www.elsevier.com/locate/worlddev

World Development Vol. 69, pp. 6–18, 2015
0305-750X/© 2014 UNU-Wider. Published by Elsevier Ltd.

http://dx.doi.org/10.1016/j.worlddev.2013.12.016

Assessing Foreign Aid's Long-Run Contribution to Growth and Development☆

CHANNING ARNDT and SAM JONES
University of Copenhagen, Denmark

and

FINN TARP*
UNU-WIDER, Helsinki, Finland

Summary. — This paper confirms recent evidence of a positive impact of aid on growth and widens the scope of evaluation to a range of outcomes including proximate sources of growth (e.g., physical and human capital), indicators of social welfare (e.g., poverty and infant mortality), and measures of economic transformation (e.g., share of agriculture and industry in value added). Focusing on long-run cumulative effects of aid in developing countries, and taking due account of potential endogeneity, a coherent and favorable pattern of results emerges. Aid has over the past 40 years stimulated growth, promoted structural change, improved social indicators, and reduced poverty.

Key words — growth, foreign aid, aid effectiveness, simultaneous equations

1. INTRODUCTION

Significant volumes of foreign aid have been channeled to developing countries for more than four decades. Not surprisingly, a large literature considers aid effectiveness particularly from the perspective of the impact of aid on aggregate economic growth. While Rajan and Subramanian (2008) find no systematic evidence that aid has contributed to economic growth, the weight of evidence is shifting to a positive contribution of aid to growth. Arndt, Jones, and Tarp (2010a) employ the same approach and raw data as Rajan and Subramanian (2008). After strengthening the prediction of supply side variation in aid, including correction for a misinterpretation of OECD/DAC bilateral aid data, they find a positive long run effect of aid on growth which lies in the domain predicted by neo-classical growth theory (e.g., Solow, 1956). Clemens, Radelet, Bhavnani, and Bazzi (2012) revisit the dynamic panel evidence, focusing on aid that is expected to have an "early impact" on growth—e.g., via infrastructure development. The authors conclude that: "[such] aid inflows are systematically associated with modest, positive subsequent growth in cross-country panel data". (p. 23). More recently, Frot and Perrotta (2012) suggest a new instrument for aid identified by the timing of the initiation of bilateral aid relationships. They come to a similar conclusion that foreign aid is associated with a moderate growth bonus. Finally, time series evidence for a range of African countries (Juselius, Møller, & Tarp, 2013) support a view that aid has played a positive aggregate developmental role in most instances; and meta-analysis of the aid–growth relation leads to a similar conclusion (Mekasha & Tarp, 2013). This macro-level evidence comes on top of meso- and micro-level evidence that has long been viewed as broadly positive (Mosley, 1987; see also Mishra & Newhouse, 2009; Riddell, 2007; Temple, 2010). However, despite increasing evidence that meso-level outcomes can add up to substantial macroeconomic effects (Cohen & Soto, 2007), these micro- and meso-level findings have not been deployed to argue that aid is effective on aggregate (one exception is Sachs, 2006).

In this article we aim to provide a broader assessment of aid effectiveness. While a focus on the effect of aid on macroeconomic growth is necessary, it is not sufficient. A growing literature considers the contribution of aid in specific social sectors, such as education. Indeed, many outcomes are valued independently of their contribution to growth. Access to "merit goods," such as basic health care and primary education, are viewed as essential human rights and fundamental to the development process. Accordingly, these outcomes should be included when considering the accomplishments of aid.

* The authors are grateful to three anonymous referees for their comments which helped strengthen the paper significantly. We also acknowledge substantive questions and suggestions from Alan Winters as well as very helpful comments and advice from Tony Addison, Ernest Aryeetey, David Bevan, Arne Bigsten, Tove Degnbol, Augustin Fosu, Henrik Hansen, Jane Harrigan, Rolph van der Hoeven, Yongfu Huang, Katarina Juselius, Richard Manning, Tseday Jemaneh Mekasha, Oliver Morrissey, Paul Mosley, Lasse Møller, John Page, Jean-Philippe Platteau, Lant Pritchett, Gustav Ranis, Roger Riddell, David Roodman, Jeff Round, and Adrian Wood. The same goes for participants to the UNU-WIDER Conference on Foreign Aid: Research and Communication held in Helsinki, September 30–October 1, 2011, the Nordic Conference in Development Economics (NCDE) held in Copenhagen, June 20–21, 2011, and the Joint AERC & UNU-WIDER Conference on the Macroeconomics of Aid held in Nairobi, December 01–02, 2011. The usual caveats apply.

