19-10-06

Conditional aid effectiveness: A meta study

Hristos Doucouliagos, Department of Economics, Deakin Univ., Melbourne, Australia*

Martin Paldam, Department of Economics, Univ. of Aarhus, Aarhus, Denmark#

Abstract:

The AEL (aid effectiveness literature) is the body of about 100 empirical studies of the effect of development aid on savings, investment and growth. It tries to catch the effects by a small set of formally homogeneous models. Thus, the AEL is ideal for meta-analysis. A third of AEL analyzes conditional models where aid effectiveness depends upon z, so that aid only works for a certain range of the variable. The key term in this family of AEL models is thus an interaction term of z times aid. The leading candidates for z are a good policy index and aid itself. In this paper, meta-analysis techniques are used (i) to determine whether the AEL has established the said interaction terms, and (ii) to identify some of the determinants of the differences in results between studies. We find no support for conditionality with respect to policy, while conditionality regarding aid itself is dubious.

JEL.: B2, F35, O35

Keywords: Aid effectiveness, meta study, economic growth, policy conditionality

Acknowledgements: Our cooperation was supported by the Aarhus University Research Foundation. Pia Wichmann Christensen has been a very competent research assistant. This paper benefited from ideas and comments from Peter Sandholt Jensen and Mehmet Ulubasoglu. It has also benefited from thorough discussions of the papers in the project at seminars at the University of Aarhus, Deakin University, Hendrix College and the Kiel Institute of World Economics, as well as from presentations at the Public Choice Society 2006 Meetings in New Orleans and Turku.

* Post: School of Accounting, Economics and Finance, Deakin University, 221 Burwood Highway, Burwood, Victoria 3125, Australia. Mail: douc@deakin.edu.au.

Post: School of Economics and Management (building 322), Aarhus University, 8000 Aarhus C, Denmark. Mail: mpaldam@econ.au.dk.

"… we, as fallible researchers, are trying to glimpse through the opaque window of imperfect empirical studies" (Rubin, 1990, p. 157)

1. Introduction: A research process driven by two challenges

Development aid is meant to contribute to the economic development of poor countries. This is certainly a very noble and desirable purpose. However, the correlation between the share of development aid and economic growth is essentially zero.[1] The zero correlation result has been known for 30 years, and it has remained *the old challenge* in the field.

From a development perspective, the zero correlation result is highly undesirable, and many have also found it implausible and risen to the challenge, by searching for a model that allows the data to tell a "nicer" story. The amount, magnitudes and variability of the data for aid and growth make them ideal for this kind of research. Research is often driven by challenges, and it has generated the AEL, Aid Effectiveness Literature, of almost 100 papers.[2] The present paper is part of a meta study of the process of research in the AEL.

Research is a process of truth revelation, which works through a mixture of *innovations* (in data, models and estimators) and *independent replications* as new data becomes available.[3] A single study rarely resolves an important issue in any science. Trust has to be build by replication. Consequently, we need techniques that can analyze the evidence contained in a sequence of studies. Unfortunately, by definition traditional qualitative reviews cannot be used to make statistical inferences from the available evidence and, consequently, they are prone to subjective interpretation (see Stanley 2001 and Hunter and Schmidt 2004). Meta-analysis techniques can be applied to study an empirical literature, which analyzes the *same effect* with *methods* that are so similar that the differences *can be coded*.

In macroeconomics, replications are rarely on fully independent samples. In addition to new data becoming available, new estimators are introduced, and models are often amended. Add also sampling error, and the outcome is often apparent differences in findings between studies. Meta-analysis is developed for quantitative summaries of partly dependent

1. See Herbertsson and Paldam (2005), Rajan and Subramanian (2005) and Doucouliagos and Paldam (2006b), as well as Figure 2 in this paper.
2. Till 1/1 2005 the AEL had reached 97 papers listed in Appendix 2 of Paldam and Doucouliagos (2006a). Appendix 2 lists the 31 of these papers that estimate conditional models.
3. A study is an *independent replication* if it is conducted by a new researcher on new data. *Dependent replication* is by a new researcher on the same data, or the same researcher on new data. The AEL has problems with independent replication, but not with dependent replication.

results. Meta-analysis is used to draw statistical inferences from the available pool of empirical studies, and to identify the factors that lead to differences in results across studies.

In the introduction to their textbook on meta-analysis Hunter and Schmidt (2004, p. xxxi) write that their "book examines in detail the severe distorting effects of statistical, measurement, and other methodological artifacts on the outcome of the *individual study*." If the reader studies Figures 4-6 below, this point is illustrated by our study as well.

The meta-approach is an *alternative* to the standard methods of sensitivity and robustness analysis. Rather than relying on a single study, meta-analysis commences with all available studies. For example, in the AEL, extensive sensitivity analysis has been reported by authors such as Easterly *et al.* (2004), Jensen and Paldam (2006) and especially Roodman (2004). However, we assess that at least 100 man-years of research has been invested into the AEL. This means that the pattern in the many results contains a lot of valuable information that is worth considering. Since we analyze the process of truth finding, it is important that our analysis covers all papers (and all estimates) that have been made available in the field.

A meta study asks two questions:[4]

Q1: Do the estimates of the effect *converge* to something we can term truth?

Q2: Can we identify the main *innovations* causing/preventing the convergence?

At present we deal with the newest family of aid effectiveness studies. It is the group of conditional aid studies, which started in 1995, and had grown to 31 papers at the start of 2005, where our data collection stops. The common feature of these studies is that they accept the zero correlation result, but take it to mean that aid works in some cases and fails in others. Hence, *the new challenge* of the AEL is to find the *condition* that determines when the good outcome results.

The operational meaning of the term *conditional* in the AEL is that the estimating equation contains a second order term, where aid is multiplied with another variable termed the condition. In growth empirics the term conditional normally means that the estimate contains a set of variables, which controls the relation for country heterogeneity. This paper uses the word in the AEL sense. Thus we distinguish between a *condition* that enters multiplicatively with aid and *the control set* that controls the estimate of country heterogeneity and other "disturbing" factors.

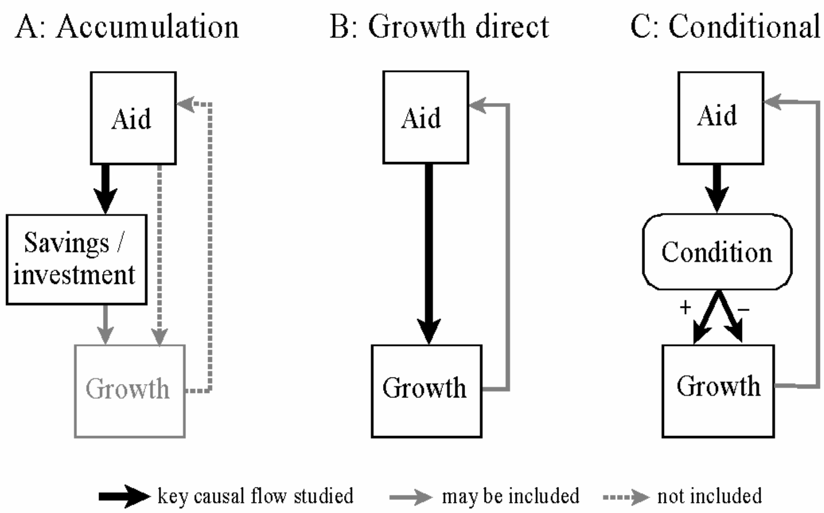
4. Meta studies can also be used to study the effect of priors and the existence of publication bias, see for example Roberts and Stanley (2005) and Doucouliagos and Paldam (2006c).

Section 2 classifies the AEL and discusses the theory of the conditionality family of models. Meta-analysis is used in section 3 to determine if the key effects of the two main conditionality models have been established, and in section 4 to explain the differences in reported results. Section 5 summarizes the rather sad findings of the paper. Appendix 1 is an introduction to meta techniques, especially to the tests used. Appendix 2 lists the AEL.

2. From aid effectiveness to conditional aid effectiveness

A thorough search produced the 97 AEL papers listed in Doucouliagos and Paldam (2006a).[5] The papers bring many models, which can be divided by their causal structure, into 3 families as shown on figure 1. Note that half the papers bring models from more than one family.

