Journal of Globalization and Development

Volume 1, *Issue* 2 | 2010 | *Article* 5

Aid, Growth, and Development: Have We Come Full Circle?

Channing Arndt, *University of Copenhagen*
Sam Jones, *University of Copenhagen*
Finn Tarp, *University of Copenhagen and UNU-WIDER*

Recommended Citation:
Arndt, Channing; Jones, Sam; and Tarp, Finn (2010) "Aid, Growth, and Development: Have We Come Full Circle?," *Journal of Globalization and Development*: Vol. 1: Iss. 2, Article 5.

DOI: 10.2202/1948-1837.1121

Aid, Growth, and Development: Have We Come Full Circle?

Channing Arndt, Sam Jones, and Finn Tarp

Abstract

The micro-macro paradox has been revived. Despite broadly positive evaluations at the micro- and meso-levels, recent literature doubts the ability of foreign aid to foster economic growth and development. This paper assesses the aid-growth literature and, taking inspiration from the program evaluation literature, we re-examine key hypotheses. In our findings, aid has a positive and statistically significant causal effect on growth over the long run, with confidence intervals conforming to levels suggested by growth theory. Aid remains a key tool for enhancing the development prospects of poor countries.

KEYWORDS: foreign aid, growth, aid effectiveness, causal effects

Author Notes: We have benefited greatly from editorial advice and a careful anonymous referee report. We also thank Tony Addison, Ernest Aryeetey, Pranab Bardhan, Bruce Bolnick, Imed Drine, Jan Willem Gunning, Gerry Helleiner, Paul Isenman, Homi Kharas, Dirk Krueger, David Roodman, Erik Thorbecke, Alan Winters, and Adrian Wood for encouragement and most valuable comments. The same goes for the participants at conferences and seminars held by the African Economic Research Consortium (AERC), the Bergen Resource Center for International Development at CMI Norway, the Brookings Institution, central ministries in Mozambique, Tanzania and Vietnam, Cornell University, the European Union Development Network (EUDN) (organized by CERDI, Clermont-Ferrand, France), the Helsinki Center for Economic Research (HECER), NORAD Norway, OECD Paris, the United Nations HQ (organized by UNU-ONY), the University of Copenhagen, the University of Ghana, the UNU Conference of Directors (CONDIR), and UNU-WIDER. Thanks are also due to Tseday Jemaneh Mekasha for excellent research assistance and to Raghuram G. Rajan and Arvind Subramanian for sharing their original data and STATA files. The usual caveats apply.

1. Introduction

The extent to which foreign aid can be a decisive factor in the economic development of low-income countries remains controversial. In 1987, Paul Mosley suggested that while aid seems to be effective at the microeconomic level, any positive aggregate impact of aid is much harder to identify (Mosley, 1987). He labeled this the "micro-macro" paradox. Now, after two decades of intense analytical work using new theory, new data, and new empirical methodologies, it would appear that the paradox has been revived. At the macro-level, Rajan and Subramanian (2008) conclude "it is difficult to discern any systematic effect of aid on growth". Nevertheless, evaluations of aid effectiveness at the microeconomic level continue to indicate positive rates of return (World Bank, 2008). Also, an increasing number of rigorous microeconomic impact evaluations have demonstrated the potential for well-designed project interventions to generate positive results (Banerjee and Duflo, 2009).

This paper has two objectives. First, we attempt to provide a balanced, up-to-date assessment of the aid-growth literature. We observe that, while aid may have very high returns in specific circumstances (Collier and Hoeffler, 2004), expectations surrounding the average potency of aid have been excessive. Second, we turn our attention to the fundamental empirical evaluation challenge: identifying the counterfactual. Using observational data, there is no way of identifying a plausible counterfactual without making assumptions that are bound to be debatable, in theory and practice. Applying recent methodological advances in the program evaluation literature to a macroeconomic question, we propose and implement a number of empirical improvements. These include strengthening the aid instrument, improving the model specification, and employing a new doubly robust estimator. We show, contrary to other recent results, that there is no firm basis on which to reject the prior that foreign aid exerts a modest long-run positive effect on growth in developing countries.

The remainder of this paper is structured as follows. Section 2 provides a literature review. Section 3 considers empirical strategies used to investigate the aid-growth relationship. Section 4 implements our proposed enhancements and presents results. A final section summarizes and concludes that there is no micro-macro paradox.[1]

[1] Please refer to Arndt *et al.* (2010) for an earlier working paper version of this paper for further details on background, analytical methods, data, sensitivity tests, and results.

2. Literature review

Scholarship on the relationship between aid, growth, and development is voluminous. This section provides a brief survey of how the aid-growth debate has evolved.

2.1. Earlier generations

Studies of the aid-growth relationship from the 1970s until recently have been classified into three generations, each influenced by dominant theoretical paradigms as well as available empirical tools (Hansen and Tarp, 2000). The first two generations were inspired by simple models of the growth process, i.e. the Harrod-Domar model and the two-gap Chenery-Strout extension. The underlying idea behind the Harrod-Domar model is of a stable linear relationship between growth and investment in physical capital. Assuming all aid is invested, it is straightforward to calculate how much aid is required to achieve a target growth rate. The impact of aid is positive and helps plug either a savings or a foreign exchange gap. Empirical studies in this tradition consequently focused on the extent to which aid increased savings and investment in recipient countries (Papanek, 1972, 1973). Overall, first generation studies show that aid tends to increase total savings, but not by as much as the aid flow. Quite reasonably, this suggests a non-negligible proportion of aid is consumed rather than invested.

Retaining the focus on capital accumulation, a second generation of literature explored the impact of aid on growth via investment. Using data for a cross section of countries, a large number of studies of this kind were produced during the 1980s and early 1990s. These studies consistently pointed to a positive link between aid and investment. While a majority of the aid-growth studies of this generation also suggested a positive impact, the result that captured attention was Paul Mosley's micro-macro paradox. An influential line of critique of the Harrod-Domar and two-gap approach was the argument that growth is less related to physical capital investment than often assumed (Easterly, 1999). If the productive impact of aid depends more on incentives and relative prices, as well as the policy environment more generally, then it becomes important to consider these broader effects. The second generation of studies also introduced the problem that poorly performing countries may receive more aid precisely *because* of their poor growth performance. Empirical analyses that do not account for the endogeneity of aid will not reveal aid's causal impact. Most second generation studies, however, did not deal with this issue.

From the early 1990s a third generation of more sophisticated econometric studies came to dominate the discourse about aid. This was motivated by the availability of panel data, allowing analysts to look at changes both across and within countries over time. Insights from new theories of economic growth also influenced the research agenda. Mindful of the weaknesses of previous studies, the aid-growth relationship came to be perceived as (possibly) non-linear and the endogeneity of aid was taken more seriously. Among the numerous studies of this generation, the contribution by Burnside and Dollar (2000) came to exert a significant influence on policy. These authors made an argument for conditional aid effectiveness, specifically: "aid has a positive impact on growth in developing countries with good fiscal, monetary and trade policies ... [but] ... in the presence of poor policies, aid has no positive effect on growth" (2000: 847).

However, these results were subject to criticism. Hansen and Tarp (2001) found that a story of diminishing returns to aid best captured the non-linear relationship between aid and growth. In a later contribution, Easterly *et al.* (2004) added that the Burnside-Dollar aid-policy result is fragile when the dataset is expanded (by years and countries). Dalgaard *et al.* (2004) found that aid has been less effective in tropical areas over the last 30 years and called for further research to help disentangle the channels through which aid matters for productivity. In an empirical review of these contributions, Roodman (2007) argued that the results of this generation are extremely sensitive to methodological choices, concluding that while some aid is likely to increase investment and growth, aid "is probably not a fundamentally decisive factor for development" (2007: 275).

2.2. Recent studies

More recently, a fourth generation of literature has emerged. A distinctive aspect of this generation is the view that aid's aggregate impact on economic growth is non-existent. A leading paper that appears to establish this result is Rajan and Subramanian (2008). They find no systematic effect of aid on growth regardless of the estimation approach, the time period. and the type of aid. Explanations for non-positive aggregate effects of aid often refer to political economy dynamics. For example, Djankov *et al.* (2008) argue that aid has effects that are analogous to a natural resource curse. Similarly, Rajan and Subramanian (2007) find that the rate of growth of value added by the manufacturing sector in developing countries has been undermined by a detrimental effect of aid inflows on governance quality.