☆ This study paper has been prepared within the UNU-WIDER project "Foreign Aid: Research and Communication (ReCom)" directed by Tony Addison and Finn Tarp and supported by specific contributions from the governments of Denmark (Ministry of Foreign Affairs, Danida) and Sweden (Swedish International Development Cooperation Agency – Sida).

This broader assessment provides enhanced insight into the aid–growth relationship in three further ways. First, we extend the analysis of Arndt *et al.* (2010a) by adding 7 years of more recent data to the series. Second, we investigate the consistency of the growth evidence with changes in other domains, particularly proximate determinants of growth, thus providing a coherence test for the aid–growth relationship. If no robust evidence of a relationship can be found between aid and important growth determinants such as investment and human capital, then the impact of foreign aid on growth becomes much harder to explain. Third, consideration of a wide range of alternative outcomes also provides a means to validate the robustness of the methods employed to address the likely endogeneity of aid.

As with many empirical questions in the economics literature, studying aid effectiveness is beset by difficulties in determining causality. In order to address these challenges, we outline a general framework that clarifies aid's potential role in contributing both to intermediate outcomes (e.g., human capital accumulation) and final outcomes (e.g., growth). The model also indicates how these effects can be identified from observational data and precisely what feasible empirical estimates will capture. The empirical analysis is then pursued in four steps: we (i) calculate reduced form estimates of the impact of aid on a range of final economic outcomes (growth, poverty, inequality, and structural change); (ii) apply the same reduced form approach to a set of intermediate economic outcomes (such as investment, consumption, and tax take) as well as social outcomes (such as health and education); (iii) run a set of sensitivity and falsification tests; and (iv) interpret the economic magnitude of the findings as well as their consistency with previous literature. In presenting a broader assessment, this analysis responds, at least in part, to the challenge set forth by Bourguignon and Sundberg (2007) to unpack the causal chain from aid to final outcomes.

We find no evidence that nearly 40 years of development assistance has had an overall detrimental effect on development outcomes. Rather, a coherent and favorable picture emerges. Aid has promoted structural change, reduced poverty, and stimulated growth. Aid also has supported proximate growth determinants, in particular by building human and physical capital. This does not mean that aid works well at all times and in all places. Also, the impact of aid is no doubt heterogeneous. Nevertheless, these findings are consistent with significant strands of the existing literature and add further weight to the conclusion that, while perhaps less potent than initially hoped and certainly not a panacea, aid has registered significant accomplishments in helping to achieve development goals.

2. METHODOLOGY

(a) *Analytical framework*

A variety of approaches have been developed to address questions of economic causality. These issues are at the core of assessing the impact of aid and are reflected in the ongoing debate concerning the suitability of the various instruments for aid that have been employed in the literature (see Clemens & Bazzi, 2009). A useful starting point for thinking about these issues is a graphical depiction of the principal (generic) impact channels assumed to be at play. A simple version of this is provided in Figure 1, which is inspired by the directed acyclic graphs (DAGs) that are central to the Structural Causal Model (SCM) approach to analyzing causality due to Pearl

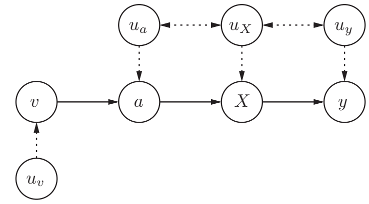

Figure 1. *General causal diagram summarizing the linkages between aid and final outcomes. Notes: This figure is a simplified causal directed acyclic graph (DAG) of the relationship between aid (a) and aggregate outcomes (y), via intermediate outcomes (X); v is a single exogenous determinant of aid; u terms are unobserved, possibly errors; solid lines represent directed relationships between observed variables; broken lines represent directed relations due to unobserved variables (errors).*

(2009) (*inter alia*).[1] Solid lines in the figure represent directed relationships between observed variables, which themselves are depicted by the nodes (circles). Dotted lines represent effects emanating from unobserved variables (u), which can be thought of as composite error terms. Consequently, the figure assumes that aid (a) affects a single final outcome such as income (y) through a vector of intermediate outcomes (X).[2] In this and the subsequent discussion, it is helpful to think of intermediate outcomes as component inputs in a generic production function for final outcomes. In the case of income, these would be so-called proximate sources of growth such as physical and human capital inputs (see Mankiw, Romer, & Weil, 1992).

