

The aid effectiveness literature

The sad result of 40 years of research

Hristos Doucouliagos
Department of Accounting,
Economics and Finance
Deakin University
Melbourne
Australia

and

Martin Paldam
Department of Economics
University of Aarhus
Aarhus
Denmark

Abstract: The AEL consists of empirical macro studies of the effects of development aid. At the end of 2004 it had reached 97 studies of three families, which we have summarized in one study each using meta-analysis. Studies of the effect on investments show that they rise by $1/4$ of the aid – the rest is crowded out by a fall in savings. Studies of the effect on growth show a small positive effect that is insignificant insignificant. Studies of the effect on growth, conditional on something else, have till now shown weak results. The Dutch Disease effect of aid has been ignored. The best aggregate estimate is that since its start in the early 1960s aid has increased the standard of living in the poor countries by 20% – this however is based on insignificant evidence.

Jel.: B2, E21, E22, F35, O35

Keywords: Aid effectiveness, meta study, accumulation, growth

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, Tryggvi Thor Herbertsson and Niels Haldrup.

The welfare loss due to mass poverty in the LDC world is one of the largest problems facing mankind. Many in the rich countries want to do something to reduce the problem, and all DCs give development aid. Mass poverty is now falling more than ever, but it is due to the fast growth of China and India, which receive little aid, while poverty is falling much slower – if at all – in the main aid recipient countries. This has caused many to doubt the effectiveness of aid, and in economics doubt is a main reason for the large empirical AEL, aid effectiveness literature, based on macro data.

The literature considers a model $g = g(h)$ between growth g and h the aid share. Effectiveness means that the coefficient $\mu = \partial g / \partial h > 0$. The model may be uncontrolled or controlled for country heterogeneity, which we term *absolute and conditional aid effectiveness* respectively. All results – also those in section 2 – reject absolute aid effectiveness.

A thorough search showed that the AEL as of 1/1-2005 consists of 97 studies. Below we divide the studies in three main families of models. Doucouliagos and Paldam (2005a, b and c) analyze each family using the tools of meta-analysis. Consequently, we now know precisely what the AEL says. The results vary remarkably, but the aggregate results are fairly sad as summarized in table 1. Even when the average effect of aid is positive, it is small and of dubious significance statistically. With the accumulation of more data, the results have grown gradually worse. The latest disappointment was the collapse of the once promising Good Policy Model, discussed in section 6.

Table 1. Main conclusions from our three meta studies

Type	Causal link	Conditional on	Conclusion	Significance	See	Our study
Family A	Aid → investment		App. 0.25 ^{a)}	Dubious	Section 4	(2005a)
	Aid → savings		App. -0.75 ^{a)}	Dubious		
Family B	Aid → growth		Positive, small	No	Section 5	(2005b)
Family C	Aid condition → growth	Good policy	Rejected	No	Section 6	(2005c)
		Aid itself (aid squared)	Positives small	Dubious		
		Eight more models	Not replicated			

Note: Aid either leads to increased investments or is crowded out through a fall in savings. Hence the investment effect minus the savings effect should add to 1. Significance means that the investment effect is larger than 0 and that the savings effect is larger than -1.

The present essay discusses the implications of our three studies. Section 1 looks at the raw data for aid and growth and argues that the impression from the data contrasts with standard

economic theory. Section 2 divides the AEL into three families of models. Section 3 briefly enumerates the purposes of the meta-analysis. Sections 4 to 6 present the results of the meta-analysis of the three families of models. Section 7 discusses an overlooked parallel to a literature that may explain the findings of the AEL. Section 8 shows what the average results imply, if insignificance is disregarded, and finally section 9 concludes.

1. Absolute aid effectiveness and the aid paradox

The most widely used comparative data for aid and growth are shown in figure 1. They are all the available data from the WDI, World Development Indicators. We have used the series for real economic growth per capita, g , and the share of development aid of GNI since 1960. Both series are then averaged to the 10 four-year periods: 1961 to 64, 1965 to 68 ... 1997 to 2000. WDI covers 156 LDCs, so we should get 1,560 observations. About 35% are missing, but this still leaves the 1,008 observations shown.

The basic regressions between the data are given in table 2. These “no frills” regressions are far from state of the art, but then there are the 1,025 regressions by 104 researchers of the 97 papers of the AEL. The regressions of table 2 show that the raw data for aid and growth have no connection. There is no *absolute* aid effectiveness. Even the significant leaded regressions (shaded in gray) that show the maximum size of the reverse causality bias have a negligible size. The basic regressions between these variables, controlling for level of development, fixed effects for countries and time periods, as well as for cyclical effects are all very similar, see Herbertsson and Paldam (2005). This is strange indeed.

Table 2. Simple regressions between aid and growth

Same data as figure 1		(1) Lag +1 growth before aid		(2) Lag 0 aid and growth		(3) Lag -1 aid before growth	
		Coeff.	p-value	Coeff.	p-value	Coeff.	p-value
All data	Constant	1.816	0.000	1.579	0.000	1.504	0.000
	Effect/slope	-0.039	0.023	-0.010	0.935	0.003	0.364
	N	895		1,008		876	
	R ²	0.006		0.000		0.000	
In box	Constant	1.843	0.000	1.676	0.000	1.578	0.000
	Effect/slope	-0.052	0.007	-0.022	0.207	-0.010	0.559
	N	841		945		839	
	R ²	0.009		0.002		0.000	

Note: Bolded estimates are significant at the 5% level.