Fourth generation scholars have also become increasingly skeptical about our ability to make valid causal inferences with respect to complex aggregate phenomena, such as the determinants of economic growth. In particular, previous methods used to deal with endogeneity have been subject to criticism. There is increasing awareness that dynamic panel (system) GMM methods – frequently employed in the third generation – are not a panacea. The concern that weak instruments typically bias coefficient estimates towards their unadjusted counterparts (e.g., OLS or panel fixed effects estimates) applies as much to panel GMM as to cross-section estimators. Bun and Windmeijer (2010) show that the weak instrument problem (previously attributed mainly to the Arellano-Bond estimator) may be equally problematic in the system approach. Also, for the Blundell-Bond (system GMM) estimator to be valid, both country fixed effects and omitted variables must be orthogonal to the lagged differences of the right-hand side (RHS) variables which are used as instruments for the level equation. This assumption cannot be tested and may be suspect given (i) the highly complex nature of the growth process, and (ii) that country fixed effects are expected to incorporate determinants of steady-state income levels that may correlate with growth along individual countries' steady-state transition paths.

In a Monte Carlo investigation of the robustness of different panel estimators, Hauk and Wacziarg (2009) conclude that the principle issue for system GMM is not one of strong or weak instruments but the validity of these moment conditions. Roodman (2009) warns that the Blundell-Bond estimator may give a false sense of certainty as a large number of internal instruments can over-fit the endogenous variables and may weaken the power of Hansen/Sargan tests. Finally, internal instruments do not prevent bias arising from systematic measurement error in the endogenous regressors, which is an important limitation in the context of aid-growth regressions.

In response to these concerns, alternative methods to assess aid effectiveness have come to the fore. These often eschew cross-country macroeconomic analysis in favor of specific micro- and meso-outcomes (Temple, 2010). Mishra and Newhouse (2007), for example, uncover a small but statistically significant effect of health aid on infant mortality. Masud and Yontcheva (2005) also find that aid helps reduce infant mortality, but this effect is only significant for aid provided by non-governmental organizations (NGOs). Alongside cautious optimism surrounding the potential efficacy of microeconomic policy interventions, financed either directly by donors through projects or indirectly via budget or sector-wide support, these findings give new sustenance to the micro-macro paradox. Indeed, with few exceptions (e.g., Sachs, 2005), recent findings at the micro- and meso-levels have not been deployed to argue that aid is effective on aggregate. This is despite increasing evidence that meso-level outcomes can add up to substantial macroeconomic effects (Cohen and Soto, 2007).

2.3. Rationale for continued macro-analysis

In light of the methodological concerns raised by recent scholarship, it is helpful to reflect on the value of attempting to assess the macroeconomic impact of foreign aid. In the first place, serious empirical challenges have not dissuaded economists from investigating other complex questions (Angrist and Pischke, 2010). For example, there are numerous parallels between the problem of estimating the causal impact of aid on growth and the causal impact of schooling on earnings. Both problems are likely to be characterized by endogenous selection, heterogeneous treatment responses, and mis-measurement of treatment input (both in terms of quality and quantity). Considerable effort has been expended in the analysis of large, high-quality, schooling datasets by some of the most skilled econometricians in the profession. Even so, debate has persisted, at least until recently, with respect to the net bias of ordinary least squares (OLS) estimates of returns to education (Card, 2001).

If the profession has experienced serious difficulties estimating the causal effect of schooling on earnings in developed countries, then it should not be surprising that estimating the impact of aid on growth in developing countries is contentious. However, it is difficult to deny that the aid-growth issue is both compelling and relevant. In developed countries, policy-makers and the wider public continue to ask whether aid is a cost effective use of taxpayer money on aggregate. Today, the attention of both the aid community and decision-makers is on “Dead Aid” (Moyo, 2009), which argues for a complete cessation of aid flows to Africa. We note that the financial crisis of 2008/09 has highlighted the importance of public spending to stabilize and stimulate economic activity. While foreign aid has multiple objectives, economic growth is central among them. If the economics profession as a whole were to abandon the question of aid’s impact on growth, it would leave the issue open to speculative and potentially unhelpful contributions.

2.4. Formulating an appropriate prior

A key aim of empirical analysis is to falsify or discriminate between competing hypotheses. Consequently, it is necessary to make explicit the prior upon which empirical testing is focused. With respect to the effect of foreign aid on economic growth, relatively few studies address the issue of an appropriate prior. A recent exception is Rajan and Subramanian (2008) who consider aid in a standard neoclassical growth model. Assuming that aid only augments physical capital investment and has no effect on productivity, they derive that a one percentage point increase in the ratio of aid to GDP should be expected to raise the growth rate of per capita GDP by around 0.16 percentage points on average. In practice, at least some aid is directed towards consumption or non-growth enhancing activities. As a result, Rajan and Subramanian place the expected growth return at around 0.1 percentage point for each percentage point of aid in GDP. Thus, the implied increase in the growth rate accruing from aid inflows at 10% of GDP should be about 1%, which is considerably less than the predictions based on Harrod-Domar models. In sum, growth theory points towards more modest expectations with respect to the potency of aid (see also Dalgaard and Erickson, 2009).

A related issue is the appropriate time-frame over which any growth effects accruing from aid can be expected to materialize. Various factors may exert a cumulative but not immediate impact on the rate of income growth. For example, changes in education move only slowly at the aggregate level and have a positive influence on economic growth with a substantial lag. This follows from simple demographics whereby improvements in schooling indicators can take many years to translate into noticeable increases in average education levels among working-age adults. Changes in human capital due to improved health indicators may take even longer to translate into more rapid economic growth. Ashraf *et al.* (2008) and Acemoglu and Johnson (2007) find that the initial economic impact of gains in life expectancy from the health interventions introduced from the 1940s may be a reduction in *per capita* incomes due to the increase in population and dependency ratios. The former authors find that it can take 30 years or more for per capita incomes to return to pre-intervention levels. They also find that significant increases in life expectancy at birth only begin to have a modest positive effect on incomes after about a 35 year lag.[2]

Overall, a series of considerations indicate that the aid-growth relationship is only likely to emerge over a long time-horizon. Many aid investments, such as in education, health, and institution-building are long term in nature; and growth theory indicates that the contribution of these investments to growth is likely to be relatively modest. When these observations are combined with the volatility of growth in most developing countries and the high degree of measurement error inherent in nearly all the variables of interest, relatively long time-frames would appear to be necessary to reliably detect the aid-growth relationship.

3. Empirical strategies

Following Temple's (2010) recommendation to build explicitly on existing empirical work, our starting point for developing an appropriate empirical strategy is Rajan and Subramanian (2008) (henceforth RS08). RS08 provides a thoughtful and highly influential contribution that is widely understood to have established that aid has no impact on growth. However, a number of concerns question this fundamental conclusion. This section addresses these concerns and suggests suitable improvements. Specifically, Section 3.1 provides a brief summary of RS08's approach and main results. Section 3.2 presents a detailed analysis of the validity of their instrumentation strategy. This motivates the development of an amended instrument in Section 3.3, while Sections 3.4 and 3.5 propose modifications to the specification and regression estimators. Once these

[2] Ashraf *et al.* (2008) focus on demographic trends as a result of disease eradication. Productivity effects, demand effects, and complementary policies may speed the realization of growth benefits from health gains.

modifications are incorporated, either individually or in conjunction, the empirical aid-growth relationship is shown to conform to theoretical priors and the micro-macro paradox disappears once again. These results are given in Section 4.

3.1. RS08

Before turning to RS08's main results, it is helpful to examine simple OLS estimates of the relationship between aid and growth. They are reported by RS08 and are replicated here in column I of Table 1. As expected, they show a negative estimated coefficient on Aid/GDP. These estimates are biased downward principally due to a simultaneity problem – *because* of their slow growth, slower growing countries typically received more aid during the period. Thus, the challenge is to find valid and relevant external instruments that explain variation in aid receipts across developing countries, but which are unrelated to their growth performance.

In their principal IV approach, RS08 investigate the aid-growth relationship using long-run averages, turning the focus away from dynamic panel methods. This is sensible. Dynamic panel methods are subject to doubt given the expected cumulative effect of aid and corresponding concerns regarding the validity of internal instruments in GMM estimators. Their long-run approach echoes the average OLS estimator proposed by Mankiw, Romer and Weil (1992) as well as the long-difference approach used by Acemoglu and Johnson (2007).