As depicted in the figure, a fundamental problem of identifying the causal impact of aid arises because the unobserved error terms are correlated. In the language of the SCM approach, there are "backdoor paths" running between a and X, y. The implication is that estimates of any of the relationships $a \rightarrow X$, $a \rightarrow y$ or $X \rightarrow y$ may be biased. Specifically, this can come about due to simultaneity or other forms of omitted variables bias, even when a set of conditioning variables is included (not depicted in the figure). Measurement error in the aid variable, as explicitly acknowledged by the OECD who compile the data, is a further challenge that can lead to attenuation bias.[3] A potential solution to these problems arises when one or more instrumental variables such as v is observed. As shown in the figure, this represents a parent (ancestor node) of aid and has an error structure that is unrelated to the error structure of any other variables, indicated by the absence of arcs to any of the other unobserved error terms.

It is important to understand what can and cannot be identified when a source of external variation such as v is available. First, any of the individual relationships between aid and specific intermediate outcomes (elements of X) can be identified through separate reduced form models. In these cases the intermediate outcomes are taken as the dependent variable to be explained. Second, assuming the same broad model is valid for other final outcomes—i.e., a generic production function approach with similar proximate inputs is appropriate—then alternative outcomes can be identified in addition to y. For any chosen dependent variable, the ratio of the relationships $v \rightarrow y$ to $v \rightarrow a$, suitably adjusted for other covariates, would correspond to an instrumental variables estimator for the effect $a \rightarrow y$. This corresponds to what Angrist and Pischke (2008) refer to as the ratio of estimates from "long" and "short" regressions. Appendix A provides a more formal exposition of these ideas.

All estimates of the kind described above should be seen as reduced forms precisely because they may capture impacts through a wide variety of channels (e.g., multiple elements of X), but do not provide any information as to the specific composition of these channels. Reduced form estimates do not control for the potentially complex pattern of interactions between intermediate outcomes, as well as reverse feedback (e.g., from final outcomes to intermediate outcomes). To give an example, aid may have a positive effect on household income through a variety of channels such as public investment. Aid may also have a direct positive effect on education (e.g., by funding school-building and teacher training) but also an indirect effect via higher incomes. A reduced form estimate of the aid → education relationship would not distinguish between these direct and indirect effects due to aid. As a consequence, one cannot simply add up estimates from different reduced form results to get a "total" effect of aid.

In contrast to a reduced form approach, a structural form model would aim to isolate these direct effects in order gain insight into the structure of relationships between multiple variables. In the present case, estimation of the structural form would require multiple instrumental variables to control for unobserved correlation between intermediate and final outcomes, as well as a precise understanding of the form of interactions between intermediate variables (for elaboration see Arndt, Jones, & Tarp, 2011). Finding a host of valid instrumental variables for outcomes such as education and health is controversial (arguably, more so than for foreign aid; see Acemoglu & Johnson, 2007). Thus, in the present exercise we focus uniquely on reduced form estimates. Thus, we leave for future research the issue of exploring the details of the interconnections between aid, intermediate outcomes, and final outcomes.

(b) *Estimation strategy*

The previous sub-section argued that the effect of aid on a broad range of final and intermediate outcomes can be estimated via a series of (separate) reduced form regression models. Following Figure 1, one benefit of this approach is that the same instrument for aid can be used in each model. It follows that this instrument plays a crucial role and must be selected with care. Our point of departure is earlier work published in Arndt, Jones, and Tarp (2010b), hereafter abbreviated to AJT10. In AJT10 we generated an external instrument for aid (per capita) from a model of its supply-side determinants at the donor-recipient level. This was developed as a modification of the instrument proposed by Rajan and Subramanian (2008), which in turn was inspired by the earlier contribution of Tavares (2003).