Figure 1a. A scatter plot of the data for aid and growth

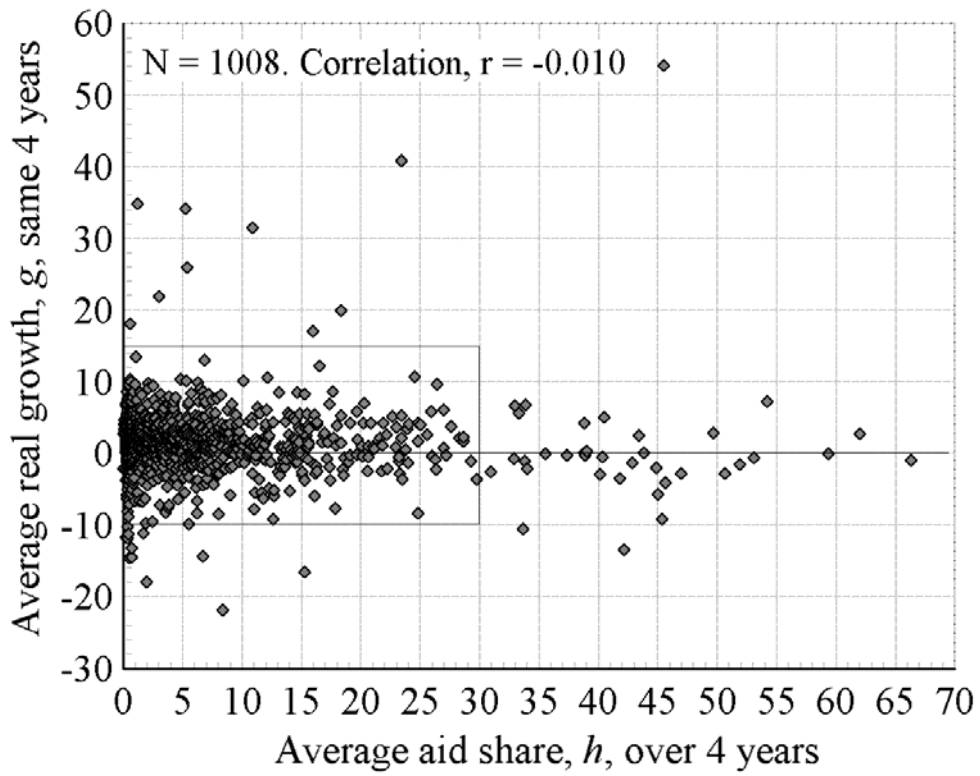
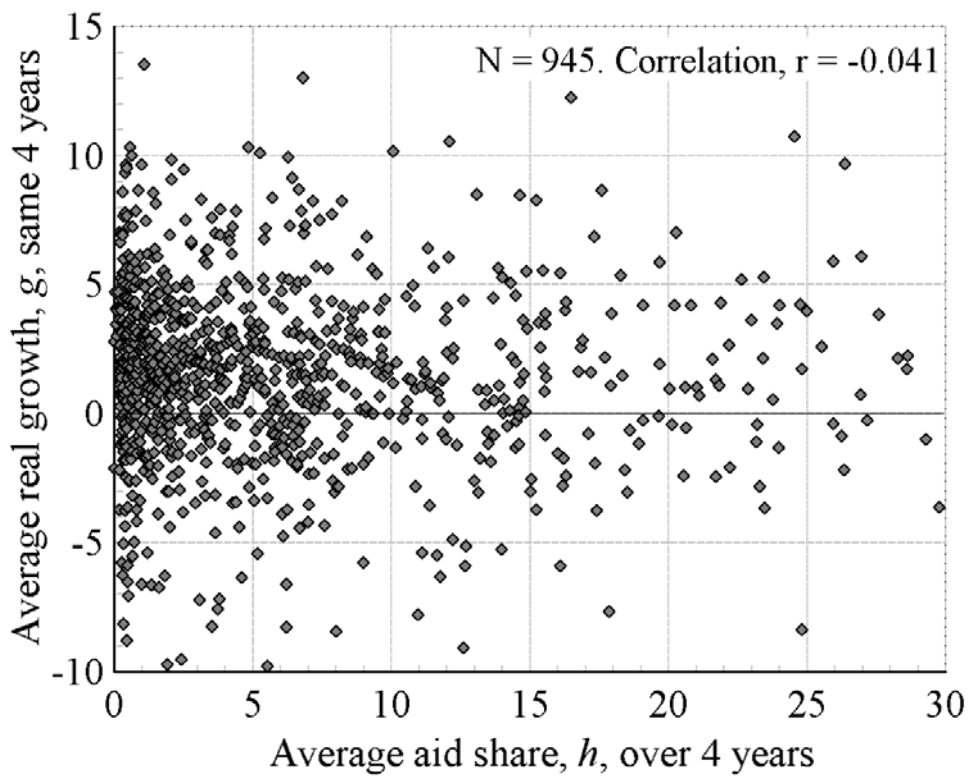


Figure 1b. The box on figure 1a enlarged



Note: An *Appendix* with similar graphs lagged to both sides is available, see Paldam (2005).

There are four good reasons, R1 to R4, why there should be a clear positive effect of aid, even if counter-arguments can be made to all four:

R1: *The micro evidence.* All aid programs have an evaluation process, and many studies have summarized the findings. Cassen (1986, 1994) is the classic survey, and the results are uncontroversial. About 50% of all development projects work, and very few of the remaining projects harm even if they fail. Simple aggregation thus predicts that aid works, even if it could work better. The contrast between the macro level results (such as the one of table 2) and the micro level findings is known as the micro-macro paradox since Mosley (1986).¹

Counter-R1: Aid is fungible, so the marginal effect of aid is somewhat different from what is financed by the aid. Donors have a good chance of selecting non-marginal projects, so the marginal effect of aid on growth is likely to be less than the effect of the projects financed.

R2: *Standard macro theory.* Aid leads to a balance-of-payments improvement and to public spending. Public spending has an activity effect, and that effect can be permitted to run in the economy due to the balance-of-payments improvement. **Counter-R2:** A whole set of arguments exists why some of the activity effect may not have its full size, but is *crowded out*. The fullest crowding out is the *Ricardian Equivalence* by which loans – that have to be paid back – will have no effect, but lead to increased savings. However, development aid has a gift element, which does not have to be repaid, and thus should have an effect.

R3: *Standard growth theory.* Both the theory of growth and growth empirics show that increased accumulation causes growth. We know that aid finances development projects, which are – in principle and often in practice – investments. **Counter-R3:** Accumulation is only one factor generating growth, and the marginal activity caused by aid is somewhat different from the activity financed. The link from aid to growth does not necessarily proceed via the accumulation effect generated.

R4: *The “why would they” argument.* Given standard rationality assumptions, an activity such as aid that has run for 40 years must do at least some of what it should. Why else would it continue? **Counter-R4:** The average aid share of the donors is actually quite small (about 0.3% of donor GDP) and has even decreased a little in the last decade due to *aid fatigue*, caused by precisely the dissatisfaction with the small effects of aid. Also, it is obvious that while aid optimism was high during the first 1-2 decades, a business was created that has the usual stakeholder interests in the continuation of the activity.

1. There are indeed several such micro-macro paradoxes. For example, there is the exchange rate disconnect puzzle where high volatility in real exchange rates is not related to fluctuations at the macroeconomic level, even though the business sector claims that there are real effects at the microeconomic level.

Seen together, R1 to R4 do suggest that aid should help – maybe not very much – but at least more than shown in figure 1 and table 2, which suggest a total lack of effect. This is the aid paradox that has driven research in the field.

A look at figure 1 allows us to draw two more conclusions: First, the data should be ideal for an analysis of aid effectiveness. They are plentiful and have great variance. The aid shares have an average of $7\frac{1}{2}\%$ of GDP. This is substantial relative to other quantities that are known to affect growth. Secondly, the fact that the raw data show nothing means that any significantly positive (or negative) effect found must be due to the imposition of structure on the data. That is, results are due to the “frills” of the analysis, as is further discussed below. It is clear that there is no absolute aid effectiveness, but the AEL is (still) looking for conditional aid effectiveness.

The analysis of aid effectiveness takes each country to provide equally good information for the analysis. The data points from India and Mauritius in figure 1 are thus of the same size. We also noted that the average aid share is $7\frac{1}{2}\%$ in the average country. However, the aid received by the average citizen in the LDC world is much smaller for two reasons: (1) A well-known fact about the cross-country distribution of aid is that its share falls with the size of the population. Most countries are small, but the bulk of the population lives in large ones. The giants, India and China, have aid shares well below $\frac{1}{4}\%$ and almost 40% of the LDC population. (2) About a third of the information for the calculation of the true average is missing. It is likely that the missing values are below average. If aid shares are weighted with population sizes of the countries, the average share falls to about one third, i.e. to $2\frac{1}{2}\%$ of GDP.² Thus the cumulated aid – over the 40 years – corresponds to one annual income (GDP per capita) of the average citizen of the LDC world.

Finally, it should be mentioned that the standard ODA measure of aid is defined as unilateral transfers with a gift element above a moderate threshold. However, Chang, Fernandez-Arias and Serven (1998) introduced the EDA measure of aid, where each grant is weighted by its gift element. EDA data are available for fewer countries and years than the ODA data, but the two data sets have a correlation of 0.83 when overlapping. Since the EDA data became available, some of the research has used EDA and some ODA data. We have dealt with that complication by converting coefficients into elasticities, which are invariant to the pure shift in scale, and then controlled for real differences in the two measures by including an EDA dummy in the meta-analyses. It becomes negative. We interpret this as

2. The missing observations for 30% of the countries cause some uncertainty, so the percentage is $2\frac{1}{2} \pm \frac{1}{2}$. For ease of calculations we have set it at $2\frac{1}{2}\%$.

evidence that policy makers are myopic. They consider the size of the ODA, and largely disregard repayments, which are likely to be the problem of later governments anyhow.

2. The three families of models and dimensions of the mining

The AEL has explored many models, but we can, as mentioned, divide them into three main families according to causal structure. This is done in figure 2, while table 3 gives the typical equations estimated for each family of models, and the variables used.