RS08 consider four periods separately: 1960-2000; 1970-2000; 1980-2000; and 1990-2000. In each period, their instrument for aid is generated from a "zero" stage regression estimated at the bilateral donor-recipient level using supply-side factors. These include past colonial relations, relative population sizes, and interaction terms. The predicted Aid/GDP ratio estimated from this regression is aggregated across donors to give a fitted average ratio for each recipient. This is then used as a single excluded instrument in a 2SLS estimation, where average growth over the period is the dependent variable. Core results from RS08 for the 1970-2000 period are replicated in column II of Table 1. The generated instrument appears reasonably strong according to conventional measures, such as the first stage partial F-statistic; also, the coefficient on Aid/GDP is exactly in line with the prediction from their growth model but is not significant.

RS08 conclude that there is no systematic (causal) effect of aid on growth, and move on to show this holds for alternative sub-periods (RS08 Table 4), alternative growth horizons (RS08 Table 6), non-linear effects (RS08 Table 7), and different types of aid (RS08 Table 8). Similarly, the same basic result emerges when the question is considered in a dynamic panel setting (RS08 Table 10).

Table 1: Alternative models for aid and growth, 1970-2000

	(I) OLS	(II) 2SLS	(III) IV-LIML	(IV) IV-LIML	(V) IV-LIML	(VI) IV-LIML	(VII) IV-LIML	(VIII) IV-LIML
Aid / GDP	-0.08*	0.10	0.10	0.07	0.08	0.12	0.21*	0.19*
	(0.04)	(0.07)	(0.08)	(0.06)	(0.06)	(0.10)	(0.11)	(0.10)
Initial per cap. GDP	-1.67***	-1.41***	-1.40***	-1.44***	-1.44***	-1.44***	-1.34***	-1.36***
	(0.32)	(0.43)	(0.39)	(0.37)	(0.36)	(0.30)	(0.33)	(0.33)
Initial level of policy	2.28***	2.14***	2.13***	2.16***	2.15***	2.28***	2.29***	2.28***
	(0.47)	(0.62)	(0.56)	(0.52)	(0.53)	(0.46)	(0.52)	(0.51)
Initial life expectancy	0.02	0.08*	0.08**	0.07**	0.07**	0.04	0.05	0.05
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Geography	0.39**	0.61**	0.62**	0.58***	0.58***	0.25	0.29	0.28
	(0.18)	(0.26)	(0.24)	(0.22)	(0.22)	(0.22)	(0.24)	(0.23)
Excluded instruments	1	1	7	3	1	7	3	1
Regional dummies	SSA, EA	SSA, EA	SSA, EA	SSA, EA	SSA, EA	SSA, A, LA	SSA, A, LA	SSA, A, LA
N	78	78	78	78	78	78	78	78
R-squared	0.70	0.59	0.58	0.62	0.61	0.69	0.65	0.66
Weak identification stat.	-	31.6	5.50	10.40	33.41	4.32	5.06	14.32
Stock-Wright LM S stat.	-	-	11.85	3.05	2.12	18.17	5.40	5.20
(probability)			0.158	0.384	0.145	0.020	0.145	0.023
Hansen J stat.	-	-	7.72	0.29	-	11.88	0.31	-
(probability)			0.358	0.865		0.104	0.857	

significance level: * 10%; ** 5%; *** 1%

Notes: only selected variables reported; columns (I) and (II) replicate results in Rajan and Subramanian (2008; RS08); column (III) employs a full set of aggregate instruments (see Table 2) in place of a single generated instrument; column (IV) restricts the aggregate instrument set to the mean population ratio, the colony dummy and their interaction; column (V) uses only log initial (recipient) population as the instrument; columns (VI) to (VIII) replicate columns (III) to (V) but use a preferred set of conditioning variables (not shown, see Section 3.4); regional dummies are included as indicated, SSA = sub-Saharan Africa, EA = East Asia, A = Asia, LA = Latin America & Caribbean; weak identification statistic is the first stage partial-F statistic in col. (II), and the Kleibergen-Paap Wald F statistic elsewhere; initial policy refers to the Sachs-Warner trade policy index; geography refers to the average of the number of frost days and tropical land area; standard errors, given in parentheses, are robust to arbitrary heteroskedasticity; dependent variable is mean real growth rate; Aid/GDP is treated as endogenous in all models except column (I); in columns (I) to (V) Aid /GDP is taken from RS08, in columns (VI) to (VIII) it is re-estimated from OECD-DAC (2008) data treating possible missing values as zeroes (see Section 3.3).

Source: authors' estimates.

3.2. Instrument validity

The supply-side approach to instrumentation developed by RS08 represents the state of the art in the aid-growth literature. Nevertheless, it has been subject to criticism. Clemens and Bazzi (2009) note that different authors have used the same variables as exogenous instruments for a wide range of endogenous variables. This raises the possibility that these exogenous instruments are correlated with other omitted variables, thereby invalidating the exclusion restriction on which valid causal inference depends. They direct specific attention to the reliance of the RS08 (fitted) instrument on the natural logarithm of the aid recipient's population size. They find that log population has a "statistically significant partial relationship with several variables that are plausible growth determinants" (2009: 11) and that are omitted from RS08's specification.

While the existence of a partial correlation between omitted explanatory variables and the chosen instrument indicates that the coefficients in a regression specification *may be* biased, the extent of bias is, ultimately, an empirical matter. This is recognized by Clemens and Bazzi (2009), leading them to advocate application of a range of empirical tests for instrument validity. On the face of it, straightforward validity checks of the RS08 (generated) instrument based on Sargan or Hansen tests are not possible because their IV model is just identified – i.e., the number of excluded instruments equals the number of endogenous variables. Nonetheless, recalling that the zero stage of the RS08 approach generates a single instrument as a linear combination of variables, it is possible to use modified versions of these same variables as excluded instruments directly in the aggregate aid-growth regressions. This provides for a large number of potential instruments, and therefore permits Hansen tests to be run either on the full set or on specific sub-sets of instruments.

Following this logic, we collapse the bilateral aid dataset along the donor dimension, thereby transforming the explanatory variables used in the bilateral zero stage regressions for use at a more aggregate level. For continuous zero stage regressors, such as the donor-recipient population ratio, the corresponding "aggregate" instrument is the mean of the population ratio for each recipient across all donors. For dummy regressors, such as the specific colonizer, it is more appropriate to take the maximum value of the dummy for a given recipient (again, across all donors). Ignoring relatively minor variables, such as currently being a colony and the population-colony interaction terms employed in RS08, this yields a set of eight possible instruments as per the rows of Table 2.

Column I of Table 2 verifies whether these aggregate instruments are adequate proxies for the fitted instruments generated from the zero stage regressions. As expected, the explanatory power is high. Moreover, underlining the contention of Clemens and Bazzi (2009), a driving force behind the fitted aid instruments appears to be the population ratio term. Thus, a fundamental issue for the RS08 instrumentation strategy is the validity of the exclusion restriction as it applies to the population-based instruments. Nevertheless, the results from column I of Table 2 indicate that other variables make some (albeit smaller) contribution to the overall fitted instrument. Thus, to further test instrument validity, we re-estimate the RS08 model employing the full set of eight aggregate instruments. The results, reported in column III of Table 1, closely replicate column II; and the Hansen J test reports a probability of 0.358, which fails to reject the validity of the exclusion restriction assumption.

Table 2: Instrument validity checks, 1970-2000

	Fitted coefficients		Residual coefficients		RS08 model		AJT model	
	RS08 (I)	AJT (II)	RS08 (III)	AJT (IV)	C stat. (V)	Prob. (VI)	C stat. (VII)	Prob. (VIII)
Population ratio	0.85***	0.59***	-0.18	-0.03	0.40	0.53	0.24	0.63
Colony (ever)	0.10**	0.01	-0.19	-0.25	1.11	0.29	1.96	0.16
Pop. ratio × colony	0.11	0.32***	0.10	0.09	0.48	0.49	0.06	0.81
Common language	0.04	0.21***	0.25*	0.26*	2.36	0.12	0.27	0.60
Spanish colony	-0.08	0.03	0.26	0.18	0.29	0.59	1.36	0.24
Portuguese colony	0.03	0.04*	0.27**	0.08	3.28	0.07	1.77	0.18
French colony	-0.01	0.11***	0.53**	0.71***	2.14	0.14	8.13	0.00
UK colony	-0.25***	0.11**	0.35	0.42**	0.02	0.89	1.02	0.31
R-squared	0.95	0.99	0.12	0.23	-	-	-	-

significance level: * 10%; ** 5%; *** 1%

Notes: columns (I) and (II) report standardized OLS regression coefficients in which the dependent variable is the fitted aid instrument taken from Table 3 columns (I) and (V) respectively; columns (III) and (IV) report standardized OLS regression coefficients from regressions of residuals saved from columns (IV) and (VI) of Table 1 against the row variables; columns (V) to (VIII) report individual difference-in-Hansen C statistics and corresponding probabilities associated with each individual row instrument (relative to the full instrument set); all OLS regression specifications incorporate the relevant set of included instruments as additional controls (not reported); significance from OLS regressions are based on robust standard errors.