We adopt the same strategy here. As in AJT10 we specify a supply-side model for aid as follows:

$$\begin{aligned}\mathrm{Aid}_{dr}/\mathrm{POP}_r = {} & \beta_0 + \beta_1\mathrm{COLONY}_r + \beta_2\log(\mathrm{POP}_d/\mathrm{POP}_r)\\ & + \beta_3\mathrm{COLONY}_r \times \log(\mathrm{POP}_d/\mathrm{POP}_r)\\ & + \beta_4\mathrm{CURCOL}_{dr} + \beta_5\mathrm{COMLANG}_{dr}\\ & + \theta_d\mathrm{DONOR}_d + \epsilon_{dr} \qquad (1)\end{aligned}$$

where d indexes donors, r recipient countries; CURCOL is a dummy variable taking the value of one if the recipient country is currently a colony of the donor; COLONY is a dummy variable taking the value of one if the recipient country was ever a colony (of any country); COMLANG is a dummy variable taking the value of one if the recipient country has a language in common with the donor; POP is population size; and DONOR are donor fixed effects. In AJT10 the variables entering this model were averaged over the same periods considered by Rajan and Subramanian (2008), with the preferred specification referring to 1970–2000. In the present study we take advantage of new data and extend the period of interest to 2007 (i.e., up to the start of the global financial crisis). Predicted aid receipts from this model are aggregated upward by recipient to give a total predicted aid inflow for each country over the period. This variable, denoted as $\hat{v}_r = \sum_d (\widehat{\mathrm{Aid}_{dr}/\mathrm{POP}_r})$, constitutes the generated instrument for aid.

The remaining aspects of our empirical approach can be set out in general form as follows:

$$a_r = \gamma_0 + \gamma_1\hat{v}_r + T_r'\gamma_2 + \eta_r \qquad (2)$$

$$y_r = \delta_0 + \delta_1\hat{a}_r + T_r'\delta_2 + \varepsilon_r \qquad (3)$$

Eqn. (2) is the familiar first-stage of a two-stage least squares system, where a_r refers to the aid variable of interest, assumed endogenous; $\hat{v}_r$ is a source of exogenous variation in aid, described above; and T_r is a vector of additional control variables for initial conditions. Eqn. (3) is the second-stage equation, where δ_1 constitutes the effect of aid on the outcome of interest (y), which is identified assuming the generated aid instrument is relevant and mean-independent of the outcome error term—i.e., we require $E(\varepsilon|T,\hat{v}) = E(\varepsilon) = 0$. Eqn. (3) also can be recognized as a reduced form model of the effect of aid on the outcome. This is because the system does not specify the intermediate channels through which aid affects the outcome. Consequently, estimates for δ_1 refer to the total effect of aid regardless of the channels or pathways through which this comes about.

Following the previous discussion, a range of outcome variables can be used to represent y. In so doing, and presuming that the same set of control variables (T) are employed, then the first stage regression will not change, meaning that the strength of the instrument will also remain the same. However, since the instrument is derived from observational data and the "true" set of exogenous background variables (such as initial conditions) is unknown, there is no guarantee that the instrument will be valid in all cases. Put differently, as the instrument does not derive from a randomized design, there may be some outcomes for which ε is not independent of $\hat{v}$ or T. Consequently, a metric of instrument validity would be of use.

As shown in the above system of equations, only a single excluded instrument for aid is employed in the first stage. This means that over-identification tests cannot be employed. To get around this constraint, we replicate the tests employed in AJT10, whereby aggregated versions of the underlying supply-side variables used to generate the aid instrument are employed directly as the instruments in the first stage (aggregate-level) regressions. Specifically, we employ the first three terms on the RHS of Eqn. (1) namely, relative population sizes, a dummy for whether the (recipient) country was ever a colony, and their interaction term. Hansen/Sargan tests deriving from the same instrumental variables regressions as above, but now using the disaggregated instrument set, thus provide some insight as to instrument validity. Although such tests should only be considered indicative, the point is that the overall coherence of our results derive from a consideration of the impact of aid (and the suitability of our instrument) over a broad range of outcomes. Thus, in contrast to earlier work, an important contribution of the present analysis is that it does not rely exclusively on the relationship between the aid instrument and a single outcome. If the aid–growth results of Arndt *et al.* (2010a) were driven by an invalid or weak instrument, then our use here of an updated dataset and consideration of alternative outcome variables provides ample opportunities to expose these properties.