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



A: 43 papers study the impact of aid on savings or investment; see Doucouliagos and Paldam (2006a). About ¾ of aid is crowded out by a fall in savings, mainly due to increases in public consumption. The remaining ¼ causes increasing investments. Unfortunately, these results are so variable that they have remained statistically insignificant. The result that approximately ¾ of the marginal effect of aid is increased public consumption is used below.

5. Extensive searches of Econlit, Proquest, Web of Science and Google were undertaken, from which we could track citations backward. Papers available only after 1st of January 2005 are not included.

B: 68 papers contain a total of 543 *direct estimates*, using reduced form models of the effect of aid on growth. They are covered in Doucouliagos and Paldam (2006b), which shows that the 543 estimates of the direct effect of aid on growth reported scatter considerably, and add up to a small positive, but insignificant effect on growth. Thus the zero correlation result has yet to be overcome.

C: 31 papers – see Appendix 2 – contain *conditional estimates*, where the effect of aid on growth depends upon a conditional variable *z*, which is scaled so that if *z* is positive, the result is positive growth, and vice versa if *z* is negative. This is the family analyzed in this paper.

The AEL was started in the early 1970s by papers in the A-group. The early studies found no effect of aid on accumulation, and the AEL then moved on to the B-group papers, where most of the research in the 1980s and 1990s was done. The C-group started in 1995. It is where most of the action has been since then. This wave of papers is still strong, so we are discussing an ongoing process.

2.1 Conditional aid effectiveness: The models

In the mid 1990s the C-family of models appeared. The C-family of studies is based on the idea that aid ineffectiveness may be due to aid having a positive effect on growth in some countries and a negative effect in others, so that aid is *conditionally* effective. Till now, three conditions have been proposed, leading to three models of which only the first two are sufficiently homogeneous for a meta study:

(1) *Good Policy Model*: Aid works if the recipient country pursues good polices, and is harmful in countries pursuing bad policies. The model was proposed by Burnside and Dollar (1995, 2000), and has been developed by a group in the World Bank. The model has been analyzed in 22 more papers.[6]

(2) *Medicine model*: Aid works if given in moderation, and harms if taken in excess, just like most medicine. This model was first proposed by Hadjimichael, Ghura, Mühl-eisen, Nord and Ucer (1995), but it has mainly been developed and publicized by the

6. They are Svensson (1999); Hansen and Tarp (2000; 2001); Collier and Dehn (2001); Dalgaard and Hansen (2001); Guillaumont and Chauvet (2001); Hudson and Mosley (2001); Lensink and White (2001); Lu and Ram (2001); Collier and Dollar (2002); Brumm (2003); Dayton-Johnson and Hoddinott (2003); Burnside and Dollar (2004); Chauvet and Guillaumont (2004); Collier and Hoeffler (2004); Dalgaard, Hansen and Tarp (2004); Denkabe (2004); Easterly, Levine and Roodman (2004); Jensen and Paldam (2004); Ram (2004); Roodman (2004); and Shukralla (2004).

group of Tarp, Dalgaard and Hansen from DERG at Copenhagen University. Most members of this group are associated with Danida (the Danish Aid Agency). This model has been further analyzed in 15 papers.[7]

(3) *Institutions models*: A residual of 10 papers contains models that condition for various institutions: Two studies link aid with *GDP* (Bowen 1995; Svensson 1999); two papers condition for *democracy* (Svensson 1999; Kosack 2002); one for *external vulnerability* (Guillaumont and Chauvet 2001); two for *quality of institutions* (Collier and Dehn 2001; Collier and Dollar 2002); one for *trade openness* (Teboul and Moustier 2001); and one for *economic freedom* (Brumm 2003). Finally one conditions for *political instability* (Chauvet and Guillaumont 2004).

Several of the papers analyze two (or more) models. While the future development may be within the institutional models, the two first models have been the prominent ones till now, each leading to a stream of papers. This in particular applies to the Good Policy Model, which dominated the macroeconomic aid discussion for almost a decade from 1995.

It is noteworthy that the two most prominent models are advocated by a group which is closely associated with an aid agency. It is clearest in the case of the Good Policy Model, where most proponents are World Bank staff, and the model was advocated to a broader audience in a World Bank Policy Research Report (see World Bank 1998). The DERG group that is the main advocate for the Medicine Model is almost as closely associated with *Danida* (Danish Aid Agency).[8] Thus, these two models can be said to have an *institutional home*.[9]

2.2 The data

Most of these papers are estimated on the standard ODA data for Official Development Aid as compiled by the OECD. It contains all gifts and loans on concessional terms (which contain a grant element of at least 25%) from the OECD countries to LDCs (Less Developed Countries). Chang, Fernandez-Arias and Serven (1998) has introduced a new set of EDA-data

---

7. They are Durbarry, Gemmell and Greenaway (1998); Hansen and Tarp (2000; 2001); Dalgaard and Hansen (2001); Hudson and Mosley (2001); Lensink and White (2001); Collier and Dollar (2002); Gomanee, Girma and Morrisey (2002); Denkabe (2003); Moreira (2003); Ovaska (2003); Collier and Hoeffler (2004); Dalgaard, Hansen and Tarp (2004); Jensen and Paldam (2004); and Roodman (2004).

8. The DERG was/is organized by Finn Tarp, who holds the Danida Chair in development. It is/was financed by the Danida Research Fund; Hansen and Tarp are frequent Danida consultants; and the model was popularized by a special grant to Tarp and Hjertholm (2000).

9. The term *institutional home* is used here in the sense that the models are proposed by research financed by a specific institution, which also contributed to making the model known. Also, in both cases the model is broadly consistent with the thinking and policies of the institution. It has not, however, been officially adopted as the basis for the policy of the institution.

It reduces the standard ODA data to their gift equivalent. The two data sets are highly correlated, and we find that ODA data normally gives more significant results.[10]

Figure 2. Interpreting the data to get positive aid effectiveness

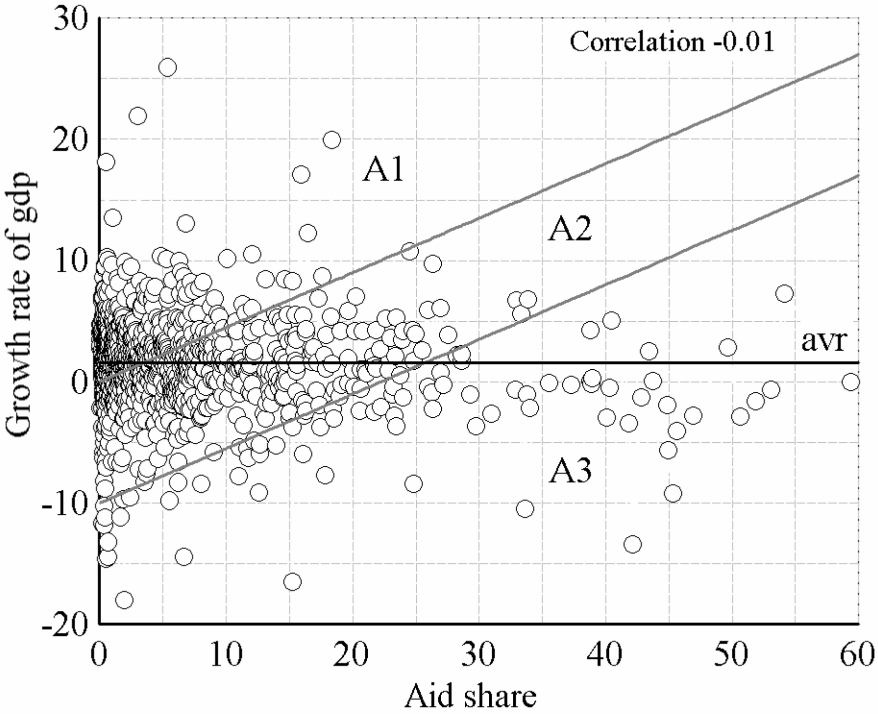


Figure 2 shows the distribution of the raw aid growth data for 156 LDCs (based on Paldam 2005). The data are averaged to 4 years and cover 1008 observations (10 observations are outside the frame).The graph is divided in three areas by the 2 gray lines: A1, A2 and A3. To find positive aid effectiveness, the conditional studies have to give areas A1 and A3 a separate explanation. Good Policy generates growth, so the Good Policy Model explains most of A1 by a separate term, while aid squared explains most of A3 by a separate term. In both cases one may hope to get a "nice" coefficient to aid.