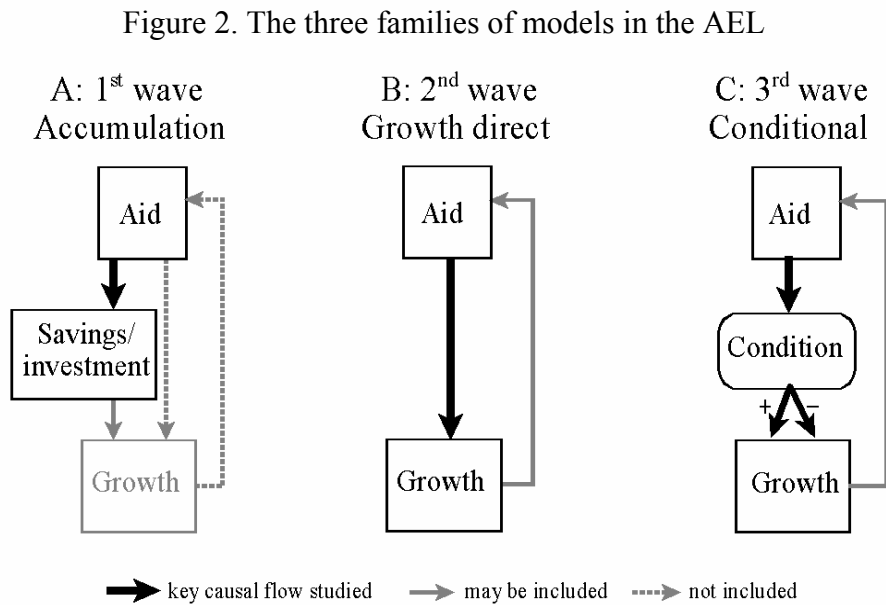


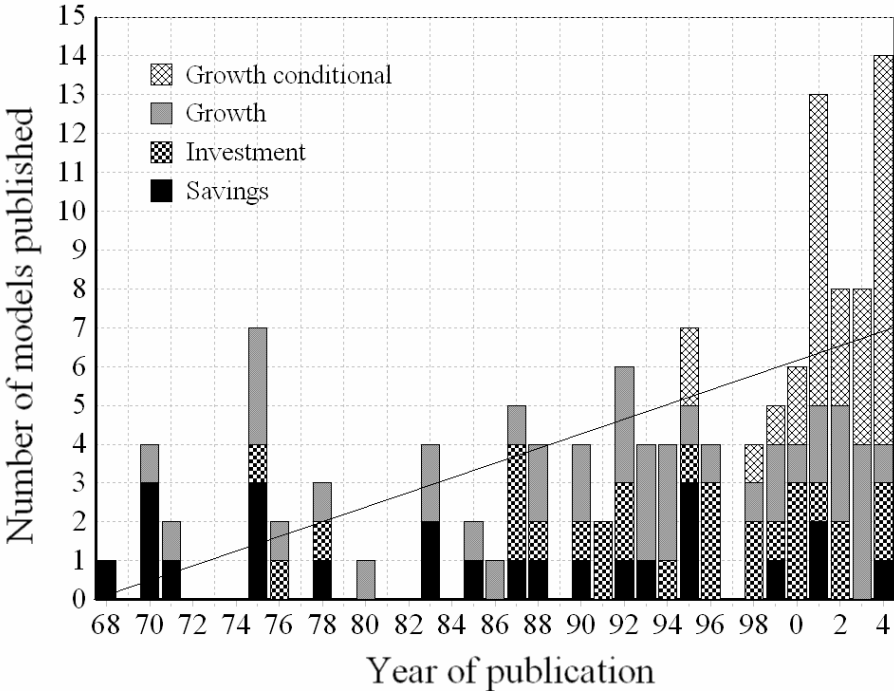
Table 3. The models and variables of the AEL

Family of models	Key model:		
A: Accumulation	$a_{it} = \alpha + \mu h_{it} + \gamma_j x_{jit} + u_{it}$		
B: Growth	$g_{it} = \alpha + \mu h_{it} + \gamma_j x_{jit} + u_{it}$		
C: Conditional growth	$g_{it} = \alpha + \mu h_{it} + \delta z_{it} + \omega h_{it} z_{it} + \gamma_j x_{jit} + u_{it}$		
Variable	Definition	Variable	Definition
i	index for countries	a_{it}	rate of savings/investments (of GNP/GNI)
t	index for time period (of 3-10 years)	g_{it}	real growth rate
j	index for control variables	h_{it}	aid share (of GNP/GNI)
α	constant, may be divided into	z_{it}	conditional variable
$\alpha = (\alpha_i, \alpha_t)$	fixed effects for countries and years	x_{jit}	vector of j control variables
$\mu, \delta, \omega, \gamma,$	coefficients to be estimated	u_{it}	residuals

Note: Most of the early models were pure cross-country models, i.e., had no time index, and some models are estimated for one country only, so they have no country index.

Figure 3 shows the development over time in the production of the models of the AEL. It has a significantly rising trend. It started with a wave of type A models – first savings models and then gradually investment models. Then came the larger wave of type B models, and finally, since 1995, C type models have emerged. Papers in the later waves often contain estimates of models of one or even two of the previous families. Type C papers always give results of type B as well, so the growth papers in figure 3 are only the pure growth papers. The present wave of papers is still on the upswing, so we are likely to see many more papers in the field.

Figure 3. Production over time of papers in the AEL



Note: The line included is a linear trend-line through the number of models published. It has a significant slope, but it exaggerates the slope, as the last 5 years include some working papers, which may or may not be published later, while no working papers are included the first 30 years.

Aid programs started during the 1960s, and aid statistics for that decade are scanty; but since 1970 aid data have rapidly accumulated, and now they grow by about 130 observations per year. The period from about 1970 to now has seen dramatic technological progress in the two relevant fields: The power and availability of computers have increased many thousand times, and econometric techniques have advanced greatly. In the early AEL papers so few data were available that only cross-country estimated were possible, but panel econometrics was not

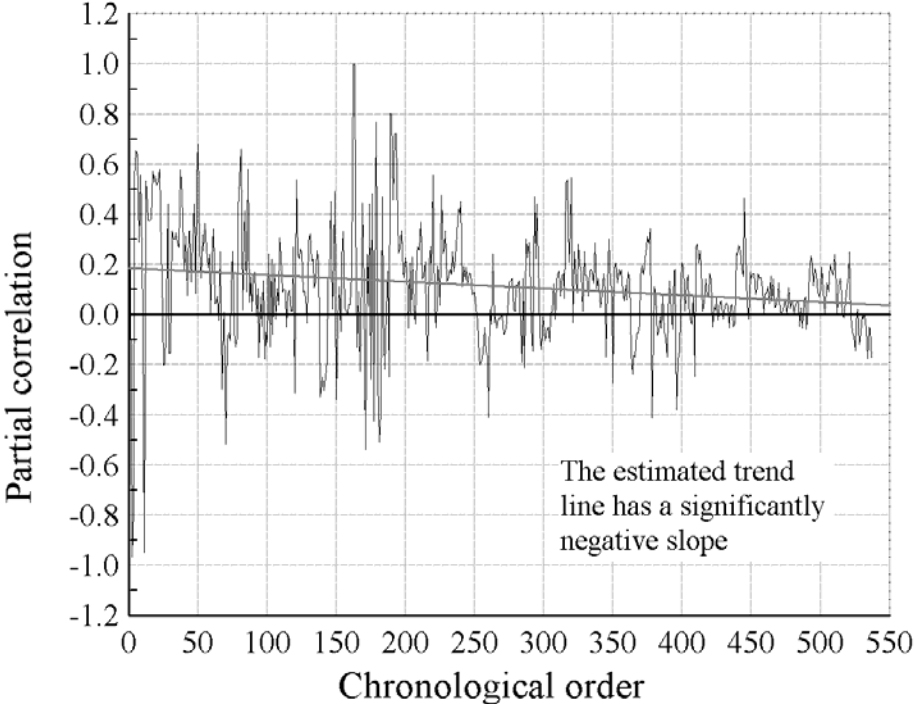
developed, anyhow. Now data have multiplied, and panel methods are routinely used, with TSIV or GMM estimators to take care of simultaneity, etc.