Source: authors' estimates.

Nevertheless, we take one further step to investigate the exclusion restriction. Following the intuition of Sargan-type tests, we save the residuals and regress them against the set of excluded instruments. This provides initial insight as to *which* variables in the instrument set may be suspect. Standardized coefficients from these regressions are given in column III of Table 2. They show that neither the population ratio term nor its interaction with the (ever being a) colony dummy is significantly correlated with the unexplained components of the growth models. In contrast, the Portuguese and French colonizer effects are significant. In our alternative specification (denoted AJT) shown in column IV of Table 2, the French and UK colonizer terms are both significant.

Following this simple but intuitive OLS approach, columns V to VIII of Table 2 report formal tests of the orthogonality of each of the individual aggregate instruments to the growth regression errors. Specifically, they report the difference-in-Hansen C statistic associated with excluding each row instrument (individually) from the full set. For example, in the first row the C statistic corresponds to the reduction in the overall Hansen J test statistic when the population ratio term is excluded from the instrument set; the corresponding probability is also shown. These findings corroborate the residual-based OLS results. The most suspicion falls on the colonizer terms; however, the population ratio variables do not give cause for concern, providing comfort as to their suitability as exogenous instruments in these models.

3.3. Improved instrumentation strategy

The results of Section 3.2 indicate that the RS08 instrumentation approach is broadly convincing but is weakened by inclusion of suspect variables in the zero stage. From a theoretical point of view, the validity of using colonizer dummies and their interactions as instruments is questionable. The institutional transplants and broader colonizing strategies pursued by imperial powers were not alike, and they may have a persistent effect on income levels to the present day. This notion is at the heart of the debate concerning the effect of different legal origins (La Porta *et al.*, 2008), historical events (e.g., Nunn, 2008), and other institutional forms on contemporary economic outcomes. Put simply, the colonial relations variables are not orthogonal to growth and therefore should not be included in the zero stage regression explaining aid.

As a first step towards improving the RS08 instrument, we re-run their aid-growth model using a smaller and "less suspect" sub-set of the aggregate instruments used in Section 3.2. These are the population ratio, a dummy for ever having been a colony and their interaction. Results are given in column IV of Table 1, showing that the Hansen J test is now passed with a high level of confidence. Nevertheless, compared against column II, the results also suggest a trade-off between efficiency and transparency in instrument selection. While the use of multiple aggregate instruments is more transparent, it does not exploit the full information about bilateral aid flows contained in the zero stage. This may be one reason why the weak identification statistics are considerably lower in column IV versus column II of Table 1. In fact, as shown in column V of Table 1, even if only one aggregate instrument is employed, namely the log of the recipient's initial population, the strength of the instrument returns to similar values to those in column II and all coefficients are essentially unchanged.[3] Consequently, using a single instrument is likely to be more efficient but there are also potential information gains from employing a zero stage, especially in small aggregate samples such as those used in (static) cross-country regressions.

Thus, to strengthen the instrumentation approach, we return to the zero stage regressions. Aside from removal of suspect terms, additional concerns motivate further modifications. First, there are errors in the calculation of average Aid/GDP in all stages of RS08's regressions. The OECD-DAC aid dataset used for bilateral aid flows includes numerous missing values. While in some cases these genuinely refer to absent data, in most cases they represent unreported null

[3] This result further underlines the reliance of the RS08 instrumentation strategy on population size (also see Clemens and Bazzi, 2009).

values.[4] RS08 incorrectly treat these as missing. This is material because it distorts estimates for average bilateral aid flows over time. Consequently, it is necessary to re-estimate the bilateral aid variables and calculate period averages for Aid/GDP and aid per capita, setting missing entries to zero. This affects the dependent variable employed in the zero stage regression as well as the endogenous aid variable used in the IV estimations.

Second, in the RS08 strategy, recipient GDP occurs in the denominator of the dependent variable in the zero stage regressions. Following Kronmal (1993), inappropriate use of ratio variables may lead to substantial misinterpretation (or bias) in least squares regressions. This may arise if the denominator of the dependent variable is correlated with the RHS variables independently of the numerator of the dependent variable. In the present case, this could arise if donor decision rules do not target the Aid/GDP ratio, and/or if there is a direct association between recipient GDP levels and population size or past colonial experiences.

Third, it is apparent that individual donor countries exhibit distinct attitudes to giving foreign aid (Alesina and Dollar, 2000), which reflect cultural and historical factors. These time-invariant influences can be understood as fixed effects and may be included as RHS variables in the zero stage regression. Notably, and unlike the RS08 explanatory variables, these fixed effects may explain a part of the variation in aid allocations that is unrelated to purely strategic or political motives. As such, their inclusion may strengthen the overall validity and interpretation of the generated instrument.

To address these concerns, we modify the RS08 specification of the zero stage regression. In place of Aid/GDP, we use aid *per capita* (Aid/POP) as the dependent variable which accords closely with the explicit aid allocation rules used by donors, such as the World Bank (see Annex 1 of IDA15, 2008).[5] We drop the colonizer-specific variables (and interactions) and only include a dummy for whether a country was ever a colony (COLONY). Adding donor-specific fixed effects (DONOR), our zero stage regression emerges as follows:

$$\frac{Aid_{dr}}{POP_r} = \beta_0 + \beta_1 CURCOL_{dr} + \beta_2 COLONY_{dr} + \beta_3 COMLANG_{dr} + \beta_4 \log(POP_d/POP_r) + \beta_5 COLONY_{dr} \times \log(POP_d/POP_r) + \vartheta_d DONOR_d + \varepsilon_{dr} \quad (1)$$

where the subscripts *d* and *r* represent donors and recipients respectively; CURCOL indicates whether the recipient is a current colony of the donor.

---

[4] Confirmed in correspondence with the OECD DAC Secretariat.

[5] Note that in all subsequent regression stages the endogenous variable of interest remains Aid/GDP.

Table 3: Alternative zero stage regressions, 1970-2000

	(I) OLS	(II) OLS	(III) OLS	(IV) OLS	(V) Heckman
Colonial relationship (dummy)	1.65***	2.09***	11.95***	-0.55	-0.88
	(0.24)	(0.19)	(1.62)	(2.08)	(2.20)
Currently a colony (dummy)	-0.97*	0.63	9.88***	14.14	24.48
	(0.56)	(0.45)	(3.81)	(21.15)	(36.71)
Common language (dummy)	0.07*	0.09***	1.36***	1.30**	1.30*
	(0.04)	(0.03)	(0.27)	(0.60)	(0.67)
Ratio of (initial) log population	0.09***	0.05***	0.40***	0.32***	0.45***
	(0.01)	(0.00)	(0.04)	(0.06)	(0.08)
Ratio of log population x colony	0.62***	0.77***	7.16***	3.32***	3.36***
	(0.11)	(0.08)	(0.69)	(0.72)	(0.77)
Dependent variable	Aid/GDP	Aid/GDP	Aid p.c.	Aid p.c.	Aid p.c.
Treatment of “missing” aid values	Unknown	Zero	Zero	Zero	Zero
Metropole fixed effects & interactions	Yes	Yes	Yes	No	No
Donor fixed effects	No	No	No	Yes	Yes
Outcome and selection independence	-	-	-	-	9.56***
Number of obs.	3288	3286	3328	3328	3328
R-squared	0.42	0.31	0.26	0.21	-
F statistic	185.93	113.55	90.49	10.65	-

significance level: * 10%; ** 5%; *** 1%

Notes: column (I) replicates Rajan and Subramanian’s zero stage regression (2008, Table 4); columns (II) and (III) retain the same RHS specification, but alter the dependent variable (denoted in the table); column (IV) revises the specification, dropping metropole (colony-specific) fixed effects & interactions (coefficients not shown); column (V) implements a Heckman correction, based on the specification in column (IV); Heckman estimator uses full information maximum likelihood (FIML); the Heckman selection equation (not shown) includes all outcome covariates and a dummy for the number of colonial relationships experienced by the recipient; test for independence of outcome and selection equations refers to a Wald test that the correlation (*rho*) between the residuals in the two equations is equal to zero; intercept not shown; standard errors are robust to arbitrary heteroskedasticity and intra-group correlation between aid recipients (except for columns I to III where standard errors assume homoskedasticity in order to replicate Rajan and Subramanian, 2008).

Source: authors’ estimates.

