal of Development Studies* 36, 30-48

Lensink, R., White, H., 2001. Are there negative returns to aid? *Journal of Development Studies* 37, 42-65

Levy, V., 1988. Aid and growth in Sub-Saharan Africa: The recent experience. *European Economic Review* 32, 1777-95

Lu, S., Ram, R., 2001. Foreign Aid, government policies, and economic growth: Further evidence from cross-country panel data for 1970-1993. *Economia Internazionale/International Economics* 54, 15-29

Mahdavi, S., 1990. The effects of foreign resource inflows on composition of aggregate expenditure in developing countries: A seemingly unrelated model. *Kyklos* 43, 111-37

Mbaku, J.M., 1993. Foreign aid and economic growth in Cameroon. *Applied Economics* 25, 1309-14

Moreira, S.B., 2003. Evaluating the impact of foreign aid on economic growth: A cross-country study (1970-1998). WP for 15th Annual Meeting on Socio-Economics, Aix-en-Provence, France

Mosley, P., 1980. Aid, savings and growth revisited. *Bulletin of the Oxford University Institute of Economics and Statistics* 42, 79-95

Mosley, P., Hudson, J., Horrell, S., 1987. Aid, the public sector and the market in less developed countries. *Economic Journal* 97, 616-41

Mosley, P., Hudson, J., Horrell, S., 1992. Aid, the public sector and the market in less developed countries: A return to the scene of the crime. *Journal of International Development* 4, 139-50

Most, S.J., Berg, H.v.d., 1996. Growth in Africa: Does the source of investment financing matter? *Applied Economics* 28, 1427-33

Murthy, V.N.R., Ukpolo, V., Mbaku, J.M., 1994. Foreign aid and economic growth in Cameroon: Evidence from cointegration tests. *Applied Economics Letters* 1, 161-63

Ovaska, T., 2003. The failure of development aid. *Cato Journal* 23, 175-88

Papanek, G.F., 1973. Aid, foreign private investment, savings, and growth in less developed countries. *Journal* *of Political Economy* 81, 120-30

Ram, R., 2003. Roles of bilateral and multilateral aid in economic growth of developing countries. *Kyklos* 56, 95-110

Ram, R., 2004. Recipient country's "policies" and the effect of foreign aid on economic growth in developing countries: Additional evidence. *Journal of International Development* 16, 201-11

Rana, P.B., Dowling, J.M., 1988. The impact of foreign capital on growth: Evidences from Asian developing countries. *The Developing Economies* 26, 3-11

Reichel, R., 1995. Development aid, savings and growth in the 1980s: A cross-section analysis. *Savings and Development* 19, 279-96

Roodman, D., 2004. The anarchy of numbers: Aid, development and cross-country empirics. WP 32 Center for Global Development

Shukralla, E.K., 2004. Aid, incentives, polices, and growth: Theory and a new look at the empirics. WP Western Michigan University

Singh, R.D., 1985. State intervention, foreign economic aid, savings and growth in LDCs: Some recent evidence. *Kyklos* 38, 216-32

Snyder, D.W., 1993. Donor bias towards small countries: An overlooked factor in the analysis of foreign aid and economic growth. *Applied Economics* 25, 481-88

Stoneman, C., 1975. Foreign capital and economic growth. *World Development* 3, 11-26

Svensson, J., 1999. Aid, growth and democracy. *Economics and Politics* 11, 275-97

Tebouel, R., Moustier, E., 2001. Foreign aid and economic growth: The case of the countries south of the Mediterranean. *Applied Economics Letters* 8, 187-90

## Appendix 2: Introduction to meta techniques, especially to the tests used

Meta-analysis uses both descriptive statistics and significance tests, which are developed for the purpose of research synthesis. Note especially that the significance tests have to take into account that all studies are based on a common pool of available macro data.

**Descriptive statistics**

Two methods are in general use: Counts of coefficients with different signs and significance and calculations of various averages.

*A1 Vote counting*

All existing reviews of the aid-growth literature have either explicitly or implicitly used vote counting. Vote counting is similar to Extreme Bounds Analysis (EBA). In EBA, researchers use the same data set and explore the impact of different control variables (see Barro and Sala-i-Martin 2004). The difference here is that rather than the same sample being used with different specifications, we use different samples as well as different specifications. This is effectively a meta-extreme bounds analysis (MEBA).

However, counting the number of signs should not be given too much weight, as it does not provide a method for research synthesis. Moreover, it ignores information provided by the confidence intervals. For example, Ram's (2004) estimate of the aid-growth elasticity has a 95% confidence interval of -0.37 to +0.33, while Economides *et al.* (2004) estimate a confidence interval of +0.16 to +0.94. From Ram we conclude that there is a zero elasticity and from Economides *et al.* that there is a positive elasticity. The two intervals do, however, intersect – they both agree that it is possible that the elasticity may be between +0.16 to +0.33. With meta-analysis we can combine all studies and avoid the potential problems of sign counting.