We studied whether the use of more advanced econometrics had any effect on results, by including binary dummies for techniques when we tried to explain the pattern in the results (see below). These dummies had no effect. It might be due to the low level of significance throughout, but also we found one instance only where an author noted that results changed when a more advanced method was used. It was a case where the introduction of an econometric refinement changed a coefficient from (just) insignificant to (just) significant.

What changed the results are not the advances in econometric techniques, but the increase of data available. This story appears to generalize.

The growing data sample has not improved the results. Figure 4 shows the development of the most numerous comparable statistic of aid effectiveness. Not all studies give all information, but we have managed to calculate 537 partial correlations of aid on growth.

Figure 4. The 537 partial correlations of aid on growth



The figure has two characteristics: (C1) The variation is falling over time as it should when the sample size grows. (C2) The average result is steadily decreasing, and is now +0.04 only. When regression number 685 is published (in 4-5 years), the trend will intersect with the zero

axis. This is puzzling: Over time, the technical skills of the personnel in the aid sector increase, and experience is steadily cumulating; consequently aid effectiveness should increase over time, so the trend should be upward. However, the sad fact is that the trend is downward.

In our studies we present a set of tests that strongly suggest two explanations for the downward trend: (e1) As small samples give a larger variation in results, it is easier to mine small samples than large ones. (e2) Most researchers want to find positive results. We hence conclude that the downward trend is an artifact – or rather that most of the positive average result in the early literature is problematic. The results now appear to be converging to the true association – which appears to be the zero line!

Table 4 gives the dimensions of the data set contained in the 97 papers. They hold 182 models of 7 types. Thus the average paper contains models from 1.88 families. The papers publish 1,025 regressions – it is 10 regressions per paper. If we assume that 25 estimates have been made for each published regression, the AEL is based on 25,000 regressions.

The models all try to explain subsets of the same observations. We can see the relation between the 26,422 data of the samples analyzed relative to the 1,008 data available as a crude measure of the data mining. Another way to see the same point is to compare the 1,008 data available with the number of 1,025 regressions published and the 25,000 that has probably been made. Clearly, the AEL has a data mining problem, as indeed does most empirical research in macroeconomics.

Data mining is a *common resource pool problem*. The individual researcher does not make a problematic dent in the degrees of freedom available by running, e.g., 250 regressions, on the 1,008 observations. However, the 104 authors of the AEL constitute a *mining collective*, who has mined the data thoroughly, and a dense net of cross-citations does exist in the field. Everybody has read some of the literature, and has thus joined the mining collective.

Table 4. Statistics of the AEL

Regressions	A: Accumulation		B: Growth	C: Conditional			Proxy	Sum
	Savings	Investments		Good Policy	Medicine	Others		
Best-set	21	37	68	23	15	10	8	182
All-set	61	122	543	162	85	23	29	1,025
Sample size	1,890	3,872	11,312	5,523	4,284	663	2,264	29,976

Note: *Proxy* studies are done using data – such as capital inflows – instead of aid, but nevertheless draw conclusions regarding aid. This was often done in the early papers where few aid data existed. *Best-set* is the regression preferred by author of paper, *all-set* are all regressions published.

Data mining is a process that eats degrees of freedom, but it is forbiddingly difficult to calculate the precise amount of mining done and the resulting loss of degrees of freedom. The first data published are mined by most of the 97 papers, while more recent data are mined by the most recent papers only. Nearly all researchers have chosen to disregard the problem.³ Even if we cannot calculate the true test limits, in the (very common) presence of data mining, we know that there is no problem as regards Type I errors (rejection of true model), but it increases the likelihood of Type II errors (acceptance of false model). That is, some models are reached by refining a random quirk in a certain data set.

We conclude: These data have been so thoroughly mined that it is *highly likely that some Type II errors have occurred in the papers published*. This is why independent replication of models is essential for their credibility. As more data accumulate, the literature should reveal whether any models survive. Also, the techniques of meta studies contain tests developed to be mining-proof, though it is, of course, difficult to design a fully mining-proof test.

3. Meta-analysis: Data and main purposes⁴

A meta-analysis should cover a literature in its totality, and it deals with two data sets:

The best-set: Each model provides one data point, the empirical result preferred by the author.

The all-set: Each empirical result – that is each regression – is taken as a data point.

For the AEL the best-set gives 182 data points, while the all-set provides 1,025 data points. They are analyzed with three purposes in mind:

Purpose 1. To sum up a literature. When an effect has appeared in a literature the meta-analysis asks: *Has the literature established that effect?* We thus ask: Have the 97 papers of the AEL established that aid works by increasing economic growth (or accumulation)? This can be done in several ways and both for the best-set and the all-set: The results may be weighted by sample size, significance, and the quality of the analysis. It can be assessed using the rating of the journal or citations of the paper as weights. In our case, the standard weights produce the same results. A key point of the summation is that data accumulates, so that as time passes

3. It is a firmly established tradition in economics (and other social sciences) that all statistical tests are presented as if calculated by one analysis run on virgin data. Thus everything looks much better than it is. This is also the case in the AEL, which follows the tradition of never mentioning this problem.

4. The three papers have an appendix with a brief introduction to the techniques and main tests of meta-analysis.

the results should converge toward the true values, while results due to the polishing of a random quirk in the data should collapse.

Purpose 2. The studies use different methods, i.e., models, data, regression techniques, etc., and these differences can be coded. It can then be analyzed to what extent differences in outcomes of the studies can be explained by differences in methodology. These results allow a discussion of methodologies with the purpose of choosing the superior ones, and hence perhaps adjust the aggregation of the results to obvious faults in the studies.

Purpose 3. To check a literature for biases due to priors. Table 5 lists the four main priors that often become significant in meta studies. All four priors appear in the AEL.

Table 5. Four priors analyzed in meta studies – all four also applies to journals

Prior	Explanation	Bias found in the AEL
<i>Ideology</i>	Authors of an ideological school that predicts an outcome have a prior for that outcome	Some authors do express political-ideological views and find results in accordance
<i>Author history</i>	Authors who have (or who belong to a group that has) previously written paper(s) in the field have a prior for similar results	50% of AEL authors participate in more than one paper. Several groups fighting for <i>their</i> model can be identified
<i>Institutional interests</i>	Authors working for an institution have a prior for results supporting its interests ^{a)}	35% of the research is financed from aid budgets. This gives a financial asymmetry
<i>Publication polishing</i>	Researchers have to publish to flourish, and journals want clear results, hence results are polished	General problem: Significance rises less with sample size than it “should”

Note a. Three reasons may be given: (1) Loyalty within organizations. (2) Selection/self-selection of organizations and employees. (3) Career pressures and other pressures of organizations on employees.

Authors often suspect that priors apply to journals – not to themselves. A journal may have an ideology: The *Journal of Women’s Liberation Studies* may accept a paper showing that women are *not* discriminated in a certain field, but the paper must then be of a very high quality, while the said journal accepts papers of a more dubious quality if they show that women are harmed by discrimination. Also, journals often have a history specializing in a certain family of studies, pushing a certain point of view. Further, some journals receive grants, and may not like to bite the hand that feeds them. And, we certainly know of authors who have been forced to shorten and sharpen papers by referees and editors. Hence meta studies normally include both variables for author characteristics and for publication outlet, and it is common that some of these become significant.