2.3 The Good Policy Model: Condition is the Good Policy Index

In both models g_{it} is the real growth rate, and h_{it} is the aid share. The *j* controls are, $\boldsymbol{x}_{jit}$, and u_{it} are the residuals. Greek letters are the coefficients estimated. The two indices are *i* for countries and t for time.

10. The EDA data cover fewer years and countries than the ODA set, and therefore the competition between the new models came to take place on about 1/3 of the available data. See Doucouliagos and Paldam (2006b) on the relative effect of the two data sets.

Conditional Aid Effectiveness

The Good Policy Model has 2 equations. (1) gives the Good Policy Index, z_{it}, as a weighted sum of the budget balance, B_{it}, the inflation rate, p_{it}, and trade openness, T_{it}, while (2) is the aid effectiveness relation.

(1) $z_{it} = \lambda_0 + \lambda_1 B_{it} + \lambda_2 p_{it} + \lambda_3 T_{it}$ Good Policy Index[11]

(2) $g_{it} = \alpha + \mu h_{it} + \delta z_{it} + \omega z_{it} h_{it} + \boldsymbol{\gamma}_{jit} \mathbf{x}'_{jit} + u_{it}$ Aid Effectiveness Relation

In the original findings (Burnside and Dollar 2000), μ is insignificant, while both δ and ω are positive and significant.[12]

The coefficients for the Good Policy Index, z, is scaled (estimated), so it is fairly symmetrical around zero for the countries considered, and z is outcome oriented, so it is not surprising that δ becomes significant and positive. What is non-trivial is that the interacted variable $z_{it} h_{it}$ produces a significantly positive coefficient, ω. It means that aid to a country that pursues good policies increases growth, which is already high due to the good policies. Aid to a country with bad policies decreases growth, which is already low due to the bad policies. Obviously, the policy implications are that aid should be given to countries pursuing good policies only. It is cruel to give aid to countries pursuing bad policies.

Thus, in the Good Policy Model, the crucial coefficient is ω to the interactive term. However, also, μ, the coefficient to aid matters, as it shifts the relation up or down: If μ is large and positive, aid may be preferable to no aid, even in countries with bad policies, and reversely, if μ is negative, no aid may be preferable even in countries with good policies. Aid effectiveness thus depends upon both coefficients μ and ω.

One of the trust building features in the presentation of the model is that Burnside and Dollar (from the start in 1996) did not hide that the model recommended that aid should be redistributed away from the countries pursuing bad policies, where it harms, to countries pursuing good policies where it helps. They even calculate the welfare gain to the world from such redistribution. This clearly involves a break from the poverty orientation towards an efficiency orientation of aid, which is contrary to much of the rhetoric of aid. In World Bank (1998) the argument is presented in a more diplomatic way.

11. The original estimates for the coefficients of (1) are: $\lambda_0 = 1.28$, $\lambda_1 = 6.85$, $\lambda_2 = -1.20$ and $\lambda_3 = 2.16$.

12. Some studies estimate more general specifications that include squared aid, and aid policy and squared aid policy interactions, but the essence of the model is captured by equation 2.

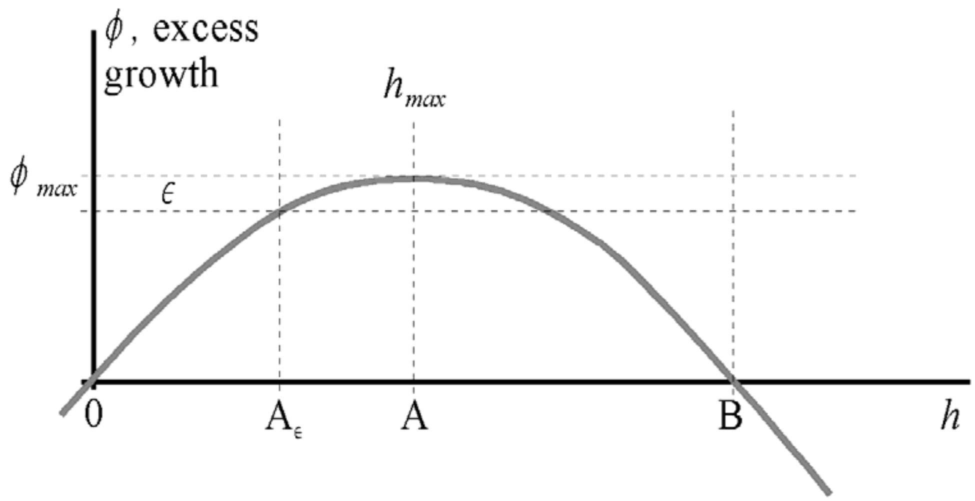
2.4 The Medicine Model: Condition is aid itself

This model needs one equation only as it uses aid itself as the condition:

(3) $$g_{it} = \alpha + \mu h_{it} - \omega h_{it}^2 + \boldsymbol{\gamma}_{jit} \mathbf{x}'_{jit} + u_{it} = \alpha + \phi(h_{it}) + \boldsymbol{\gamma}_{jit} \mathbf{x}'_{jit} + u_{it}$$

The proponents of the model find that μ, $\omega > 0$. The size and robustness of both μ and ω, are important for this model. The quadratic curve $\phi(h) = \mu h - \omega h^2 = h(\mu - \omega h)$ shows the excess growth due to aid. The ϕ-curve is zero at $h = 0$ and μ/ω. And $\phi' = \mu - 2\omega h$ and $\phi'' = -2\omega$. The maximum $(h_{\max}, \phi_{\max}) = (\mu/(2\omega), \mu^2/(4\omega))$. The ϕ-curve is drawn on Figure 3.

Figure 3. Excess growth, ϕ, due to aid in the Medicine Model



We can study the welfare properties of the model, by considering aid as a game between a donor agency, D, and a recipient country, R. D makes an offer of aid, which R accepts or rejects. Assume the ϕ-curve to be proportional to the welfare gain of R. The two standard alternative assumptions about D are:

(1) D is an ideal bureaucracy, which maximizes world welfare. There are some little costs, ε, in the donor country as well, and thus it offers A_ε. Around A the ϕ-curve is flat, so A_ε is at a bit lower than A. If D only considers Rs welfare, it offers A. In both cases R accepts. If aid is constrained, welfare is maximized, when ϕ_i' is the same for all Rs. As $\phi' = \mu - 2\omega h$, the condition is that all h_i are the same. Aid should be redistributed to make all aid shares the same. This is different from the advice from the Good Policy Model.

(2) D is a Niskanen-type bureaucracy. It wants to maximize the production/budget, i.e. *h*. Consider here intervals for *h*: (i) From *h* = 0 to A, the welfare of both D and R rises. (ii) From *h* = A to B, Ds welfare increases, while Rs welfare decreases compared to the optimum, but R still has a welfare gain, and thus accepts. (iii) From *h* > B, R looses and rejects aid. Thus B is D's the optimal point. Here D has captured all of Rs potential welfare gain.

Thus, when we consider the Medicine Model it is crucial where the A and B are located. Hansen and Tarp (2000) and Jensen and Paldam (2006), find that the best estimate of B is somewhere between 20% and 30% for *h.* If the EDA data are considered it is closer to 20% and for the ODA data closer to 30%. This means that A is between 10 and 15%.

Table 1 Average aid shares to 45 African countries 1990/94, 1995/99, 2000/04

Aid share From	Aid share To	Number of observations		% of cases Now	% of cases Proposal
0%	10%	52	Below A	51%	27%
10%	15%	31	↑↓		
15%	20%	15	Between A and B	30%	24%
20%	30%	17	↑↓		
30%		17	Above B	19%	49%
		132		100%	100%

Note: Aid shares are ODA/GDP in %. Countries between the Sahara and South Africa.

If we confront these values of A and B with the picture on Figure 2 it is clear that the great majority of aid shares are well below A. Also, they are very different. If the model is true, a large potential gain can be made by making them equal. Further, Table 1 gives the level of aid shares to Africa, 1990-2004. The observations are aggregated to 3 averages of 5 years each. Only three observations are missing. It appears that approximately half the aid shares are below A and half are above. There is even 19% over B. The column "proposal" shows the effect of doubling aid to Africa, as proposed by Bono, Jeffrey Sachs, and other idealists. It appears that aid has already increased considerably in 2005. According to the Medicine Model this would be harmful for Africa.