Results from these modifications are given in Table 3. Column I replicates the RS08 specification (only selected coefficients shown); column II employs the revised dependent variable in which missing Aid/GDP values are set to zero. This change has a moderate impact and the pair-wise correlation between the fitted values from these two models is 0.83. Column III also retains the original RHS specification, but introduces aid per capita as the dependent variable (with missing aid values set to zero). All core coefficients retain the same sign and significance, but there is a minor fall in explanatory power, indicating there may have been some unwanted independent correlation between GDP in the dependent variable and the RHS variables. Column IV employs the new RHS specification, as per equation (1). Again, there is a small loss of explanatory power, but the population ratio and its interaction with the colony dummy remain highly significant. Also, the donor fixed effects (coefficients not shown in the table) vary in sign and many are significant. Overall, the RHS variables continue to explain a reasonable share of observed aid allocations.

The existence of zero-value aid inflows points to a final possible weakness. In principle, the decision by a donor to provide aid involves at least two distinct choices (Tarp *et al.,* 1999): (i) which recipients should receive aid; and (ii) how much to supply – i.e., the distribution of bilateral aid flows reflects an unobserved selection process. In the absence of an explicit model, one way to address potential bias from unobserved selection effects is to use Heckman's correction (Heckman, 1979). Column V of Table 3 employs a Heckman selection model (estimated by full information maximum likelihood) to the specification in column IV, where the existence of zero or non-zero aid flows is used as the binary selection variable. Despite these changes, the direction of the results and their interpretation are largely unchanged. However, we reject the hypothesis that there is no selection bias. We therefore retain the Heckman estimator employed in Column V as our preferred zero stage regression.

3.4. Improved specification

Before presenting the results of the aid-growth IV regressions using the improved instrument, it is appropriate to discuss additional areas where the RS08 approach can be strengthened. The first of these is the choice of covariates. Given the relatively small sample available in the aggregate regressions (78 countries), inclusion of redundant variables may lead to a loss of efficiency and/or contribute to undesirable multicollinearity. In the case of RS08, we note that the three macroeconomic initial conditions (inflation, money supply, and budget balance) as well as ethnic fractionalization are insignificant in RS08's cross-section outcome regressions for all periods. In addition, and as Wooldridge (2005) clarifies, inclusion of contemporaneous outcome variables – i.e., variables which may also be affected by the level of treatment – can invalidate the unconfoundedness assumption required for valid causal inference (Angrist and Pischke, 2008). This is pertinent as RS08's chosen specification includes two variables that capture average outcomes during the period of analysis – institutional quality and the number of forced changes in the top government elite, labelled "revolutions". Inclusion of these variables is puzzling in light of the literature which examines the effects of aid on growth through institutional performance. Controlling for such outcomes blocks potential channels through which aid may affect growth and thereby restricts the estimated coefficient on aid to a partial as opposed to a general effect. Such variables may also introduce unwanted reverse causality.

It is also helpful to consider the appropriate role of regional fixed effects. In RS08's specification, only East Asia and sub-Saharan Africa are included as regional dummy variables. This appears to be an *ex post* choice in the sense that prior to the 1980s there was no particular reason to identify these as "special". Including regional dummy variables helps absorb intra-regional correlations and captures omitted spatial fixed effects such as those arising from geography, shared historical experiences, and trade relationships. *A priori*, a more plausible approach is to include a fuller set of regional dummies. Finally, it is appropriate to include additional variables that reflect initial socioeconomic conditions such as education and health indicators, as well as additional geographic characteristics such as trading distances. These variables are frequently seen as important determinants of growth and may also proxy for initial conditions; as such, they may explain some of the variation in the expected growth returns to aid.

Consequently, we propose a revised covariates specification (denoted AJT). This involves dropping contemporaneous outcome covariates and redundant variables, adding an alternative set of regional dummies and including additional controls. These are selected following Sala-i-Martin *et al.* (2004) who undertake comprehensive Bayesian averaging of long-run growth estimates. We include variables identified by these authors that are among those with the highest posterior probability of inclusion and refer to initial conditions. To this, we add civil liberties in 1972 and distance to major ports. The first of these captures additional dimensions of initial institutional quality, including the ability of citizens to bring the government to account, which is often deemed relevant for aid effectiveness. Air distance is associated with export transaction costs, and ease of access to developed markets and has recently been identified by Moral-Benito (2009) as a robust correlate of growth.

3.5. Alternative estimators

Another area that can be strengthened concerns the choice of IV estimator. In light of the expected complexity of the growth process as well as the different properties of alternative estimators, it is valid to investigate whether or not empirical results hold across different estimators. While RS08 employ a 2SLS estimator, this is not the only option. Other possible estimators, which offer moderate differences, include LIML (limited information maximum likelihood), Fuller's modified LIML (with alpha = 1), and a continuously updated GMM estimator (GMM-CU).

In the program evaluation literature, the "doubly robust" estimators of Robins and Rotznitzky (1995) are attractive. Imbens and Wooldridge describe these estimators as "best practice" (2009: 25). Various doubly robust estimators have been proposed (see Imbens, 2004); however, none of these can be applied straightforwardly to the current aid-growth problem. They assume a binary treatment/control framework in which receipt of the treatment is conditionally independent of potential outcomes. Indeed, Imbens and Wooldridge (2009) note that, compared with the binary case, much less is known about settings with continuous treatment variables even though such settings are common in practice. Nevertheless, some simple extensions enable doubly robust estimators to be extended to the instrumental variables context (see Arndt *et al.*, 2010). Principally, this involves dichotomization of the (generated) instrument, whilst maintaining the endogenous variable in continuous form. This is our preferred estimator, and we denote this IV-IPWLS (IV inverse probability weighted least squares).

4. Empirical results

Section 3 motivated and proposed three main improvements to RS08's empirical approach. These refer to their instrument, their specification, and their chosen estimator. Tables 4 and 5 present results for different combinations of these modifications. We focus on the 1970-2000 period, allowing the effects of aid to be considered over a generation of elapsed time. The shorter periods (1980-2000 and 1990-2000) may not allow sufficient time for the aid-growth relationship to emerge. With respect to 1960-2000, many countries had not attained independence by 1960, particularly those in Africa. Further, even though the majority of French colonies achieved independence in 1960, the shift to independent administration was very gradual in most cases. In contrast, by 1970 the large majority of developing countries had achieved independence and had operated independently for at least a few years, with Portuguese colonies being the prominent exception.

Column I of Table 4 reports regression results using the RS08 specification, the new preferred instrument (based on the zero stage regression in column V of Table 3), and the IV-LIML estimator. Column II introduces the doubly robust IV-IPWLS estimator, while columns III and IV replicate columns I and II with our new specification. Columns V and VI continue with the same model, but respectively use the GMM-CU and Fuller estimators. Note that in all columns the improved instrument is employed.

DOI: 10.2202/1948-1837.1121

Table 4: Modified IV regressions, 1970-2000

	(I) IV-LIML	(II) IV-IPWLS	(III) IV-LIML	(IV) IV-IPWLS	(V) GMM-CU	(VI) Fuller
Aid / GDP	0.22*	0.21*	0.25**	0.13***	0.25**	0.24**
	(0.12)	(0.13)	(0.12)	(0.05)	(0.12)	(0.12)
Initial per cap. GDP	-1.34***	-1.92***	-1.03***	-1.33***	-1.03***	-1.05***
	(0.40)	(0.39)	(0.38)	(0.27)	(0.38)	(0.37)
Initial level of policy	2.14***	2.58***	2.12***	2.44***	2.12***	2.12***
	(0.60)	(0.62)	(0.54)	(0.46)	(0.54)	(0.53)
Initial life expectancy	0.09**	0.05	0.04	0.03	0.04	0.03
	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
Geography	0.63**	0.48**	0.29	0.25	0.29	0.29
	(0.25)	(0.24)	(0.26)	(0.21)	(0.26)	(0.25)
Coastal pop. density in 1965			0.00**	0.00***	0.00**	0.00**
			(0.00)	(0.00)	(0.00)	(0.00)
Primary schooling in 1960			2.58**	2.26**	2.58**	2.56**
			(1.15)	(0.88)	(1.15)	(1.13)
Malaria risk in 1966			-1.50*	-1.06*	-1.50*	-1.49*
			(0.85)	(0.58)	(0.85)	(0.83)
Invest. goods price, 1960-64			-0.01	-0.01	-0.01	-0.01
			(0.00)	(0.00)	(0.00)	(0.00)
Civil liberties in 1972			-1.28*	-0.98*	-1.28*	-1.24*
			(0.70)	(0.50)	(0.70)	(0.68)
Air distance (log)			0.09	-0.03	0.09	0.09
			(0.38)	(0.33)	(0.38)	(0.38)
Specification	RS08	RS08	AJT	AJT	AJT	AJT
Scale of excluded instrument	Continuous	Binary	Continuous	Binary	Continuous	Continuous
Regional dummies	SSA, EA	SSA, EA	SSA, A, LA	SSA, A, LA	SSA, A, LA	SSA, A, LA
N	78	78	78	78	78	78
R-squared	0.57	0.70	0.59	0.77	0.59	0.60
Kleibergen-Paap Wald F stat.	29.48	24.42	17.28	39.78	17.28	17.28
Stock-Wright LM S stat.	4.33	3.53	5.77	6.49	5.77	5.77
(probability)	0.037	0.060	0.016	0.011	0.016	0.016