*A2 Average effects*

The effect between two variables (holding other effects constant) established by a literature can be derived as a weighted average of the associated estimates:

(1A) $$\varepsilon = \sum [N_i \varepsilon_i] / \sum N_i$$

where ε is the *standardized* effect (elasticity or partial correlation) from the i$^{th}$ study and N is the sample size.

**Regression based tests**

The data for the two following tests are a set of *n* estimates of the same effect, *ε*, with the associated tests statistics ($t_i$, $s_i$, $d_i$), where $t_i$ is the t-statistics; $s_i$, is the standard error; $d_i$ is the degrees of freedom of the estimate. All *n* estimates use variants of the same estimation equation and sub-samples of the same data. Both tests use the population of observations and are somewhat robust to data mining.

*A5 Meta-Significance Testing: The MST (Card and Kreuger 1995; Stanley 2001; 2005)*

The idea is that a connection between two variables, such as foreign aid and economic growth, should exhibit a positive relationship between the natural logarithm of the absolute value of the t-statistic and the natural logarithm (ln) of the degrees of freedom in the regression:

(3A) $$\ln |t_i| = \alpha_0 + \alpha_1 \ln df_i + u_i$$

As the sample size for the i$^{th}$ study rises, the precision of the coefficient estimate for the i$^{th}$ study rises also, i.e., standard errors fall and t-statistics rise. Stanley (2005) shows that the slope coefficient in equation (3A) offers information on the existence of genuine empirical effects, publication bias, or both. If $\alpha_1 = 0$, then there is no association between the two variables of interest. If $\alpha_1 < 0$, the estimates are contaminated by selection effects, because t-statistics fall as sample size rises. That is, studies with smaller samples report larger t-statistics, indicating that authors mine smaller samples in order to increase the prospects of publication. If $\alpha_1 > 0$, there is a genuine association between aid and economic growth, since t-statistics rise as sample size increases.

*A6 Funnel-Asymmetry Testing: FAT (Egger et al. 1997; Stanley 2005)*

Smaller samples have larger standard errors. If publication bias is absent from a literature, no association between a study's reported effect and its standard error should appear. However, if there is publication bias, smaller studies will search for larger effects in order to compensate for their larger standard errors, which can be done by modifying specifications, functional form, samples and even estimation technique. FAT is given by:

(4A) $$\varepsilon_i = \beta_0 + \beta_1 s_i + \upsilon_i$$

where $\varepsilon_i$ is the *standardized* effect, and $s_i$ is its associated standard error. Since the explanatory variable in equation (4A) is the standard error, heteroscedasticity is likely to be a problem. Equation (4A) (from Stanley 2005) is corrected for heteroscedasticity by dividing it by the associated standard error. This produces equation (5A):

(5A) $$t_i = \varepsilon_i / s_i = \beta_1 + \beta_0 (1/s_i) + v_i$$

If publication bias is present, the constant, $\beta_1$, in equation (5A) will be statistically significant.

*A7 Meta-Regression Analysis*

The meta-regression model (known as MRA) has been developed to analyze the multi-dimensional nature of the research process. The impact of specification, data and methodological differences can be investigated by estimating an MRA of the following (linear) form:

(6A) $$\varepsilon_i = \alpha + \gamma_1 X_{i1} + \ldots + \gamma_k X_{ik} + \delta_1 K_{i1} + \ldots + \delta_n K_{in} + u_i$$

where $\varepsilon_i$ is the standardised effect derived from the i$^{th}$ study (in our study we use the partial correlation, r), $\alpha$ is the constant, $X_j$ are dummy variables representing characteristics associated with the i$^{th}$ study, $K_j$ are continuous variables associated with the i$^{th}$ study, $\gamma$ and $\delta$ are the unknown regression coefficients, and $u_i$ is the disturbance term, with usual Gaussian error properties.

Equation 6A is a fixed effects MRA and assumes that variation in $\varepsilon_i$ can be explained by sampling error and *systematic* differences between studies (the $X$ and $K$ study characteristics variables). The random effects MRA is given by:

(7A) $$\varepsilon_i = \alpha + \gamma_1 X_{i1} + \ldots + \gamma_k X_{ik} + \delta_1 K_{i1} + \ldots + \delta_n K_{in} + u_i + e_i$$

Equation 7A assumes that in addition to sampling error, the source of some of the variation in $\varepsilon_i$ is due to *random* differences among studies that cannot be identified.The regression coefficients in 6A and 7A quantify the impact of specification, data and methodological differences on reported study effects ($\varepsilon_i$). Both the MST and FAT tests can be combined with the MRA.