Table 6. Some characteristics of the AEL authors

Papers	Participation in		Probability of more ^{a)}	Origin of author	Nr
	Number			DC (OECD country)	
1	75	No more	50.0 %	Mixed ^{b)}	27
2	17	1 more	22.7 %	LDC	4
3	8	2 more	16.0 %	<hr/> Financing of research	
4	3	3 more	8.0 %	University	72
5	1	4 more	3.3 %	International organiz.	17
6+	0	5+ more	0 %	Other aid	12
All	104			Other	3

Note: a. Probability that author appears no more in the AEL, in 1 paper more, etc.

b. Author with non-DC origin now working in DC (mainly the USA).

Another point to note is that only 9 of the 104 authors are female.

The clearest *ideological/political factor* is that a handful of the early AEL writers state that they have a Marxist/left-wing ideology, or a Libertarian one.⁵ These authors have a prior for finding that aid harms the recipient country, though for different reasons. Left-wingers explain the poverty of LDCs by exploitation by DCs, and see aid as a factor in that process. Libertarians note that aid is given to public sectors in the LDCs, and see it as an inducement for government growth, planning and ultimately for socialism. Authors of the above two persuasions normally report negative aid effectiveness.

As seen from table 6, half of the 104 authors of the AEL appear in 2-5 papers. Also, many are members of groups. Recently several groups, with *institutional homes within the aid industry*, have been prominent in the research. These groups fight for a model that is *liked* by their institution to the extent that it has supported its propagation. The most prolific such group is the *World Bank Group* around David Dollar and Paul Collier, which produced seven papers, presenting and defending the Good Policy Model.⁶ The second is the *Danida Group* around Finn Tarp and Henrik Hansen, which produced four papers defending the Medicine Model.⁷ Both models are discussed in section 6.

Such groups typically keep demonstrating that their model is right, and that the other group's model is wrong. It obviously poses a problem for tests of significance of the findings of a literature that group behavior reduces the degrees of freedom in the tests. This is a prob-

5. The Marxist/left is a large and heterogeneous group, which has few economists among the *gurus*. However, the main *gurus* of the Libertarian school are economists. Some of those – notably Friedman (1958) and Bauer (1971) – have written very clearly about development aid.

6. Three more papers have been produced by “renegade” members of that group who have left the World Bank and now refute the group's model.

7. The DERG (Development Economics Research Group) at Copenhagen University is financed by Danida (Danish Development Aid Agency). The model was propagated by a grant to Tarp and Hjertholm (2000), and its institutional home is thus Danida. The World Bank Group similarly published World Bank (1998).

lem for the analysis covered in section 6, but it is less of a problem in sections 4 and 5, where the number of studies is high, and the results are clear anyhow.

The clearest cases of *institutional interests* in the AEL are that those working for/financed by the aid industry may have a prior for showing that aid works. Many donors reserve a small fraction of the aid budgets (currently of about \$60 billion) for development research. Even ½% of \$60 billion is still \$300 million. In the author statistics of table 6 we are not able to fully identify the institutional interests of all 104 authors, though for the last decade many researchers have created home pages with biographic information, and most papers have a note of acknowledgement. 78 of the researches only give a university affiliation. However, many university researchers do get outside funds, and even if a certain paper did not use such funds the author may have other grants, so it is likely that some of the 78 belong in the next two rows of the table. A conservative estimate is that 35% of the researches in the field work for the aid industry.

It is not unusual that research is financed by donations from organizations with a clear interest in the findings. However, most fields have other organizations with the reverse interests, working as a counterweight. It is worrying that even when other funds are available to finance AEL research, they have no correspondingly large countervailing interests, so funding affects the priors in this research in an *asymmetrical* way. The effect of the asymmetry of priors is an empirical question. We are glad to report that although it works in the direction predicted, and it often becomes significant, it is not very strong.⁸

Finally, there is the possibility that the pressures of academic careers make published results too good. We think that everybody who works in the “racket” of academia knows that the trade has its little tricks, like any other trade. We like to think that we can resist the temptation to use such tricks, and that the present paper is completely honest.

The temptation comes from the fact that a paper where the key coefficient α is estimated to 0.25 (1.69), where the bracket holds the t-ratio, is harder to sell than if α were 0.45 (2.57). So maybe, from looking at the residuals, we can discover, why economic theory makes it essential that a couple of observations are handled by a special dummy. Maybe a footnote could be inserted in the data appendix that these observations were omitted. Or maybe a variable could be squared or logged. Econometrics is a toolbox with many great tools that can persuade data to confess.

8. The average aid effectiveness, μ , is small too. We have tried to assess whether the bias generated by the “interest prior” can explain the size of the average μ . It is difficult to assess, but it is probably not the case.

Also, there is a moral hazard problem in the literature that uses models that contain a set of control variables that have to be chosen from a large set of possible variables. The voluminous literature on Barro growth empirics has now tried about 400 controls and of these about 60 has been tried in the AEL. This gives a large number of models to experiment with.⁹

The **FAT**-tests (funnel asymmetry tests) have been developed to show the presence of exactly that bias. The logic is that we know that small samples should have more variation in the results if the sample is small than if it is large, so estimated coefficients should lie within a funnel-shaped area, of a form that follows from the large sample variation in results. The funnel is wide for small samples and narrows as samples grow larger. Polishing is easier the smaller the sample, so the test for polishing is to study if there is an asymmetry in the funnel plot (Egger *et al.* 1997 and Stanley 2005).

These tests are normally significant in meta studies. They certainly are in our studies. It is a fact of life that people polish their goods to make them as shiny as possible to attract customers. It follows that results based on small samples are even less credible than indicated by the low number of observations.

The next three sections summarize our studies, following the classification in figure 2.

4. Results from Family A: Does aid cause increasing accumulation?

Family A of the AEL started around 1970, when development economists used Harrod-Domar models. They saw accumulation as the crucial factor in growth. The savings rate and subsequently the balance of payments are thus the key constraints for growth (see Chenery and Strout, 1966). Aid was meant to finance accumulation and hence moved both gaps outward, and it should thus greatly contribute to growth.

Several papers just before 1970 started a discussion of savings functions in poor countries and the effect of external inflows.¹⁰ As a part of the new savings literature, Griffin and Enos (1970) and Weisskopf (1972b)¹¹ demonstrated (on the scanty data then available) that aid flows decreased savings in the recipient countries by the same amount. The fungibility of

9. The typical number of controls is about 5. This gives $\binom{400}{5} \approx 8.3 \cdot 10^{10}$ or $\binom{60}{5} \approx 5.5 \cdot 10^6$ possible models.

10. The discussion was started by a remark by Haavelmo (1965) about the savings function in LDCs, which led to a discussion, see Rahman (1968); Ahmed (1971); Griffin (1970); Weisskopf (1972a). We now know that even if accumulation is an important factor of growth, it only explains between 25% and 50% of growth. Human capital, health, macroeconomic stability, technical progress and spillovers, social capital, etc., all contribute too.

11. The challenge was supported by arguments building on the “imperialism” school of thought, and notably Weisskopf was a supporter of the (then) New Left, and saw his results as a confirmation of his views.

aid permitted the marginal activity generated by aid *not* to be increased accumulation. If the key constraint for growth was accumulation, this was a major challenge to the justification of aid. This challenge corresponds to the one of Boone (1996) discussed in a moment.

Table 7. Two challenges to aid

Marginal activity caused by aid	Origin of challenge
C1 Aid reduces domestic savings by the same amount	Griffin and Enos (1970) and Weisskopf (1972b)
C2 Aid increases public consumption by the same amount	Boone (1996)

Note: While (C2) \rightarrow (C1) the reverse causality does not hold.