Thus both models have strong policy implications if they are true, so it is no wonder that they have been carefully analyzed in a total of 31 studies.

3. Are the substantial terms of the models established?

In order to make sense of the many conflicting results of the AEL, we use the tools of the rapidly emerging field of meta-analysis, see Appendix 1. As mentioned in the introduction we look at two questions: (Q1) Do the estimates of the effects *converge* to something we can term truth? (Q2) Can we identify the main *innovations* causing/preventing the convergence?

Regarding (Q1) we want to know if an empirical literature has established certain facts. In our case we want to know what the AEL says about the size of the two key coefficients, μ and ω, of the two main models. (Q2) We want to know if we can explain the observed variation in the estimates by methodological differences of the studies. The present section considers (Q1), while section 4 turns to (Q2).

The term *methodological differences* is broadly defined to mean: (m1) Differences in *models*, notably in control variables included. (m2) Differences in the *technique*s used for the estimation – notably if the relation is controlled for endogeneity. (m3) *Path dependency* occurs as the two models are proposed and defended by research groups that find significantly different results on the same data. We want to explain why. Finally we consider two variables for the *quality of the research*. (m4) The size of the *data samples* on which models are estimated – we consider both the number of years and countries covered.[13] (m5) The publication outlet for the paper is the assessment by the market of the quality of the research.

3.1 The data of the meta-analysis and the methods of analysis

The 31 studies listed in Appendix 2 provide the data of the meta-analysis. From the studies we derive a data set for each of the two models, which are the estimates of the two substantial coefficients μ and ω:[14]

The ***best-set*** of the 31 regressions chosen by the author(s) of the papers. In some cases it is unclear what the authors' prefer – in these cases we had to assess; but then the candidates for the best-set are normally close to each other.

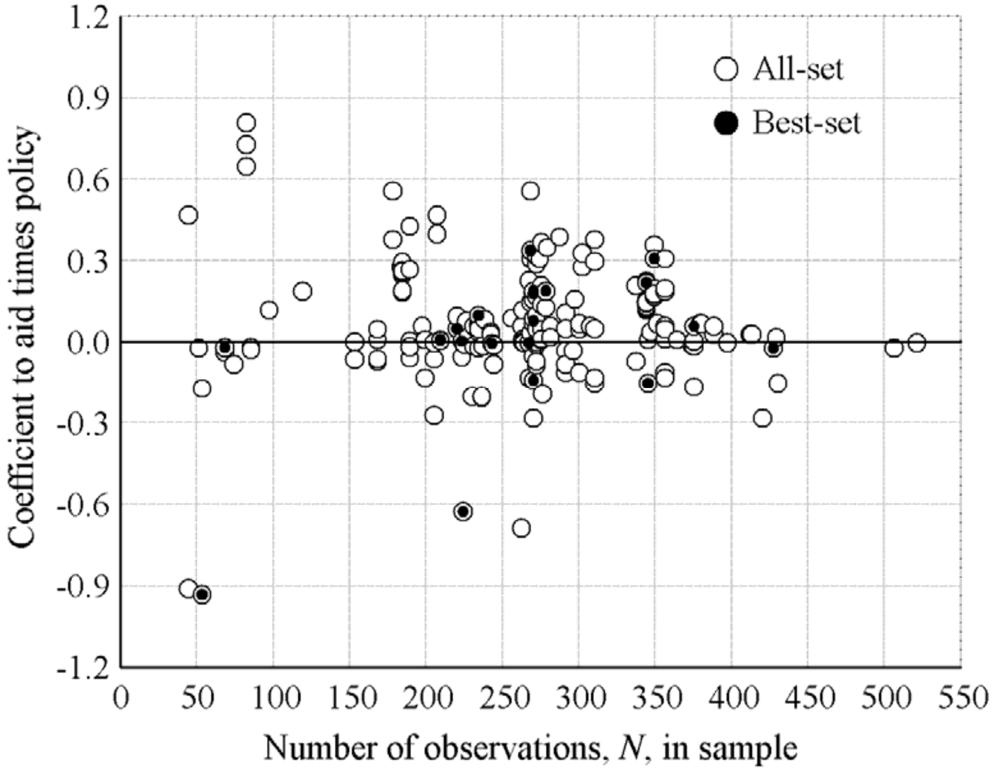
The ***all-set*** of all 232 regressions reported for the Good Policy Model and 123 for the Medicine Model. This increases the data available for tests, but it gives some interdependence between data points.

To each of these estimates we attach a vector of variables (as a check list) that characterize the methods by which the result is reached in the 5 ways (m1) to (m5).

13. Aid started in the early 1960s. It has accumulated with 100-150 observations annually since the mid 1970s.
14. We use all estimates reported by researchers, including their attempts to replicate the findings of others.

To get a 'feel' for the data, consider figure 4, which is a funnel plot of the 232 aid policy interaction estimates. All 23 studies have economic growth (as a percentage) as the dependent variable, and likewise all use a similar measure of aid (as a percent of GDP). Hence, the estimates are directly comparable across studies.

Figure 4. Funnel plot of aid and good policy interactive coefficients, ω



Note: One extreme point at (98, -3.68) from Svensson (1999) is outside the frame of the figure.

The funnel plot shows the association, if any, between the estimates of ω and its accuracy (proxied here by sample size). The coefficient should converge toward the true result as N increases, with smaller studies showing greater variation, so the point scatter should look funnel-like, but it does not.[15] There is a clear cluster of coefficients around the zero mark – especially for high Ns – suggesting that the aid policy variable has a coefficient close to zero, as is shown in tables 2 and 3. There are, however, many positive as well as some negative coefficients. To see if a policy effectiveness result has been established, we need to address two problems.

15. Using only the 23 best-set observations also does not produce a funnel shaped plot. See Doucouliagos and Paldam (2006b) for better examples of funnel shaped distributions.

First, should all or only some estimates be included? We explore the all-set (which includes results relating to robustness checks) as well as the best-set. Second, should all studies be treated equally? In the present section we use the sample size to assign weights to studies.[16] As this literature is relatively young, many of the papers are still working papers. We control for publication status of papers in section 4.

3.2 Have the coefficients μ and ω of the Good Policy Model been established?

The Good Policy studies have a best-set of 23 observations, and the all-set has 232 observations.[17] The key coefficient in the Good Policy model is ω on the aid-policy term and μ on aid. Table 2 reports the distribution of the Good Policy Model results for both coefficients. It is telling that of the 23 studies only (8/23 =) 35% found a positive and statistically significant aid-policy interaction.

Table 3 reports basic tests. Appendix 1 explains these tests. The Meta-Significance Test (MST) is a test for the existence of a genuine effect between two variables, using all the available empirical evidence.[18] A genuine effect will reveal itself through a positive and statistically significant association between the absolute value of t-statistics on the aid-policy interactions (the dependent variable) and the associated degrees of freedom, *df*, (the key explanatory variable in the MST). See Card and Kreuger (1995) and Stanley (2001, 2005) for details. If aid-policy interactions exist, then larger studies should have larger t-statistics.

The MST results for both Good Policy model datasets (top part of Table 3) show that while the coefficient on $\ln df_i$ is positive it is not statistically significant, indicating that the aid-policy interaction effect is not meta-significant (all p-values > 0.10). That is, taking all the available information there is *no* evidence of a genuine effect between aid-policy interactions and economic growth.