significance level: * 10%; ** 5%; *** 1%

Notes: the endogenous variable is Aid/GDP, re-estimated from OECD-DAC (2008) data treating possible missing values as zeroes; in columns (I) and (II) the specification follows Rajan and Subramanian (2008) (only selected covariates shown); all remaining columns use a modified specification, removing contemporaneous and redundant covariates and adding additional initial conditions; chosen estimator is in the column title; initial policy refers to the Sachs-Warner trade policy index; geography refers to the average of the number of frost days and tropical land area; intercept not shown; standard errors, given in parentheses, are robust to arbitrary heteroskedasticity; dependent variable is the average real growth rate.

Source: authors' estimates.

With respect to the implementation of the IV-IPWLS estimator (columns III and IV), a binary instrument is required. This is derived by taking the fitted instrument from RS08's zero stage regression, sorting countries in ascending order (from lowest to highest predicted aid shares), and then selecting the first 30 for the "control" and the rest for the "treatment" group. The motivation for this choice is to identify a sub-sample of countries with the smallest possible average value for predicted aid inflows while still maintaining statistical viability. Thus, in practice, the control group approximately corresponds to all countries falling below the 40th percentile. Besides permitting application of doubly robust techniques, dichotomization of the instrument represents a useful robustness check. If results arising from the binary instrument were not comparable with its continuous counterpart, this might indicate that the latter findings were driven by peculiarities in the distribution of the instrument. It also relaxes the assumption of a constant linear relationship between aid and growth, instead placing emphasis on the average difference between treatment and control groups regardless of the shape of growth's response to aid. Consequently, possible non-linear effects due to diminishing returns to aid are addressed by this dichotomization. Finally, as the instrument is derived from a zero stage regression, dichotomization provides a check against measurement error or misspecification in the zero stage.

Turning to results, the range of test statistics reported in Table 4 indicates that the new instrument continues to perform strongly across different specifications and estimators. Under-identification tests (not shown), which can be interpreted as testing the null hypothesis of a zero correlation between the instruments and the endogenous regressors, are all rejected. The weak identification test (the Kleibergen-Paap Wald F statistic, which uses a finite-sample adjustment of the standard F-statistic to assess the strength of the partial correlation between the excluded instruments and the endogenous variables in first-stage regressions) not only exceeds critical values in all cases but is comparable to the levels achieved using RS08's original approach (Table 1, column II). Perhaps more importantly, the Stock-Wright S statistic, which is based on the reduced form regression and is robust to the presence of weak instruments (see Baum *et al.*, 2007), finds a significant (partial) correlation between the instrument and dependent variable in all cases.

Moving across the columns of Table 4, we note that the treatment effect – i.e., the coefficient on the endogenous aid variable – is consistently positive, significant, and in a domain that is consistent with the RS08 prior (Section 2.4). The main effect of using the new and strengthened instrument (column I) is that the treatment effect estimate edges upwards (from 0.l0 to 0.22). The doubly-robust estimator leaves this result almost unchanged, but enhances the overall explanatory power of the model. According to the Kleibergen-Paap Wald F statistic, switching to the modified specification (column III onwards) slightly reduces the strength of the instrument in the IV-LIML first stage. However, by placing greater emphasis on the most informative observations, the strength of the instrument is considerably improved for the IV-IPWLS estimator. Finally, the

DOI: 10.2202/1948-1837.1121

alternative IV estimators (columns V and VI) are virtually identical to the results of column III.

The preferred estimate presented in Column IV of Table 4 represents a new estimator, a new specification, and a new instrumentation strategy.[6] In contrast to RS08, we find robust statistically significant empirical support for a positive aid-growth relationship for the 1970-2000 period. Robustness and sensitivity tests, documented in Arndt *et al.* (2010), include application of a flexible doubly robust estimator, exclusion of (possibly) influential observations, as well as implementation of sample restrictions and additional covariates. They show that, despite some sensitivity to different choices, there is no firm basis on which to reject our main results.

Table 5: Summary of results from model modifications, 1970-2000

Instrument	**Specification**	**Estimator**	
		2SLS / IV-LIML	IV-IPWLS
RS08	RS08	0.10	0.15*
	AJT	0.10	0.10**
AJT	RS08	0.22*	0.21*
	AJT	0.25**	0.13***

significance level: * 10%; ** 5%; *** 1%

Notes: AJT refers to our preferred instrument (Table 3, Column V) or specification (Table 4 Column III). Cells show the estimated coefficient on Aid/GDP from IV regressions involving different combinations of specifications, instruments, and estimators; column (I) provides estimates from standard IV estimators (2SLS or IV-LIML); column (II) employs the IV-IPWLS estimator, as described in the text; standard errors on which statistical inference is based are robust to arbitrary heteroskedasticity; dependent variable is the average real growth rate.

Source: authors' estimates.

To get a better sense of the individual and joint impact of alternative combinations of our three main modifications, Table 5 provides a summary of the various models. Each cell reports the estimated impact of aid on growth over the 1970-2000 period for a given combination of modifications. Reinforcing the previous results, we find that the RS08 result is not robust. With the exception of employing the AJT specification alone, all other modifications (either individually or jointly) yield a significant aid-growth relationship. Moreover, the probability of falsely rejecting the null hypothesis falls with the number of modifications

[6] Note that adding the three initial macroeconomic conditions employed by RS08, which had been excluded for redundancy, to the models estimated in columns III and IV of Table 4 leaves all results essentially unchanged; moreover, these three variables continue to be redundant.

employed. Thus, when all three modifications are incorporated, the estimated coefficient on Aid/GDP of 0.13 becomes significant at the 1% level.

Given our preferred approach employs a single generated instrument, it remains to be established whether the underlying instrumentation strategy remains valid in the context of the new set of conditioning variables. Thus, maintaining the improved specification and IV-LIML estimator, we replace the generated instrument with different sets of aggregate instruments as per Sections 3.2 and 3.3. These results are reported in columns VI to VIII of Table 1, which respectively employ a full set of aggregate instruments (used in the RS08 zero stage), a preferred sub-set of three "least suspect" instruments, and initial log population size only. The most important result is that the Hansen J statistic is considerably strengthened when three as opposed to eight aggregate instruments are used, passing all conventional test thresholds. Moreover, in the last two columns, the coefficient on Aid/GDP remains positive and significant, once again supporting our principal results.

We now return to the other periods analyzed in RS08 (1960-2000, 1980-2000, and 1990-2000), using our preferred specification and instrumentation strategy as well as both the IV-LIML and IV-IPWLS estimators. Results for each alternative period are presented in Table 6. For the 1960-2000 period (columns I and II) both the point estimate and variance of the estimated treatment effect are squarely in the domain found in Table 4 (columns III and IV) for the 1970-2000 period. The long-run impact of foreign aid comes across as well established. With respect to the shorter run effects of aid, given in columns III to VI of Table 6, we cannot reject the hypothesis that the treatment effect is zero. This is confirmed by the (very weak) relation in the reduced form given by the Stock-Wright S statistic. The plausible range for the treatment effect is much wider for these periods, reflected by larger standard errors on the treatment effect. This is most apparent for 1990-2000 where the standard error on the FDR treatment effect estimate is almost five times larger than that for the 1960-2000 period. As suggested in Section 2, a meaningful and robust average short-run effect of aid on growth may be very difficult to discern from the available empirical data.

At this point, it is helpful to reflect on what exactly the estimated regression parameters represent. As has been established in the literature, instrumental variables estimators typically cannot be interpreted as average treatment effects. This is only appropriate under strong additional assumptions – in this case, homogeneity across countries in their response to aid. Rather, what is actually recovered depends on the instruments chosen as well as the extent of heterogeneity in responses to changes in these instruments. In the case of a single binary instrument and endogenous response variable, IV estimators often recover a local average treatment effect (LATE), defined as the average treatment effect for the sub-population that switches from the control to the treatment group on account of the switch in the instrument. This can be understood as a weighted average of the marginal treatment effects for the sub-population that is responsive to the instrument (Heckman, 2001).