The savings challenge led to a wave of studies – see figure 3 – and this tradition has continued till this day, with many studies of the AEL having a section with some regressions analyzing the effect of aid on the rate of savings or investments. As listed in table 4, a total of 29 studies bring 90 savings regressions (incl. proxy studies), and 37 studies bring 122 investment regressions. 211 of these results can be made comparable. The standard savings-investment bookkeeping identity for an open economy looks as follows:

$$(1) \quad I - S = (I_p - S_p) + (I_G - S_G) = -XMB,$$

where I is investments, S is savings and XMB is the surplus on the goods and service balance and the subscripts p and G indicate the private and the (general) government sector.

In this framework aid, H , is a device that allows XMB to turn negative by H . It is, of course, given for that purpose. This will allow investment to rise by the amount of H , provided that S does not fall. If S falls by H the rise in I is crowded out. With the normalized variables $(s, i, h) = (S/Y, I/Y, H/Y)$ this gives the possible effects listed in table 8.

The challenge of Boone (1996) is that aid leads to an increase in public consumption only. As follows from equation (1) this causes the government savings rate to fall correspondingly, and it is thus one mechanism that explains the challenge of Griffin and Enos.

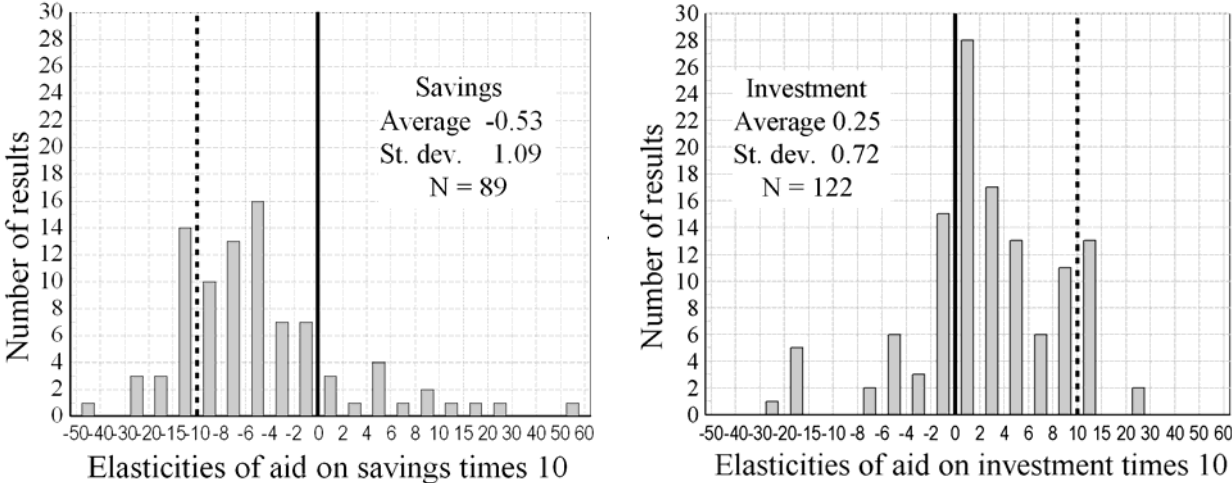
Table 8. Interpreting possible effects of the aid on savings and the investment

Effectiveness	Super	Full	Some	None	Harmful
Savings effect	effect > 0	0	0 < effect < -1	-1	effect < -1
Investment effect	effect > 1	1	1 < effect < 0	0	effect < 0

Note. The effects are expressed in percentage points of shares of the GDP, i.e. as elasticities.

Figure 5 shows the results. Both picture show an amazing range of results, and thus give a rather unclear picture. The investment graph has its highest peak just above zero, but then there is a secondary peak around 1, so on average there is probably a positive effect, but it is clearly well below 1. The savings graph is very similar, though shifted downward by 1 as it should. Most, but not all, of the effect is crowded out by a fall in savings. After further analysis we conclude that aid increase accumulation by about 25% of the aid, and that most of the remaining 75% lead to an increase in public consumption, and hence to a fall in public savings. Both effects are of dubious significance relative to 0 and -1 respectively.

Figure 5. The estimated effect of aid on either savings or investments



However, accumulation is only one explanation for growth, so the total effect on growth depends on what the remaining 75% of the aid do to the economy. As it basically leads to public consumption, it is likely to be a problem, because we know that public consumption has a negative effect on growth (see Barro and Sala-i-Martin 2004; pp 525-26). Also, papers in the AEL that include public consumption get a negative coefficient to that control.¹² Hence, from the analysis so far it is unclear if aid leads to development.

If we consider growth the key goal of aid, then surely it is better to study the effect directly. This is done in the other families of the AEL, most directly in family B.

12. It is an obvious misspecification to include public consumption among the controls, but it does, of course, cause an increase the estimate of aid effectiveness.

5. Results from Family B: Does aid cause increasing growth?

This family of models is a subfamily of the large empirical literature on cross-country growth models. That is, the literature summarized in Barro and Sala-i-Martin (2004; chaps. 10-12). The literature started as a study of convergence, using the much researched Barro equation. Here the variable of interest is the GDP level, $\log y_{it}$, which is used to estimate the coefficient of convergence β :

$$(2) \quad g_{it} = \alpha + \beta \log y_{it} + \gamma_{jit} \mathbf{x}_{jit} + u_{it}$$

This equation was then amended by adding $\log y_{it}$ to the controls and singling out another variable of interest, such as the aid share, h_{it} .

$$(3) \quad g_{it} = \alpha + \mu h_{it} + \gamma_{jit} \mathbf{x}_{jit} + u_{it}$$

Model (3) is the basic model that has been estimated to give the effect of aid, μ , in a total of 537 versions we have found in the B family of the AEL. Obviously very much can be said about (2) and (3), but at present we shall take the basic set-up as given, and just concentrate on the results, as summarized by the meta-analyses. For easy comparison we have converted all coefficients to elasticities. They thus show the effect in percentage points on the growth rate of an increase of one percent of aid.

Figure 6. The estimated effect of aid on growth

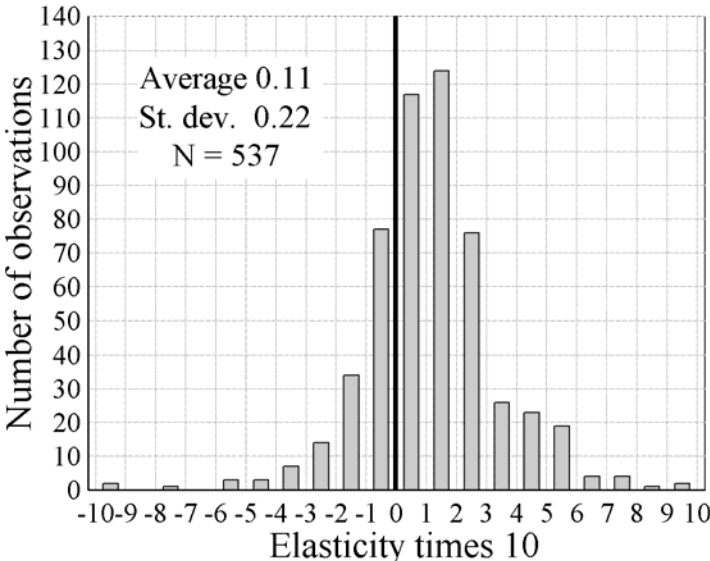


Figure 6 shows a remarkably broad range of results. We can read the figure as a monument to the ingenuity of our profession. From a set of raw data showing nothing (see figures 1a and b), it has proved possible to generate the distribution of results shown. The 537 observations have a positive average, but it is small relative to the variation, so it is not surprising that it proves insignificant in tests that adjust for dependencies between the results. When the elasticities are analyzed in the form of a funnel plot where the horizontal axis is the number of observations used in the calculation, it is obvious that most of the extreme values are found for small samples. This is also why the results are getting smaller and smaller as data have accumulated, as shown on figure 4.