The Funnel Asymmetry Test (FAT) is a test for publication bias in a given literature. Publication bias is present if smaller studies (with larger standard errors) report larger aid-policy coefficients (Egger *et al*. 1997, Stanley 2005). For the all-set the p-value < 0.10, there is clear evidence of publication bias[19] – smaller studies report larger coefficients and hence

16. A larger sample should on average give more accurate estimates and should, hence, be assigned a larger weight (see Hunter and Schmidt 2004).
17. Some studies report only a single estimate, e.g. Lensink and White (2001) and Svensson (1999, but the average study reports 7 regressions.
18. The MST and FAT results for the all-set were derived using the bootstrap with 1000 replications. Stata 9.1 was used for all the meta-regression analysis.
19. This finding is not unique. Except for a couple of investigations, the majority of studies have detected publication bias in empirical economics research (see, for example, Card and Krueger 1995, Ashenfelter *et al*. 1999, Görg and Strobl 2001, Roberts and Stanley 2005 and Monkerjee 2006).

larger t-statistics. The sign on the constant in the FAT test indicates the direction of publication bias. Table 3 shows that the bias is in favor of positive aid-policy interaction terms. FAT also offers a second test for the existence of a genuine empirical effect. This would be revealed through a statistically significant coefficient on 1/SE. Aid-policy interactions fail this test as well.[20] For comparison purposes, we test also the 149 estimates of the impact of policy on growth reported by this literature. These are reported in the lower panel of Table 3. The MST here indicates that there is a genuine empirical effect on policy on growth, although it is of weak significance.[21]

Table 2. Sign counts of the effects of aid, μ, and aid times policy, ω, in the Good Policy and Medicine Models

Good Policy Model	All-set, N = 232				Best-set, N = 23			
	Positive		Negative		Positive		Negative	
	Signif.	Not	Not	Signif.	Signif.	Not	Not	Signif.
Aid, μ	25	56	60	7	8	6	9	0
Aid · policy, ω	90	75	55	12	8	5	8	2
Medicine Model	**All-set, N = 123**				**Best-set, N = 16**			
	Positive		Negative		Positive		Negative	
	Signif.	Not	Not	Signif.	Signif.	Not	Not	Signif.
Aid, μ	61	26	11	2	12	2	2	0
Aid squared, ω	0	7	27	89	0	0	4	12

Signif. is statistically significant at the 10% level, while *Not* is insignificant.

Of particular interest to us is whether there is any detectable difference between the results reported by researchers affiliated with the aid industry and those employed by universities. While it may be controversial, we expect *a priori* that the former group is more likely to either (a) find that foreign aid has a positive effect on economic growth, or (b) find that the impact of foreign aid on economic growth is dependent on policy. We have no prior expectations with respect to the direction of findings by researchers from universities who are not engaged with the aid industry.

20. We estimated also MST and FAT models with the addition of various controls for study characteristics (see Table 5 for the list of these). These results confirm the results from Table 3, and are available from the authors. Doucouliagos and Paldam (2006b) apply MST and FAT to the direct estimates of aid on growth (but not the conditionality estimates) and conclude that aid has no effect on growth.
21. However, this is not supported by the coefficient on 1/SE in the FAT. We are considering only those policy-growth effects reported by the aid conditionality literature. There are surely many more estimates of the impact of policy on growth, the meta-analysis of which is beyond the scope of the current essay.

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

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

Variable	(1) MST All-set	(2) MST Best-set	(3) FAT All-set	(4) FAT Best-set
		Good Policy Model		
Constant	-0.01 (-0.02)	-0.55 (-0.33)	**0.88** (6.85)	0.57 (1.09)
ln(*df*)	0.01 (0.08)	0.07 (0.22)	-	-
1/SE	-	-	0.01 (0.95)	-0.01 (-0.61)
Adjusted R^2	0.00	0.00	0.02	0.00
N	232	23	228	22
Average Y	+0.05	-0.18	+1.01	+0.60
		Medicine Model		
Constant	0.03 (0.03)	**2.55** (2.47)	**-2.08** (-19.46)	**-2.16** (-5.93)
ln(*df*)	0.08 (0.44)	**-0.34** (-1.73)	-	-
1/SE	-	-	0.0001 (0.84)	0.0001 (0.53)
Adjusted R^2	0.00	0.07	0.00	0.01
N	123	16	120	16
Average Y	+0.49	+0.65	-2.00	-2.08
		Policy Effects		
Constant	0.01 (0.01)	1.06 (1.00)	**3.32** (19.59)	**3.11** (4.78)
ln(*df*)	**0.19** (1.73)	-0.04 (-0.18)	-	-
1/SE	-	-	-0.02 (-1.54)	-0.06 (-1.56)
Adjusted R^2	0.01	0.00	0.02	0.00
N	149	17	149	17
Average Y	+1.01	+0.87	+3.18	+1.07

Explanation: If aid interaction terms have an effect on growth, ln(*df*) in the MST should have a *positive* and statistically significant coefficient. This fails for both models. If the literature is free of publication bias, the constant in the FAT should not be statistically significant. It is significant in 3 cases. The 1/SE term is also a measure of the existence of a genuine empirical effect. It fails in all cases.

Notes: The ln(*df*) coefficient relates to formula (3A) and the constant to formula (5A) of Appendix 1. Bolded estimates are statistically significant, at least at the 10% level. t-statistics in brackets, using robust standard errors. The bootstrap was used for the All-Set. Average Y reports the average value of the dependent variable (natural logarithm of t-statistic for MST and t-statistic for FAT). Some observations are lost due to missing data in some cases.

The aim of meta-analysis is to explain apparent differences in study findings. Theoretical debates are characterized often by several groups who develop and propose competing models. This competition results in evidence presented either in favor or against the alternative models, giving rise to heterogeneity in results. In this context, institutional affiliation can be seen to be a proxy for the competing model specifications. An analysis of these factors does not in any way imply nor suggest that any of the published results are in anyway deliberately manipulated.

Table 4. Meta-analysis of the estimates of the Good Policy Model

(1) Group	(2) No.[a)]	(3) Coeff.	(4) Simple average	(5) Median	(6) Weighted average[b)]	(7) 95% interval [c)] Lower	Upper
All-set of 232 regressions							
All Studies	232	μ	-0.01	+0.01	-0.01	*-0.09*	*+0.07*
		ω	**+0.07**	**+0.07**	**+0.08**	*+0.05*	*+0.10*
Non-Aid Business	93	μ	-0.06	+0.01	-0.06	*-0.21*	*+0.08*
		ω	+0.03	+0.01	+0.04	*-0.01*	*+0.08*
Aid Business	139	μ	+0.04	+0.01	+0.03	*-0.05*	*+0.11*
		ω	**+0.09**	**+0.12**	**+0.09**	*+0.08*	*+0.15*
Danida Group	22	μ	**+0.31**	**+0.14**	**+0.32**	*+0.13*	*+0.57*
		ω	+0.05	+0.01	+0.06	*+0.02*	*+0.10*
World Bank Group	62	μ	**-0.11**	**-0.02**	**-0.09**	*-0.17*	*-0.02*
		ω	**+0.08**	**+0.15**	**+0.12**	*+0.05*	*+0.16*
Best-set of one regression per paper							
All Studies	23	μ	+0.13	+0.13	+0.10	*-0.15*	*+0.35*
		ω	0.00	+0.03	+0.04	*-0.05*	*+0.12*
Non-Aid Business	11	μ	+0.03	+0.09	+0.01	*-0.43*	*+0.26*
		ω	-0.13	-0.01	-0.06	*-0.25*	*+0.03*
Aid Business	12	μ	+0.23	+0.16	+0.18	*-0.13*	*+0.60*
		ω	**+0.12**	**+0.10**	**+0.13**	*+0.04*	*+0.21*
Danida Group	4	μ	**+0.82**	**+0.80**	**+0.85**	*+0.21*	*+1.43*
		ω	+0.03	+0.03	+0.03	*-0.002*	*+0.05*
World Bank Group	6	μ	**-0.25**	**-0.25**	**-0.25**	*-0.43*	*-0.13*
		ω	**+0.23**	**+0.22**	**+0.23**	*+0.16*	*+0.30*

Notes: μ is the coefficient on aid, and ω is the coefficient on aid times policy. The bolded estimates are significant. a) Number of estimates. b) Weighted with number of observations for each estimate. c) Confidence intervals around the weighted average are bootstrapped and are bias-corrected, see Appendix 1. Non-weighted intervals are similar.

Above we mentioned that the World Bank is the institutional home of the Good Policy Model, while Danida is the institutional home of the Medicine Model. Accordingly, we consider two groups of aid business studies. 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).[22] The Danida group papers are: Dalgaard and Hansen (2001); Dalgaard, Hansen and Tarp (2004); and Hansen and Tarp (2000; 2001). We expect that the World Bank group has a prior for a positive aid-policy term, while the Danida group has a prior against the aid-policy term, but for the aid squared term.

22. The World Bank is a large organization, but the group was concentrated in the same department, with P. Collier (director) and D. Dollar (division head). They are also two of the most prolific authors in the AEL. Note that Svensson (1999), though he (at that time) belonged to the World Bank group, rejected the aid-policy term, but we do code this as part of the World Bank group of studies.