In implementing this empirical strategy, a number of important practical decisions need to be made. First is the question of the time period over which causal effects are to be estimated. A large part of the modern aid–growth literature has employed panel data, focusing on relatively short term effects of up to 5 years. However, there are good reasons to believe that the impact of developmental aid is cumulative and long-term in nature. This notion is captured in Woolcock's metaphorical distinction between growing sunflowers *versus* oak trees (Woolcock, 2009, 2011; also Temple, 2010). For instance, the impact of aid that finances an expansion of access to education may only be visible in aggregate indicators of education outcomes after a significant proportion of the population has passed through the education system. In turn, individuals must complete their education and then find work for this expansion to have a measurable effect on growth. This implies there may be a very long lag between receiving aid and being able to distinguish any form of aggregate effects.

One way to address this challenge, adopted by Clemens *et al.* (2012), is to restrict the analytical focus to the effects of "early impact" aid on growth. By construction, this excludes aid toward many key areas, including the social sectors, and presumes that a clear distinction can be made (theoretically and empirically) between different types of aid flows. Other analysts have focused on the effect of specific types of aid on narrower outcomes such as education. However, both of these approaches have their drawbacks. Aid given to a specific sector (objective) may not exclusively affect outcomes within that sector (objective). For instance, aid to education may well bring health-related benefits (and vice versa). Sector-specific measures of aid are also problematic due to difficulties in attributing multi-sector funds to individual sectors, thereby adding to measurement error concerns. Moreover, OECD-DAC data regarding aid disbursements at the sector level are only available for a small number of recent years. This means that over longer time horizons it is necessary to impute sector-specific disbursement data from data on aid commitments, the values of which are known to diverge significantly both for individual aid components and in total (Odedokun, 2003).

In light of these issues, as well as our objective of taking a broad view of aid effectiveness, we focus on the cumulative effects of aid for a cross-section of countries over the 1970–2007 period. In doing so we do not restrict our focus to specific types of aid, nor to specific types of outcomes. Due to concerns regarding the quality of sector-specific aid data, we only use aggregate measures of aid (specifically, net aid disbursements). We recognize that this measure is imperfect and masks substantial differences in both aid quality and development intentions. However, this measure of aid is transparent and allows for fungibility between sectors. In addition, we dispense with a dynamic panel approach and consider only the long-run static effect of aid. To do so, the principal variables in Eqns. (2) and (3) (namely, aid and the chosen outcome) are measured in terms of their average values for the full period. Admittedly, the choice of this period may be sub-optimal. The period 1970–2007 could represent an insufficient or excessive window of time to fully capture the effect of aid for some outcomes. However, as we are not aware of any optimal window, we consider 1970–2007 sufficiently "long" to count as the long-run and use all available information over this period in taking period averages for the aid and outcome variables.

It should be emphasized that this long-run averaging procedure applies only to the outcomes and aid inputs specified in Eqns. (2) and (3). Background control variables, denoted by T, are measured at their observed value in 1970, or the nearest available data point. The reason for this is that it avoids confounding the impact of aid with effects that occur contemporaneously through other intermediate variables. That is, in the language of SCM, we make sure not to block any pathways through which aid may influence the outcome of interest. For a small number of chosen outcomes (e.g., for education and health), however, observations are scarce in the early years of the 1970–2007 period, but increase over time. In order to avoid the long period average being dominated by more recent observations, in these cases we use a simple arithmetic mean of the earliest and latest observations, thereby assuming a linear trend over time. In a few other cases (e.g., for poverty rates), data are unavailable in the 1970s and early 1980s. Here we define the dependent variable as the endpoint level (Appendix B lists the variables to which this applies). [4]