Table 6: Modified regressions, alternative periods

	1960-2000		1980-2000		1990-2000	
	(I)	(II)	(III)	(IV)	(V)	(VI)
	IV-LIML	IV-IPWLS	IV-LIML	IV-IPWLS	IV-LIML	IV-IPWLS
Aid / GDP	0.16*	0.09**	0.02	0.05	-0.11	0.11
	(0.08)	(0.04)	(0.14)	(0.10)	(0.19)	(0.14)
Initial per cap. GDP	-0.67**	-0.83***	-1.41***	-1.36***	-0.72	-0.12
	(0.31)	(0.24)	(0.42)	(0.35)	(0.74)	(0.57)
Initial level of policy	1.88***	2.33***	2.10***	1.51	0.65	0.99*
	(0.45)	(0.42)	(0.71)	(1.01)	(0.53)	(0.52)
Initial life expectancy	0.01	0.02	0.06	0.10**	0.12	0.13**
	(0.02)	(0.03)	(0.04)	(0.04)	(0.07)	(0.06)
Geography	0.26	0.21	0.48**	0.33	0.13	0.23
	(0.19)	(0.17)	(0.21)	(0.23)	(0.41)	(0.38)
Coastal pop. density in 1965	0.00***	0.00***	0.00*	0.00	0.00	0.01***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Primary schooling in 1960	2.76***	2.13***	1.45	1.15	-1.38	-2.08
	(0.91)	(0.74)	(1.00)	(1.04)	(1.93)	(1.68)
Malaria risk in 1966	-1.20**	-1.03**	-0.99	-1.09	-2.36**	-2.48**
	(0.57)	(0.40)	(0.79)	(0.77)	(0.96)	(1.02)
Invest. goods price, 1960-64	-0.01**	-0.01**	0.00	-0.00	-0.01	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Civil liberties in 1972	-0.33	-0.40	-0.34	-0.19	0.67	0.74
	(0.26)	(0.27)	(0.31)	(0.35)	(0.64)	(0.62)
Air distance (log)	0.31	0.28	1.19*	1.27	2.11**	1.78**
	(0.42)	(0.39)	(0.69)	(0.82)	(0.90)	(0.75)
Scale of excluded instrument	Binary	Continuous	Binary	Continuous	Binary	Continuous
N	74	74	75	75	70	69
R-squared	0.71	0.84	0.64	0.60	0.52	0.51
Kleibergen-Paap Wald F stat.	15.82	42.80	19.31	22.01	17.58	20.04
Stock-Wright LM S stat.	4.30	4.30	0.03	0.20	0.30	0.63
(probability)	0.038	0.038	0.869	0.654	0.584	0.428

Notes: the endogenous variable is Aid/GDP, instrumented as per the models in Table 3 and re-estimated from the relevant bilateral data and time period; initial policy refers to the Sachs-Warner trade policy index; geography refers to the average of the number of frost days and tropical land area; covariate values also vary according to the period chosen, however this is not the case with respect to the variables from Sala-i-Martin et al. (2004) as alternative initial values are not available; intercept and regional dummies not shown; standard errors, given in parentheses, are robust to arbitrary heteroskedasticity; dependent variable is the average real growth rate.

Source: authors' estimates.

In the present case, the complex mapping from multiple instruments to a continuous endogenous variable significantly complicates a LATE-type interpretation of the estimates. Furthermore, given the small sample sizes involved, distinguishing between key sub-populations and important observations is exceedingly difficult operationally. Nevertheless, in light of the instruments used, a possible interpretation is that the estimates reflect the growth response to aid for countries whose total aid inflows have been most influenced by differences in relative population sizes or ever having been a colony. These drivers typically would be associated with political rather than altruistic motivations. If politically motivated aid is less effective, this may bias downward the estimated aid coefficient. Finally, one notes that moving from the continuous instrument to the binary instrument (used in the doubly robust estimations) has a relatively small effect on the estimated parameter of interest. This is consistent with the observation made in Arndt *et al.* (2010) that the results are not dependent on the specific threshold chosen to define the binary instrument.

Lastly, we consider how this evidence stacks up against theory. In all the regression specifications reported so far, tests have been made against a null hypothesis of a zero relationship between the explanatory and dependent variables. Nevertheless, it is straightforward to calculate the t-statistic for testing whether the estimates differ from the theoretical prior that the point estimate for the long-run effect of foreign aid on growth is around 0.1 (see Section 2.4). With the exception of the OLS regression reported in column I of Table 1, all point estimates for the aid growth relationship across all regressions reported in Tables 1, 4, 5, and 6 are not significantly different from 0.1.[7] To put it differently, there is no basis on which to reject the theoretical prior that aid has a positive long-run effect on growth.

5. Conclusion

To conclude, we respond to the question posed in the title to this paper: has the aid, growth, and development literature gone full circle? Our answer is “no”. While in the most recent literature the pendulum has swung to deep skepticism concerning the ability of aid to contribute to economic growth, a series of important points of consensus have emerged. First, methodological advances in the program evaluation literature have improved the profession’s capacity to identify causal effects in economic phenomena. These advances are beginning to be applied at the more aggregate level, as pursued here. Second, methodological advances also highlight the serious challenges that must be surmounted in order to

[7] For these regressions, the null hypothesis that the coefficient of the estimated aid-growth parameter is equal to 0.1 is, in fact, never close to being rejected, including in RS08’s original IV specification (Table 1, Column II).

derive robust causal conclusions from non-experimental data. In many important areas of inquiry, long-standing debates with respect to causal impacts persist despite improved methods and improved data availability. Third, the formation of reasonable expectations about the likely returns to foreign assistance has been greatly facilitated by the application of growth theory. Finally, there is increasing recognition that many of the key interventions pursued by foreign aid will only result in positive growth outcomes over long time horizons.

In line with Temple (2010), we started by replicating RS08. Based on a detailed analysis of their approach, we subsequently developed a better instrumentation strategy, an improved specification, and a preferred estimator. The improved specification contains a fuller set of regional fixed effects and indicators of initial human capital and geographic conditions. These were drawn from theory and previous research. They included primary schooling, coastal population density, and malaria risks. Consistent with best practice in the program evaluation literature, we excluded covariates, such as revolutions and institutional performance, which represent potential channels through which aid affects growth. With respect to the zero stage instrumentation, we (i) excluded suspect variables; (ii) corrected errors in the implementation of the RS08 instrumentation strategy; (iii) employed aid per capita in place of Aid/GDP to preclude spurious correlation with the chosen instruments; (iv) introduced donor-specific fixed effects; and (v) accounted for selection bias through a Heckman correction. Finally, we deployed robust regression estimators which adjust for heterogeneity across countries. This involved introducing a new doubly robust estimator that can be used in instrumental variable contexts. A variety of robustness and validity checks, including of the underlying instrument, provide support to our approach.

Overall, we believe our study represents the most carefully developed empirical strategy employed in the aid-growth literature to date. Our results provide solid support for the view that the effect of aid on growth is positive in both the 1970-2000 and 1960-2000 periods. The preferred doubly robust estimator places the point estimate of the long run elasticity of growth with respect to the share of aid in recipient GDP at 0.13 (IV-IPWLS). This is below the unweighted point estimate of 0.25 (IV-LIML). In sum, our findings suggest that an inflow on the order of 10 percent of GDP spurs the per capita growth rate by more than one percentage point per annum in the long run. These estimates are consistent with the view that foreign aid stimulates aggregate investment and may also contribute to productivity growth, despite some fraction of aid being allocated to consumption. The 95% confidence interval around our estimates lies in the strictly positive domain and contains the prior, suggested in RS08 that the long-run elasticity of growth to foreign aid should be around 0.1. In the shorter term, our analysis indicates that the impact of aid is difficult to discern. Nevertheless, when the longer run macro evidence is combined with the evidence at the micro- and meso-levels, a consistent case for aid effectiveness emerges. There is no micro-macro paradox.

We find ourselves in a similar position to Winters (2004) in his review of the implications of trade liberalization for growth. While he concludes that trade liberalization stimulates growth over the long run and on average, he adds that "For a variety of reasons, the level of proof remains a little less than one might wish but the preponderance of evidence certainly favors that conclusion" (2004: F18). Similarly, we conclude that the bleak pessimism of much of the recent aid-growth literature is unjustified and the associated policy implications drawn from this literature are often inappropriate and unhelpful. Aid has been and remains an important tool for enhancing the development prospects of poor nations.