Table 4 gives a basic meta-analysis of the two data sets for the two coefficients. The estimates have been grouped and a set of averages calculated, and the obvious differences tested for significance. Table 4 combines the results from different studies into an overall result, showing what the literature has established for the all-set and for the best-set. Results are given for the coefficient ω on the aid policy interaction term and μ on aid alone. Column 2 lists the number of estimates for each sub-group. Column 4 reports the simple unweighted average of the estimates, and column 5 reports the median values. The sample size weighted averages are reported in column 6. Column 7 reports the 95% confidence intervals constructed around these weighted averages (see Appendix 1 for construction details).

The results for all studies reported in table 4 are that the effect of aid-policy conditionality on economic growth, i.e. the weighted average value of the interaction effect, ω is +0.04 for the best-set. This shows that the interactive term is of little *economic* significance; also, it is not statistically significant (the 95% confidence interval includes zero and recall also the MST from Table 3). Hence, we conclude that when all available studies are considered, there is no evidence of a positive aid-policy interaction. The table also shows that the average coefficient μ on the aid variable in this family of the AEL is also not statistically significantly different from zero. Thus, this family of studies has found that aid has no impact on economic growth, much like in our studies of the other families of the AEL (see Doucouliagos and Paldam 2006a and 2006b). Table 3 compares the results of the aid business studies to the non-aid business studies. They deviate systematically in the direction predicted, and the deviation is often significant – especially as regards ω.

3.3 Is there an aid squared effect?

The Medicine Model has been analyzed in 16 papers, and a total of 123 regressions have been presented with an aid squared term. A main problem for the analysis of this data is that a large fraction of the regressions are in papers proposing the model, Hadjimichael *et al.* (1995), Hansen and Tarp (2000; 2001), and Lensink and White (2001). Thus, relative to the good policy model, there is little independent replication of the model.

The distribution of the results is reported in the lower half of table 2. Taking the best-set of results from all 16 studies, the weighted average aid squared coefficient, ω, is -0.07, and the associated partial correlation is -0.12.[23] Taking all 123 estimates from the 16 studies, the weighted average aid squared coefficient is -0.13, and the associated partial correlation is -

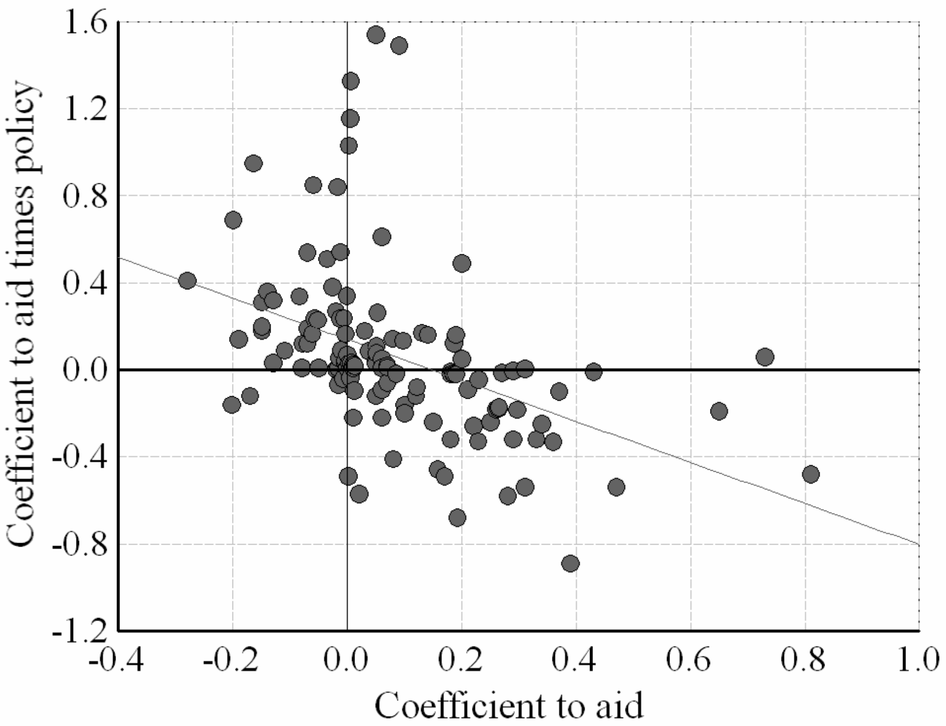
23. There is a high degree of skewness in the reported coefficients from this part of the literature. For example, while the unweighted average aid squared coefficient is -0.09, the median is -0.01.

0.12. Table 3 shows that the MST results for all the aid squared estimates had a coefficient on the natural logarithm of degrees of freedom of +0.08, and for the best-set it is -0.34, and neither is positive *and* statistically significantly different from zero. The lack of an aid*aid interaction is confirmed also by the 1/SE coefficient in the FAT test. Hence, there is no evidence of a genuine association between aid squared and economic growth. The constant in the FAT results indicates that the pattern of reporting of aid*aid interactions is biased in favour of negative coefficients. Hence, a firm conclusion on this aspect of conditionality cannot be made. Further research is clearly needed.

3.4 The debate between the two model groups

A comparison of the findings of the Danida and World Bank groups also confirms our expectations. The World Bank group reports a relatively large positive and statistically significant value for ω and a negative value for μ. In contrast, the Danida group studies find a near zero value for ω and a positive value for μ. Interestingly, the 95% confidence intervals for the best-set (reported in Table 4) do not overlap, indicating that the two groups of researchers produce quantitatively different results: Affiliation matters.

Figure 5. The relation between the coefficients to aid and aid times policy



One obvious feature of table 4 is the association between the magnitude of the coefficients μ and ω. Figure 5 is a simple scatter diagram of the coefficients on the aid variable and the coefficients on the aid-policy interactive terms, for the Danida and World Bank group of studies. It shows a clear negative association. Studies that report higher coefficients on the aid-policy interactive term tend to report lower coefficients on the aid term. As noted earlier, even if the aid policy interactive term is positive and statistically significant, it is still possible for aid to have a positive impact regardless of policy, if the coefficient on aid is large enough.

3.4 Institutional conditions: The future?

The 10 papers in the residual group are all relatively new proposals. They have not been independently replicated. They are promising, but because they are a small group they cannot be submitted to a formal meta-analysis.

Two papers condition for *democracy* (Svensson 1999; and Kosack 2002). Both suggest that aid works better in democracies, but are otherwise different, and the main thrust of Kosack's paper is to replace economic growth with growth of the human development index as the dependent variable, but he also reports results using growth, confirming that the two welfare measures are highly correlated.

Two studies by Chauvet and Guillaumont (2001; 2003) condition for various measures for *political instability* and *external vulnerability*, which attempt to catch institutional stability. This appears logical, as successful projects do need time for implementation, and hence some kind of institutional stability. Two related studies condition for *quality of institutions,* Collier and Dehn (2001) and Collier and Dollar (2002), but in different ways. Perhaps one may say that the quality and stability of institutions is the same factor for aid effectiveness. It is a main problem that we need simple and clear measures for this factor, but the proxies tried suggest that it is important for aid effectiveness.

One study conditions for *trade openness* (Teboul and Moustier 2001). The logic is here more indirect, and one gets the impression that trade openness is a proxy for a broader concept. The broadest such concept is perhaps *economic freedom* as is tried in Brumm (2003). Although it may not matter, it appears that the statistical methods used by Brumm are too far from the state of the art in the field.

Finally, two studies link aid with *GDP* (Bowen 1995 and Svensson 1999). They suggest that aid works better in more developed countries. The many studies that concentrate on a region or have regional dummies give evidence pointing in the same direction.

Table 5. Means and Standard Deviations of MRA variables. Good-Policy Model, All-set.