A separate issue concerns the scale used to measure aid in Eqn. (2). Raw values are not informative due to differences in income and population between countries. One option is to scale total aid received by a given country (over time) by its population size, suggesting per capita aid as the "treatment" variable of interest. This is an intuitive measure and is technically appealing as many intermediate outcomes are expressed in population terms (e.g., average years of schooling, life expectancy). However, aid per capita has specific limitations compared to the use of the aid to GDP ratio (Aid/GDP), which has been more commonly deployed in the literature to date. First, it is hard to give a sensible or clear interpretation to any estimated effect of aid per capita on key macroeconomic outcomes, where variables are often measured in terms of or scaled by GDP. For instance, suppose we find that an inflow of US$10 of aid per capita causes the GDP growth rate to rise by 1 percentage point. Although this may be of interest *per se*, the problem is that the implied benefit–cost ratio is ambiguous because it depends on the initial size of the economy. Second, it is reasonable to assume that the real cost of providing a given flow of public services, such as education, tends to increase with GDP. Thus, especially over long time frames, the relative purchasing power of aid over a wide range of outcomes is best considered in economic terms, not population terms. [5] For these reasons, unless noted otherwise, we employ Aid/GDP as the measure of aid.

A large number of variables might be considered candidates for inclusion as either final or intermediate outcomes. [6] However, data availability and computational limitations mean that some exclusions must be imposed *ex ante*. With respect to final outcomes, we focus on growth, poverty, inequality, and the sectoral composition of value added. The first three of these variables are intimately connected (see Bourguignon, 2003); therefore, we should expect to see a consistent pattern of effects across them. The remaining variables capture the extent of changes across different macroeconomic sectors (agriculture, industry, and services). Historical experiences indicate that sustained growth transitions are normally associated with a declining share of agriculture and a rising share of industry in value added. At the same time, there are concerns that aid may provoke Dutch Disease, which is often associated with faster growth in service sectors than manufactures (e.g., Rajan & Subramanian, 2011). By including these variables, we hope to gain insight into whether aid is associated with specific growth challenges.

For intermediate outcomes, a number of "usual suspects" emerge from previous literature. These fall into the following groups: (i) sub-components of GDP (investment, private consumption, and government consumption); (ii) components of government revenue and spending; (iii) aggregate education and health outcomes (e.g., average years of schooling, life expectancy); and (iv) monetary and financial sector effects. A number of variables from each category is employed in the reduced form analysis, thus providing coverage over a wide range of meso-level aid effects. Details of the specific variables and sources of data are given in Appendix B.

Finally, to assist comparison of estimated effects across different outcomes, the aid and outcome variables all enter the models in standardized form, meaning that they are linearly transformed to have mean zero and standard deviation one. Also, to maintain comparability with previous research, we use the same sample of 78 developing countries and the same set of control variables as in AJT10. The only exception is that we include a dummy for being an oil producer in 1960. This variable was included in robustness tests in AJT10 but is now treated as part of our core specification due to the extension of the dataset from 2000–07, which includes a period of rapid economic growth in oil-producing countries, driven by rising oil prices. That is, it controls for the spike in growth rates in the latter period for a small sub-group of countries.

3. RESULTS

This section describes the results of the modeling exercise, as well as those of a number of auxiliary sensitivity and falsification tests. A more detailed interpretation of results is given in Section 4.

(a) *Reduced form*

In presenting the main results, we focus on reporting estimates of the effect of aid on a range of development outcomes. Thus, estimates for other variables in Eqn. (3) are not discussed at length.[7] Even so, since the same RHS specification is used throughout, Table 1 reports more detailed regression estimates for the effect of aid on average real growth per capita (1970–2007). The table reports results using our preferred measure of aid, Aid/GDP (columns I–IV), as well as for (PPP adjusted) real aid per capita (columns V–VIII). Different columns apply different regression estimators and/or sets of instruments for aid. Focusing on the Aid/GDP results, column (I) uses an OLS estimator, which ignores the potential endogeneity of aid and just employs its observed values; column (II) is a limited information maximum likelihood (LIML) estimator; and column (III) is an inverse probability weighted least squares (IPWLS) estimator. Column (IV) directly employs as excluded instruments for aid three principal variables used to generate the supply-side aid instrument (now aggregated). This allows Hansen/Sargan tests to be applied (not reported in the table; see further below).