When the results are analyzed for priors, we find that we cannot reject the possibility that the average result is biased upward by about 0.11 due to institutional priors, i.e., that the average result is due to the asymmetry of finance. Thus if we so adjust the results, they come very close to the simple regressions in table 2.

The main problem is that by all standard methods of summarizing the results, the small positive coefficient has remained insignificant. It does not matter if we look at all 537 regressions or the best regression from each of the 68 studies. The estimates are so volatile that they are no different from zero in the standard significance tests.

6. Results from Family C: Is the effect of aid on growth conditional?

One way to read figure 1 is to point out that it shows that aid helps in some cases and harms in other – equally common – ones. This is well in line with the impressions of old hands in the aid business: Sometimes aid works and other times it fails. Old hands think they can predict the outcome, though perhaps the forecasts are mainly *ex post*.¹³

So perhaps one can find a criterion, z , which can be transformed so that if $z > 0$ aid works, and if $z < 0$ aid harms. That is, the interacted variable, $h_{it}z_{it}$, gets a significant coefficient, ω , when model (4) is estimated:

$$(4) \quad g_{it} = \alpha + \mu h_{it} + \delta z_{it} + \omega z_{it} h_{it} + \gamma_{jit} \mathbf{x}_{jit} + u_{it}$$

13. Let us term growth above average *head* and growth below average *tail*. Imagine that the data for aid and growth are fully independent as suggested by the no frills regressions of table 1, then in about 50% of the flips, the coin lands on head and in the remaining cases on tail. If you examine 100 conditions that changed during the day you made a series of flips, then surely you will find some connections. Maybe you note that you got most heads on the flips when you were wearing a green tie, and your luck changed when you took it off. This is interesting, but it is hard to believe before a controlled experiment is made on new data.

The AEL has pointed to 10 candidates for the role of z during the last decade. Eight are examined in 1 or 2 studies, so they allow no meta study till now, but the other two models are the *Good Policy Model* covered by 22 studies (and 150 regressions), and the *Medicine Model* covered by 15 studies (and 85 regressions).

The Good Policy Model by Burnside and Dollar (1996; 2000) uses a special Good Policy Index as z . The index is a weighted sum of the budget surplus, the inflation rate and the trade openness, scaled to symmetric around a zero mean for the sample of countries and years analyzed. The Good Policy index is outcome-related so it is almost a tautology that the coefficient δ to z_{it} becomes positive and significant when model (3) is estimated. However, it was non-trivial when Burnside and Dollar reported that ω to $h_{it}z_{it}$ became significant and positive. The implication is that aid to countries with good policies helps the country, and help to countries with bad policies harms the country.

How much this message has actually affected World Bank lending since 1995 is not known, but it has probably had an effect, especially since it was propagated in World Bank (1998), and it has certainly received a lot of attention in discussions since then. The model has been vigorously defended by researchers in the *World Bank Group* (see section 3) in no less than 7 papers, but it has been demonstrated in the ensuing literature that it is a fragile model, and when the standard tools of meta-analysis are applied to the 22 papers and 150 regressions, it appears that the key coefficient of the model (that is ω to $h_{it}z_{it}$) is insignificant. In fact the model is unusually fragile to changes in sample, control variables, etc. Also, the decisive controls seem to be unconvincing.

The Medicine Model uses aid itself as the condition, so model (4) reduces to:

$$(5) \quad g_{it} = \alpha + \mu h_{it} + \omega h_{it}^2 + \gamma_{jit} \mathbf{x}_{jit} + u_{it}$$

The result of the proponents of the model is the two coefficients $\mu > 0$ and $\omega < 0$. That produces a reverse parable for excess growth with a positive section between $h = 0$ and $h = 2h^*$, and with a maximum for $h = h^*$. The marginal contribution of aid to growth is -2ω .

The aid squared term is propagated by *the Danida group*¹⁴ (see section 3) in four papers, and about 25% of the regressions in its support are actually found in papers of the group. This model is quite robust to reasonable changes in control variables as long as the original (rather limited) data set is used. However, it fares less well when the data set is

14. The model was discovered by Hadjimichael, Ghura, Mühleisen, Nord and Ucer (1995) in a paper on Africa. It has also been advocated by Lensink and White (2001).

expanded. The 15 papers and 85 regressions analyzing this model have failed to prove decisively that the two coefficients are statistically different from zero, though the results are just around the level of significance. If the criterion of independent replication on new data is used, it fails.

In short, the two leading new conditional models may not be anything but the mining of an arbitrary quirk in the data. However eight more models have been proposed and supported by some empirics. Time will tell if they hold up.

7. A parallel literature: Resource rent and Dutch Disease

Imagine the following thought experiment: An economist who knows everything except the AEL is asked: What piece of economic theory would be your first choice for analyzing the macroeconomic effect of aid? We think this economist would say: The theory about the transfer problem or Dutch Disease.

Development aid is an *external rent* that enters into the domestic economy. Before 1950 the effect of such transfers was discussed as the *transfer problem*. However, since then it has mainly been discussed in connection with resource rents received from exported resources. Here the subject is known as the *Dutch Disease* discussion or more ominously as the *Resource Curse*.¹⁵ The key result is that while a transfer certainly does increase the income *level* of the recipient – it is “paid for” by a decrease in the *growth* rate, making it less of an advantage in the longer run than it appears at first.¹⁶

The resource rent received by the LDC world is a couple of times larger than the aid received, and it is even more unequally distributed. The typical natural resource deposit has a long exhaustion time, but resource prices fluctuate to give considerable variation over time; whether aid or resource rents fluctuate more is unknown. Both resource rent and development aid are received primarily by LDC governments, and they are used to finance public spending in much the same way. To the extent that development aid is fungible it makes virtually no difference if the rent received comes as development aid or as a resource rent. Hence, the models used in the analysis should be similar, but we have found virtually no relation between the AEL and the Dutch Disease literature.¹⁷

15. Some references are Corden (1984), Sachs and Warner (1995) and Gylfason *et al.* (1999).

16. This is a frequent theme in the aid versus trade literature, where the trade generates dynamism and efficiency in the economy, while aid is a rent with negative effects, see Huges (2003).

17. The only exceptions we found are Younger (1992) and Eldabawi (1999) – two papers that have remained largely uncited in the AEL.

The Dutch Disease literature gives the real exchange rate a main role by demonstrating that a rent transfer inevitably leads to a real revaluation of the currency of the recipient country. Hereby its international competitiveness is reduced. Consequently there are losses to the economy outside the “booming” aid sector (that is the public consumption sector). The macro effects of aid are thus less favorable than predicted by the micro-macro result (P2 of the introduction). The Dutch Disease literature predicts a *micro-macro paradox*. The interesting question is consequently not the sign of the Dutch Disease effect of aid, but only the *size* of the effect.¹⁸

We note that the LDCs have had rather more inflation than the DCs, and more flexible exchange rates as well. So there is a story waiting to be told about the – small or big – role development aid has played in this development. It is actually documented in a few cases that aid has played an important role in the dynamics of prices and exchange rates, and hereby for the real economy. One well-documented case is that Tanzania was able to keep an unrealistically low exchange rate due to aid during the first half of the 1980s, with the predictable bad effects on the growth rate, till aid was temporarily stopped (see e.g. Paldam, 1997b).

This story is ruled out in the AEL by the way the models are set up. First it used the real 2-gap models (see section 3), where the exchange rate is fixed, and inflation has no effect by definition. Then it continued in the Barro tradition, which was developed from growth theory, which is real, though inflation is often used as an exogenous control variable.

The only use we have found of the observation that aid is a transfer of rents is in the political-economy type of models discussed at the end of section 3, and here it is used for a different purpose.

8. Disregarding significance: What do the numbers found mean?

Section 1 concluded that the average citizen in the LDC world has received aid in the order of 2½% of his/her annual income (GDP per capita). Thus the cumulated aid over the 40 years corresponds to one annual income.