Variable	BD means binary dummy. It is 1 if condition fulfilled, otherwise 0	Mean	St dev
Dependent	BD if study reports significantly positive coefficients (ω)	0.40	0.49
DevJour	BD if published in development journal	0.14	0.35
AidBus	BD if author(s) employed/affiliated with aid agency	0.60	0.49
WorldBk	BD if paper from World Bank group	0.27	0.44
Danida	BD if paper from Danida group	0.09	0.29
NrCountries	Number of countries included in the sample	56	15
WorPap	BD if the research has yet to be published in journal	0.41	0.49
NrYears	Number of years covered in the analysis	25	3
Endo	BD if the aid was treated as an endogenous variable	0.16	0.37
EDA	BD if paper use EDA measure of assistance	0.56	0.50
Region	BD if paper controlled for regional effects	0.88	0.33
Subsample	BD if estimate relates to sub-sample of countries	0.18	0.38
Reproduce	BD if estimate is an attempt to replicate results	0.10	0.31
Fixedeffects	BD if Fixed Effects estimator used	0.07	0.25
Ethnic	BD if controlled for ethnic fractionalization	0.74	0.44
Polmeasure	BD if a Burnside-Dollar type measure of policy used	0.77	0.42
Finmarkets	BD if controlled for financial markets development	0.78	0.42
Institutions	BD if controlled for quality of institutions	0.91	0.29
Aidsqr	BD if included aid*aid term	0.26	0.44
*AidSqr*Policy*	BD if included aid*aid*policy term	0.10	0.30
Instability	BD if paper controlled for political instability	0.82	0.38

4. Accounting for differences in results for the two models

The previous analysis shows that priors influence results, but we also want to study the methodological differences producing the results. MRA (meta-regression analysis) can be used for that purpose (Stanley 2001 and 2005).

The dependent variable is a binary variable taking the value of 1 if the study reports a statistically significant positive aid-policy coefficient and otherwise 0.[24] Probit meta-regressions can be used to identify the determinants of conditional aid effectiveness. The number of observations is limited, so we only use the most important explanatory variables,[25] which are

24. Normally, the dependent variable in an MRA is regression coefficients, t-statistics or a partial correlations, estimated using a linear regression model (see, for example, Doucouliagos and Paldam 2006b). Our focus here, however, is on the factors that result in the reporting of a positive and statistically significant aid-policy interaction. Hence, we use a probit model.

25. We considered other aspects of studies, e.g. dummy variables for different journals, regions and period dummies. These variables were not statistically significant and can be ignored. Compared to the broader literature on the impact of aid on economic growth, the aid-policy conditionality literature uses very similar specifications. Hence, there are fewer candidates for specification variables. The data used is also broadly

defined in Table 5. The aim of our probit MRA is to identify the characteristics of studies that influence the reported results. We are still interested in exploring whether an author's association with the aid business results in qualitatively different results on the aid policy variable. We include the *Danida* and *WorldBk* dummies in order to test the findings of table 4 within a multivariate context. We add a control variable for working papers since working papers have not yet passed the referee process and, hence, may have a lower quality (i.e. *WorPap* is our binary measure of research quality).[26]

Four variables are included to capture the impact of data differences: *NrCountries*, *NrYears*, *EDA, Subsample and Polmeasure*. If the aid-policy conditionality is robust, we expect a positive association between the number of countries included in a study and the study results.[27] Similarly, we include the number of years of data from each study. *Subsample* controls for different sub-sets of data while *Polmeasure* controls for differences, if any, in the measure of policy. We include the *DevJour* variable to see if there are differences in results across the types of journals. *Reproduce* controls for estimates made by different authors with the sole aim of reproducing another researcher's results. Two variables relate to estimation techniques: *Endo* is included in order to see if accounting for the endogeneity of aid changes the results and *Fixedeffects* is the use of the Fixed Effects estimator makes a difference. Nine variables are included to capture the impact of specification differences: *Region*, *Instability, Ethnic*, *Finmarkets*, *Institutions*, *Aidsqr* and *Aidsqr*Policy.*

4.1 Results for the Good Policy Model

The MRA results reported in table 6 use the all-set of 232 estimates. The observations included in the all-set are not all statistically independent. Hence, we use the bootstrap to derive robust standard errors (see Doucouliagos 2005). We test the sensitivity of the results by running several regressions. The row labeled Wald (affilia.) presents the test results for Wald tests on the institutional affiliation variables (AidBus or WorldBk and Danida jointly). The marginal effects, reported in the last three rows, show the effect of affiliation on the probability that a study reports a positive *and* statistically significant aid-policy interaction.

similar. With the main difference been the number of countries used. See Doucouliagos and Paldam (2006b) for the impact of specification on the broader aid-growth literature.

26. Working papers may also use newer techniques and more recent data and, hence, may report qualitatively different results.

27. Several studies show that the choice of countries can influence study results, see e.g. Burnside and Dollar (2000) and Jensen and Paldam (2004).

Table 6. Meta-regression analysis, Good Policy Model, All-Set

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	**-0.58**	**-0.57**	**5.51**	**-0.38**	**-0.94**	**8.63**	**7.26**
	(-4.2, 4.2)	(-4.9, -5.2)	(3.1, 3.6)	(-1.7, 1.8)	(-1.7, -2.3)	(2.2, 3.7)	(3.1, 4.4)
AidBus	**0.52**	-	-	-	-	-	-
	(3.0, 3.0)						
WorldBk	-	**1.12**	**1.10**	**0.97**	**0.68**	0.67	**0.74**
		(5.6, 5.6)	(4.4, 4.8)	(3.7, 3.8)	(2.6, 2.7)	(1.4, 2.0)	(2.3, 2.5)
Danida	-	-0.03	-0.16	-0.33	-0.23	-1.32	**-1.16**
		(-0.1, -0.1)	(-0.4, -0.5)	(-0.7, -0.9)	(-0.5, -0.6)	(-1.3, -2.2)	(-2.6, -3.0)
NrCountries	-	-	**-0.08**	-	-	**-0.12**	**-0.12**
			(-3.0, -3.6)			(-2.2, -4.2)	(-3.1, -4.4)
NrYears	-	-	-0.03	-	-	-0.05	-
			(-0.8, -0.8)			(-0.8, -1.1)	
EDA	-	-	-0.30	-	-	-0.10	-
			(-1.3, -1.4)			(-0.2, -0.3)	
Subsample	-	-	**-1.34**	-	-	**-1.84**	**-1.95**
			(-2.6, -3.1)			(-2.1, -3.7)	(-2.8, -3.8)
Polmeasure	-	-	**-0.53**	-	-	0.12	-
			(-1.9, -2.1)			(0.2, 0.3)	
WorPap	-	-	-	-0.27	-	-0.64	-0.49
				(-1.1, -1.1)		(-1.5, -2.0)	(-1.5, -1.7)
DevJour	-	-	-	-0.15	-	-0.07	-
				(-0.4, -0.4)		(-0.1, -0.1)	
Reproduce	-	-	-	**0.97**	-	**1.21**	**1.26**
				(2.7, 3.0)		(2.2, 3.3)	(3.2, 3.8)
Fixedeffects	-	-	-	**-0.82**	-	**-1.98**	**-1.75**
				(-2.6, -1.5)		(-2.1, -2.5)	(-2.4, -2.6)
Endo	-	-	-	-0.48	-	-0.47	-
				(-1.5, -1.6)		(-1.0, -1.4)	
Instability	-	-	-	-	0.30	-0.09	-
					(0.8, 0.9)	(-0.1, -0.2)	
Ethnic	-	-	-	-	-0.65	-1.08	**-0.97**
					(-0.5, -1.1)	(-0.7, -1.5)	(-3.1, -3.5)
Finmarkets	-	-	-	-	-0.12	0.07	-
					(-0.1, -0.2)	(0.0, 0.1)	
Institutions	-	-	-	-	1.10	0.55	-
					(1.6, 2.2)	(0.5, 0.7)	
Region dummy	-	-	-	-	-0.24	-0.44	-
					(-0.6, -0.7)	(-0.5, -0.8)	
Aidsqr	-	-	-	-	0.33	**0.80**	**0.82**
					(1.3, 1.4)	(1.7, 2.4)	(2.8, 3.2)
*Aidsqr*Policy*	-	-	-	-	0.12	-0.26	-
					(0.3, 0.4)	(-0.5, -0.7)	
N	232	232	232	232	232	232	232
Pseudo R^2	0.03	0.11	0.24	0.19	0.18	0.40	0.38
Wald (affilia.)	**8.91**	**31.49**	**21.56**	**14.96**	**7.35**	3.53	**16.96**
Wald (controls)	-	-	-	-	-	2.59	
Aid business ME	**0.20** (2.99)	-	-	-	-	-	-
World Bank ME	-	**0.42** (6.09)	**0.41** (4.71)	**0.37** (3.93)	**0.27** (2.66)	0.24 (1.40)	**0.27** (2.27)
Danida ME	-	-0.01 (-0.10)	-0.06 (-0.50)	-0.12 (-0.76)	-0.08 (-0.51)	**-0.29** (-2.46)	**-0.26** (-4.14)
NrCountries ME	-	-	**-0.08** (-2.78)	-	-	**-0.04** (-2.81)	**-0.04** (-3.90)

Notes: The dependent variable is a binary variable reflecting whether the aid-policy interaction term of the study has a positive and statistically significant impact on economic growth. Figures in brackets are z-scores using bootstrap robust standard errors and z-scores without using the bootstrap, respectively. *ME* is marginal effect. Bolded estimates are statistically significant, at least at the 10% level.