Table 1. *Estimates of reduced form relation between aid and growth, 1970–2007*

	OLS (I)	LIML (II)	IPWLS (III)	LIML (IV)	OLS (V)	LIML (VI)	IPWLS (VII)	LIML (VIII)
Aid/GDP	−0.115	0.639*	0.612**	0.432*	–	–	–	–
	(0.073)	(0.382)	(0.288)	(0.249)				
Aid per capita (PPP)	–	–	–	–	0.058	0.299**	0.228**	0.239**
					(0.072)	(0.138)	(0.111)	(0.120)
GDP per capita (PPP)	−0.713***	−0.318	−0.345	−0.426**	−0.663***	−0.704***	−0.698***	−0.694***
	(0.146)	(0.245)	(0.233)	(0.185)	(0.138)	(0.133)	(0.130)	(0.128)
Primary schooling	0.146	0.158	0.243	0.155	0.140	0.110	0.237	0.118
	(0.466)	(0.516)	(0.450)	(0.470)	(0.456)	(0.414)	(0.397)	(0.404)
Trade policy index	0.768***	0.693**	0.699***	0.714***	0.767***	0.810***	0.797***	0.799***
	(0.224)	(0.278)	(0.262)	(0.244)	(0.229)	(0.213)	(0.194)	(0.208)
Life expectancy	0.034*	0.058**	0.058***	0.052***	0.037**	0.036**	0.036**	0.036**
	(0.017)	(0.024)	(0.021)	(0.019)	(0.017)	(0.017)	(0.017)	(0.017)
Geography	0.082	0.102	0.123	0.097	0.102	0.172*	0.173*	0.155
	(0.105)	(0.122)	(0.102)	(0.112)	(0.109)	(0.095)	(0.091)	(0.100)
Coastal pop. dens.	0.037***	0.035**	0.043***	0.035**	0.038***	0.044***	0.047***	0.042***
	(0.011)	(0.017)	(0.016)	(0.014)	(0.011)	(0.011)	(0.011)	(0.011)
Malaria prevalence	−0.957***	−1.190***	−1.077***	−1.127***	−1.008***	−1.069***	−0.921***	−1.054***
	(0.261)	(0.390)	(0.315)	(0.325)	(0.275)	(0.281)	(0.237)	(0.267)
Civil liberties	−0.051	−0.344	−0.385	−0.263	−0.158	−0.415	−0.393	−0.351
	(0.269)	(0.389)	(0.321)	(0.323)	(0.280)	(0.330)	(0.264)	(0.317)
Air distance	−0.259	−0.344*	−0.394**	−0.320*	−0.279	−0.306*	−0.334**	−0.299*
	(0.180)	(0.205)	(0.199)	(0.192)	(0.183)	(0.171)	(0.169)	(0.166)
Oil producer	0.438**	0.912***	1.009***	0.783***	0.527***	0.593***	0.675***	0.577***
	(0.178)	(0.283)	(0.215)	(0.218)	(0.165)	(0.157)	(0.127)	(0.149)
N	78	78	78	78	78	78	78	78
R^2 (centered)	0.74	0.54	0.55	0.64	0.74	0.69	0.69	0.71
Weak id. statistic		9.16	12.13	4.26		18.44	11.70	9.56
Anderson–Rubin test		0.02	0.01	0.06		0.02	0.01	0.06
Endogeneity test		0.012	0.018	0.005		0.034	0.071	0.029

Notes: Dependent variable is real GDP growth; intercept, investment prices, and region dummies included but not shown; growth and aid measures enter in standardized form; columns I and V estimated by OLS; columns II and VI estimated by LIML, using a single generated aid instrument; columns III and VII replicate the latter with IPWLS; columns IV and VIII use three instruments for aid taken from the zero stage regression employed to generate the single aid instrument; final two rows report probabilities; endogeneity test is Durbin–Wu–Hausman χ^2; standard errors (in parentheses) are robust.

Source: Authors' calculations; see Appendix B for variable definitions and sources.

* Significance: 0.1.

** Significance: 0.05.

*** Significance: 0.01.

The LIML and IPWLS estimators are both instrumental variables estimators. The former is a standard alternative to a two-stage least squares estimator and is numerically equivalent where one excluded instrument is employed. However, where the model is not just-identified (as in column IV) the LIML estimator is known to be more robust to the presence of weak instruments (Stock, Wright, & Yogo, 2002). The IPWLS estimator, presented in detail in Arndt *et al.* (2010b), instruments for aid but employs a binary aid instrument and applies weights to the data giving greater emphasis to the part of the empirical distribution of covariates where there is most overlap between "high" and "low" aid recipients (according to the instrument). This approach is "doubly robust" but has the disadvantage of discarding valuable information and, therefore, may lead to an efficiency loss. Thus, it should be seen primarily as a robustness check on the linearity assumption underlying the LIML results. For all results employing instrumental variables, tests of instrument strength are reported.