18. The reader is referred to the case of the unsatisfactory development of the ex-DDR after it became a heavily subsidized part of Germany after the reunification (see Sinn, 2004) and the case of Greenland (see Paldam, 1997a). In the latter case, the Dutch Disease effect of a 50% aid share for now half a century may be assessed to have a 50% effect on the real exchange rate. Thus, with a 7% aid share in the average LDC it is possible that the real exchange rate has revalued by a similar amount. It will surely differ a lot from one country to the next, but it is – a priori – likely to be a substantial effect.

From section 4 we know that about 25% of aid becomes increased accumulation. With an aid share of 2½% this gives an extra share of accumulation of ½%. If we consider that the average rate of accumulation is about 12½%, aid raises that share by 4%. With a real rate of return of 5-10% that should increase growth by around 0.03%. It is not large, but it accumulates. If the average project has a lifespan of 10 years, a permanent flow of aid should thus add up to at most 0.3% of growth.

From section 5 we know that the best estimate we can make of the elasticity of the real product to aid is about 0.11. Hence, an aid share of 2½% should generate at most 0.3% extra growth. Both calculations are uncertain, but they are of the same magnitude, and over the 40 years of aid this produce $1.003^{40} = 1.27$, i.e. 27% higher income. As the numbers seems to be a bit optimistic 20% is likely to be closer to the true number.

If we accept these numbers, we reach the result that *the average LDC citizen may be 20% richer* due to aid since its start in the mid-1960s.

This number is, of course, very uncertain. It is worrying that these estimates have remained insignificant in spite of the large effort put into mining the data. It is also worrying that we may ascribe most of the difference between the no frills regressions of table 2 and the average results cited from sections 4 and 5 to an upward bias due to the asymmetry in the way this research is financed.

It should finally be mentioned that there appears to be enough evidence to conclude that the results differ between the different regions of the world. Aid is more effective in Asia and Latin American than in Sub-Sahara Africa. In the poorest region of the world aid is not pushing a development process well under way, but trying to start such a process and that is particularly difficult.

9. What to do?

Most of the analysis above has been gloomy. In spite of a thorough mining of the data for aid and growth by no less than 104 researchers over almost 40 years, the AEL (the aid effectiveness literature) has failed to show a clearly significant effect from aid to growth. The average effect found is positive, but it is small. No doubt many economists would argue that since the coefficient is small and uncertain, we should conclude that aid has failed.

However, section 8 did show that if we disregard significance and take the average found literally, 40 years of aid may have increased the standard of living of the average LDC

citizen by 20%. This is an optimistic assessment. If it is true it is a major contribution, but it is considerably less than we have all hoped for.

When the effect is compared with the one of the not-aid-related growth of China and India, then surely an increase of 20% over 40 years is modest indeed. The growth of the two giants over the last 10 years alone has increased the GDP of the LDC world by approximately 50%, and their own GDP by well over 100%. So 20% over 40 years is not impressive. Thus, there are certainly good reasons for improving the effectiveness of aid.

Development aid is an activity that has proved difficult to do right. When something is difficult it is of paramount importance that it is transparent, i.e., that it is done by simple, clear and easily controllable rules.

However, aid is surrounded by complex politics. In order to attract popular support in donor countries, it caters to all kinds of lofty and continuously shifting goals mixed up with stakeholder and strategic interests. In the aid discourse, the air is often stale and muggy from all the big, sweet and vague words that steadily shift. Even if it would cost some support in the short run, *it would prevent aid fatigue in the longer run if aid became more effective, and it could be found to work.*

References:

- Ahmed, N., 1971. A note on the Haavelmo hypothesis. *Review of Economics and Statistics* 53, 413-14
- Barro, R.J., Sala-i-Martin, X., 1998, 2004. *Economic growth*. MIT Press, Boston
- Bauer, P.T., 1971. *Dissent on development: Studies and debates in development economics*. Weidenfeld and Nicolson, London. Sums up several older papers
- Boone, P., 1996. Politics and the effectiveness of foreign aid. *European Economic Review* 40, 289-329
- Burnside, C., Dollar, D., 2000. Aid, policies and growth. *American Economic Review* 90, 847-68. Working paper available since 1996
- Cassen, R., and associates. 1986, 1994. *Does aid work?* Oxford UP, Oxford, UK
- 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
- Corden, W.M., 1984. Booming sector and Dutch disease economics: review and consolidation. *Oxford Economic Papers* 35, 359-80
- Doucouliaqos, H., Paldam, M., 2005a. Aid effectiveness on accumulation. A meta study¹⁹
- Doucouliaqos, H., Paldam, M., 2005b. Aid effectiveness on growth. A meta study¹⁹
- Doucouliaqos, H., Paldam, M., 2005c. Conditional aid effectiveness. A meta study¹⁹
- 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
- Elbadawi, I.A., 1999. External aid: Help or hindrance to export orientation in Africa. *Journal of African Economics* 8, 578-616
- Friedman, M., 1958. Foreign economic aid: Means and objectives. *Yale Review* 47, 500-16
- Griffin, K., 1970. Foreign capital, domestic savings and economic development. *Bulletin of the Oxford University Institute of Economics and Statistics* 32, 99-112
- Griffin, K., Enos, J., 1970. Foreign assistance: Objectives and consequences. *Economic Development and Cultural Change* 18, 313-27
- Gylfason, T., Herbertson, T.T., Zoega, G., 1999. A mixed blessing: Natural resources. *Macroeconomic Dynamics* 3, 204-25
- Haavelmo, T., 1965. Comment to W. Leontieff. The rates of long-run economic growth and capital transfers from developed to underdeveloped areas. Section 13, 1039-1068 in *The Econometric Approach to Development Planning*. Pontificia Academiae Scientiarum Scripta Aaria 28. North-Holland Publishing, Amsterdam
- 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
- Herbertsson, T.T., Paldam, M., 2005. Does development aid help poor countries catch up? An analysis of the basic relations¹⁹

19. Working paper from Department of Economics, Aarhus University, or from <http://www.martin.paldam.dk>.

- Huges, H., 2003. Trade or aid? Which benefits developing countries more? *Economic Papers – Economic Society of Australia*, 22 pp, Sept
- Lensink, R., White, H., 2001. Are there negative returns to aid? *Journal of Development Studies* 37, 42-65
- Mosley, P., 1986. Aid effectiveness: The micro-macro paradox. *Institute of Development Studies Bulletin* 17, 214-25
- Paldam, M., 1997a. Dutch disease and rent seeking: The Greenland model. *European Journal of Political Economy* 13, 591-614
- Paldam, M., 1997b. *Dansk u-landshjælp. Altruismens politiske økonomi*. Aarhus UP, Aarhus
- Paldam, M., 2005. Appendix. Available from <http://www.martin.paldam.dk>
- Rahman, A., 1968. Foreign capital and domestic savings: A test of Haavelmo's hypothesis with cross-country data. *Review of Economics and Statistics* 50, 137-38
- Sachs, J., Warner, A.M., 1995. *Natural resource abundance and economic growth*. NBER Working Paper 5398. Published in several versions
- Sinn, H.-W., 2004. *Ist Deutschland noch zu retten?* Econ-Verlag, Berlin
- Stanley, T.D., 2005. Beyond publication bias. *Journal of Economic Surveys* 19, 309-45
- Tarp, F., Hjertholm, P., eds. 2000. *Foreign aid and development. Lessons learnt and directions for the future*. Routledge Studies in Development Economics: London
- WDI, annual. *World development indicators*. CD from the World Bank, Washington DC
- 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
- World Bank, 1998. *Assessing aid: What works, what doesn't, and why*. A Policy Research Report. Oxford UP, New York
- Younger, S.D., 1992. Aid and the Dutch Disease: Macroeconomic management when everybody loves you. *World Development* 20, 1587-97