Table 6, column 1 presents the regression with only the aggregate AidBus variable. It is significant. This is separated into WorldBk and Danida in column 2.[28] Data differences are introduced in column 3. Estimation and publication outlet differences are added in column 4. Column 5 uses the specification dummies. The general model (with all variables included) is presented in column 6. Column 7 presents the results of the specific model, after several statistically insignificant variables were eliminated (we sequentially removed any variable that had a t-statistic of less than one).

The Wald test for the joint insignificance of the eliminated variables is reported in the column labeled Wald (control). Column 7 is our preferred set of results from Table 6. The Wald (affilia.) tests for columns 2 to 7, show that WorldBk and Danida are jointly strongly statistically significant determinants (except in column 6).

The *WorldBk* dummy variable is positive and statistically significant. Studies by authors associated with the World Bank are more likely to report positive aid-policy effectiveness results. The *Danida* variable has the expected negative sign and the associated marginal effects are statistically significant in the results presented in columns 6 and 7. It is very important to note that these effects remain even after controlling for obvious specification and modeling differences.

NrCountries are statistically significant: the more countries included in a study, the *less* likely is a positive and statistically significant aid-policy interaction. Endogeneity is not important in determining a positive aid-policy conditionality result, however relative to OLS, the use of fixed effects results in a lower probability of reporting favorable aid*policy effects. The coefficient on *Reproduce* confirms that researchers are able, in general, to replicate prior findings – dependent replication is not a problem in the AEL. Interestingly, the inclusion of an aid squared increases, on average, the significance of the aid-policy term while ethnic fractionalization decreases it.

The probit meta-regression analysis for the best-set is problematic, as there are only 23 observations and the variables *Danida* and *WorldBk* are perfect predictor variables: as expected *Danida* has a strong negative and *WorldBk* a strong positive effect on the probability of reporting positive aid-policy interactive terms.[29]

The meta-regression analysis was repeated for the data associated with the Medicine Model (these results are available from the authors). The MRA here is not as conclusive as it

---

28. WorldBk and Danida do not exhaust the Aid Business category. If a third variable – other Aid Business – is added to any of the regressions, it is never statistically significant.

29. These estimates are available from the authors. The marginal effects show that the probability that a positive and statistically significant effect is reported is almost 1 when the researcher is associated with the World Bank.

was for the Good Policy Model. However, once again, the number of countries included in a growth regression appears to be robustly inversely related to the probability of finding a statistically significant conditionality result. The results show also that working papers are less likely to support the medicine model.[30] For the best set of only 16 observations, *Danida* is a perfect predictor variable, and it is the only variable with a significant marginal effect of +0.32 (z-statistic = +2.29).

5. Summary and conclusions

The aim of this paper was to explore one family of the AEL, aid effectiveness literature, by the tools of meta-analysis. We analyzed the conditional models, where the effectiveness of aid depends on a condition: If it is in one interval aid helps, and if it is in another aid harms.

The most researched condition is "Good Policy" (defined above). The number of studies is already large enough to permit clear conclusions on the two questions asked:

(Q1) Is the impact of aid on growth moderated by policy? The aggregate coefficient to the interaction between foreign aid and policy proves to be very close to zero. Good policies help increase growth, but they do not appear to influence the marginal effectiveness of aid. (Q2) Are the reported estimates systematically influenced by study characteristics? We established that the author's institutional affiliation does influence reported results, as do sample size, estimation technique and model specification.

The success of the Burnside and Dollar and World Bank reports was based on the evidence available at that time, but subsequent analysis has shown that their conclusions were premature. This proves Hunter and Schmidt's (2004, xxvii) statement that: "Scientists have known for centuries that a single study will not resolve a major issue. … Thus, the foundation of science is the cumulation of knowledge from the results of many studies".

The second most researched condition is aid itself, where aid works as medicine, which has an optimal dose. Here we asked the same questions. The number of independent studies is not large enough to reach a sharp conclusion, either way. It is a more robust connection, but it has proved unstable in independent replication, so more research is needed.

Finally, the literature contains a whole set of new conditional variables that have been proposed and tested once or twice. These variables are potentially quite promising, and suggest that the field is open to future research.

30. For the all-set, the key marginal effects are: for *NrCountries* -0.01 (z-statistic =-2.28); for *WorPap* -0.60 (z-statistic = -3.56); and for *WorldBk* -0.49 (z-statistic = -1.63).

References (note also Appendix 2 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, 1277-1283

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

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

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>

Collier, P., Dollar, D., 2004. Development effectiveness: What have we learnt? *Economic Journal* 114, 244

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, H., Paldam, M., 2006a. Aid effectiveness on accumulation. A meta study, *Kyklos*,

Doucouliagos, H., Paldam, M., 2006b. Aid effectiveness on growth. A meta study. Working paper, Aarhus Univ.

Doucouliagos, H., Paldam, M., 2006c. The aid effectiveness literature. The sad results of 40 years of research. Working paper, Aarhus Univ.

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

Egger, M., Smith, G.D., Scheider, M., Minder, C., 1997. Bias in Meta-Analysis Detected by a Simple, Graphical Test. *British Medical Journal* 316, 629-34

Gorg, H., Strobl, E., 2001. Multinational Companies and Productivity: a meta-analysis. *The Economic Journal* 111 (475), 723-39

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

Hedges, L. V., and 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 using the basic relations, Aarhus Univ. [26]

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

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

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

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

Paldam, M., 2004. A look at the raw data for aid and growth. Posted at <http://www.martin.paldam.dk>

Rajan, R.G., and A. Subramanian (2005). Aid and Growth: What Does the Cross-Country Evidence Really Show? IMF WP No. 05/127. June (see also WP No 05/126)

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

Rubin, B. D. 1990. A new perspective on meta-analysis. In *The Future of Meta-analysis*, edited by K. W. Wachter and M. L. Straf. New York, NY: Russell Sage.

Stanley, T.D., 2001. Wheat From Chaff: meta-analysis as quantitative literature review. *The Journal of Economic Perspectives* 15, 131-150

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, 947-973

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

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

Appendix 1: An 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.

A1. 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, common coefficients or partial correlation) from the i^{th} study and N is the sample size.

A2. Confidence intervals

Confidence intervals in meta-analysis can be calculated in several ways. Hunter and Schmidt (2004) derive the formula for the standard error in the mean correlation for a homogenous group of studies, as well as the standard error in the mean correlation for a heterogenous group of studies. Hedges and Olkin (1985) use a slightly different procedure. We prefer to follow Adams *et al.* (1997) and use resampling techniques to construct bootstrap confidence intervals. Bootstrap confidence intervals are more conservative. The 95% confidence intervals were constructed using the bootstrap, of 1000 iterations (with replacement) to generate the distribution of aid and aid-growth interaction effects (see Efron and Tibshirani, 1993). The lower and upper 2.5% of the values of the generated distribution are used to construct the 95% confidence intervals.

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.

A3 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.

A4 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 ε_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, β_1, in equation (5A) will be statistically significant. β_0 is the publication bias adjusted measure of a genuine empirical effect (Stanley 2005).

A5 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 ε_i is the standardised effect derived from the i^{th} study (in our study we use the partial correlation, r), α 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, γ and δ are the unknown regression coefficients, and u_i is the disturbance term, with usual Gaussian error properties. The regression coefficients in 6A quantify the impact of specification, data and methodological differences on reported study effects (ε_i).

Appendix 2: Studies used in the meta-analysis

Only papers in English available till 1/1 2005 are included. The whole EAL is listed in Doucouliagos and Paldam (2006a).

1 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

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

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

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

5 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

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

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

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

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

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

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

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

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

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

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

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

17 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

18 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

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

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

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

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

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

24 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

25 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

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

27 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

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

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

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

31 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