Finally, the complex and idiosyncratic process of managing aid to spark and sustain growth is subject to considerable learning. Nearly all participants in the aid-growth debate, not least these authors, recognize the potential for aid to do better, particularly in fostering productivity growth. Abolishing foreign aid, or drastically cutting it back, would be a mistake and is not warranted by any reasonable interpretation of the evidence. The challenge is to improve foreign assistance effectiveness so that living standards in poor countries are substantially advanced over the next three decades.

References

Acemoglu, D. and S. Johnson. 2007. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy* 115 (6): 925-985.

Alesina, A., and D. Dollar. 2000. "Who Gives Foreign Aid to Whom and Why?" *Journal of Economic Growth* 5 (1): 33-63.

Angrist, J.D., and J.-S. Pischke. 2010. "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics." *Journal of Economic Perspectives* 24(2): 3-30.

Angrist, J.D., and J.-S. Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Arndt, C., S. Jones and F. Tarp. 2010. "Aid, Growth, and Development: Have We Come Full Circle?". UNU-WIDER Working Paper 96.

Ashraf, Q.H., A. Lester, and D.N. Weil. 2008. "When Does Improving Health Raise GDP?". *NBER Working Paper* 14449, National Bureau of Economic Research, Inc.

Banerjee, A.V. and E. Duflo. 2009. "The Experimental Approach to Development Economics", *Annual Review of Economics* 1: 151-178.

DOI: 10.2202/1948-1837.1121

Baum, C.F., M.E. Schaffer, and S. Stillman. 2007. "Enhanced Routines for Instrumental Variables/generalized Method of Moments Estimation and Testing." *Stata Journal* 7 (4): 465-506.

Bun, M., and F. Windmeijer. 2010. "The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models." *Econometrics Journal* 13(1): 95-126.

Burnside, C., and D. Dollar. 2000. "Aid, Policies, and Growth." *American Economic Review* 90 (4): 847-868.

Card, D. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69 (5): 1127-60.

Clemens, M., and S. Bazzi. 2009. "Blunt Instruments: On Establishing the Causes of Economic Growth." *Center for Global Development Working Paper* 171.

Cohen, D. and M. Soto. 2007. "Growth and Human Capital: Good Data, Good Results." *Journal of Economic Growth* 12: 51-76.

Collier, P., and A. Hoeffler. 2004. "Aid, Policy and Growth in Post-conflict Societies." *European Economic Review* 48 (5): 1125-1145.

Dalgaard, C.-J,. and L. Erickson. 2009. "Reasonable Expectations and the First Millennium Development Goal: How Much Can Aid Achieve?" *World Development* 37 (7): 1170-1181.

Dalgaard, C.-J., H. Hansen, and F. Tarp. 2004: "On The Empirics of Foreign Aid and Growth." *Economic Journal* 114 (496): F191-F216.

Djankov, S., J. Montalvo, and M. Reynal-Querol. 2008. "The Curse of Aid." *Journal of Economic Growth* 13 (3): 169-194.

Easterly, W. 1999. "The Ghost of Financing Gap: Testing the Growth Model Used in the International Financial Institutions." *Journal of Development Economics* 60 (2): 423-438.

Easterly, W., R. Levine, and D. Roodman. 2004. "Aid, Policies, and Growth: Comment." *American Economic Review* 94 (3): 774-780.

Hansen, H., and F. Tarp. 2000. "Aid Effectiveness Disputed." *Journal of International Development* 12 (3): 375-398.

Hansen, H., and F. Tarp. 2001. "Aid and Growth Regressions." *Journal of Development Economics* 64 (2): 547-570.

Hauk, W., and R. Wacziarg, R. 2009. "A Monte Carlo Study of Growth Regressions." *Journal of Economic Growth* 14 (2): 103-147.

Heckman, J.J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (1), 153-61.

Heckman, J.J. 2001. "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture." *Journal of Political Economy* 109(4): 673-748

IDA15. 2008. "Additions to IDA Resources: Fifteenth Replenishment." Technical report, International Development Association.

Imbens, G.W. 2004. “Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review.” *The Review of Economics and Statistics* 86 (1): 4-29.

Imbens, G.W., and J.M. Wooldridge. 2009. “Recent Developments in the Econometrics of Program Evaluation.” *Journal of Economic Literature* 47 (1): 5-86.

Kronmal, R.A. 1993. “Spurious Correlation and the Fallacy of the Ratio Standard Revisited.” *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 156 (3): 379-392.

La Porta, R., F.L. de Silanes, and A. Shleifer, A. 2008. “The Economic Consequences of Legal Origins.” *Journal of Economic Literature* 46 (2): 285-332.

Mankiw, N.G., D. Romer, and D.N. Weil. 1992. “A Contribution to the Empirics of Economic Growth.” *The Quarterly Journal of Economics* 107 (2): 407-37.

Masud, N. and B. Yontcheva. 2005. “Does Foreign Aid Reduce Poverty? Empirical Evidence from Nongovernmental and Bilateral Aid.” *International Monetary Fund Working Paper* 05/100, International Monetary Fund.

Mishra, P., and D.L. Newhouse. 2007. “Health Aid and Infant Mortality.” *International Monetary Fund Working Paper* 07/100, International Monetary Fund.

Moral-Benito, E. 2009. "Determinants of Economic Growth: A Bayesian Panel Data Approach." *Policy Research Working Paper Series* 4830, The World Bank.

Mosley, P. 1987. *Overseas Aid: Its Defence and Reform.* Brighton: Wheatshead Books.

Moyo, D. 2009. *Dead Aid: Why Aid Is Not Working and How There Is a Better Way for Africa*. London: Allen Lane.

Nunn, N. 2008. “The Long-Term Effects of Africa's Slave Trades.” *The Quarterly Journal of Economics* 123 (1): 139-176.

Papanek, G.F. 1972. “The Effect of Aid and Other Resource Transfers on Savings and Growth in Less Developed Countries.” *Economic Journal* 82 (327): 935-50.

Papanek, G.F. 1973. “Aid, Foreign Private Investment, Savings, and Growth in Less Developed Countries.” *Journal of Political Economy* 81 (1): 120-30.

Rajan, R., and A. Subramanian. 2007. “Does Aid Affect Governance?” *AEA Papers and Proceedings,* 97(2): 322-327.

Rajan, R.G., and A. Subramanian. 2008. “Aid and Growth: What Does the Cross-Country Evidence Really Show?” *The Review of Economics and Statistics* 90 (4): 643-665.

DOI: 10.2202/1948-1837.1121

Robins, J.M., and A. Rotnitzky, A. 1995. “Semiparametric Efficiency in Multivariate Regression Models with Missing Data.” *Journal of the American Statistical Association* 90 (429): 122--129.

Roodman, D. 2007. “The Anarchy of Numbers: Aid, Development, and Cross-Country Empirics.” *World Bank Economic Review* 21 (2): 255-277.

Roodman, D., and J. Morduch. 2009. “The Impact of Microcredit on the Poor in Bangladesh: Revisiting the Evidence.” *Center for Global Development Working Paper* 174

Roodman, D. 2009. “A Note on the Theme of Too Many Instruments.” *Oxford Bulletin of Economics and Statistics* 71 (1): 135-158.

Sachs, J. 2005. *The End of Poverty: Economic Possibilities for Our Time*, Penguin Press: New York.

Sala-i-Martin, X., G. Doppelhofer, and R.I. Miller. 2004. “Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach.” *American Economic Review* 94 (4): 813-835.

Tarp, F., C.F. Bach, H. Hansen, and S. Baunsgaard. 1999. “Danish Aid Policy: Theory and Empirical Evidence”. Chapter 9 (pp. 149-69) in Kanhaya Gupta (ed.) *Foreign Aid: New Perspectives*. Norwell, MA: Kluwer Academic Publishers.

Temple, J.R.W. 2010. “Aid and Conditionality.” In Dani Rodrik and Mark Rosenzweig (eds.), *Handbook of Development Economics.* 5: 4415-4523.

Winters, L.A. 2004. “Trade Liberalisation and Economic Performance: An Overview.” *Economic Journal* 114 (493): F4-F21.

Wooldridge, J.M. 2005. “Violating Ignorability of Treatment By Controlling For Too Many Factors.” *Econometric Theory* 21 (05): 1026-1028.

World Bank. 2008. *Annual Review of Development Effectiveness 2008: Shared Global Challenges*, Independent Evaluation Group, available at: http://go.worldbank.org/U2T30HQKG0