The OLS results in Table 1 provide no evidence of a positive impact of aid on growth. However, once the endogeneity of aid is accounted for using instrumental variables techniques, this conclusion is rejected and a positive and statistically significant impact is found. Interpretation of the LIML point estimates in column (II) for Aid/GDP are as follows: a one standard deviation increase in Aid/GDP is expected to boost growth by 0.64 standard deviations on average, holding all other variables fixed. One can translate this estimated effect to raw units by referring to the information in Appendix B. This shows that a one standard deviation of the real GDP growth rate equals 1.79 percentage points; and a one standard deviation increase in Aid/GDP represents 3.77 percentage points. Thus, an aid–growth effect of 0.61 standard deviation units implies a $0.64 \times (1.79/3.77) = 0.30$ percentage point effect in raw terms. Put more simply, a one percentage point increase in Aid/GDP is expected to boost the real GDP growth rate by 0.30 percentage points. The IPWLS results are essentially the same.

Before proceeding to consider other outcomes, three additional comments on the results in Table 1 can be made. First, although the instrumental variables results are positive and statistically significant, the respective confidence intervals are relatively wide suggesting the effect is not precisely estimated. This is not a surprise given the nature of the data and sample size. Nonetheless, a positive effect of aid on growth is found for both the Aid/GDP and aid per capita measures, giving credence to the findings. Indeed, when applied to the full range of intermediate and final outcomes, use of aid per capita yields highly consistent results with those presented here (for full details see Arndt, Jones, & Tarp, 2011).

Second, instrument strength and validity tests give no cause for concern. The Andersen–Rubin test, which is robust to weak instruments, confirms a statistically significant partial correlation between the endogenous variable (aid) and the outcome (growth). Although they are not reported in the table, Hansen-*J* tests calculated from the estimates in columns (IV) and (VIII) are passed comfortably, supporting the validity of the generated instrument. However, the weak identification (Kleibergen–Paap F) statistic shows that when three aggregated instruments are used instead of a single generated instrument, instrument strength declines. Thus, for interpretation we focus on estimates from the LIML (and IPWLS) estimators that employ a single instrument.

Third, coefficient estimates for other covariates included in the model are plausible. All of these refer to initial conditions and are measured as the value in 1970 (or thereabouts). Thus, the interpretation is that trade openness at the beginning of the period is associated with more rapid subsequent growth, and malaria prevalence in 1970 is associated with slower average growth. In the present specification the estimate on the level of GDP per capita represents a convergence effect—the negative sign indicates that lower income countries grow faster on average. Inclusion of this term is appropriate because the dependent variable here is measured in differences. For other outcomes (see below), which are measured in levels, inclusion of the initial level of the same variable on the RHS is not necessary (e.g., see Acemoglu & Johnson, 2007; Bloom, Canning, & Sevilla, 2004). However, as we retain the same specification throughout, the GDP level term simply acts to control for initial income.

The finding of a positive effect of aid on growth is important, but it raises equally critical distributional questions. To make a contribution to development in a wider sense, economic growth should benefit poorer households. Thus it is relevant to validate the effect of aid on other aggregate welfare outcomes. These are considered in Table 2, which summarizes results from separate regressions for a chosen set of final outcomes. The same specification and instrument (s) for aid are employed as in Table 1; also, results are reported for each the three main estimators employed before—OLS, LIML, and IPWLS.[8] To assist interpretation, Aid/GDP and all outcome variables are entered in standardized form. Thus, each cell of Table 2 gives the standardized coefficient on the aid

Table 2. *Summary of reduced form results for relationships between Aid/GDP and final outcomes*

	N	OLS	Pr.	LIML	Pr.	IPWLS	Pr.	Endog. test	Hansen-*J*
GDP per capita growth	78	−0.115	0.12	0.639	0.09*	0.612	0.03**	0.01.**	0.84
Agriculture, value added (%GDP)	76	0.035	0.69	−0.652	0.08*	−1.062	0.01***	0.03**	0.89
Industry, value added (% GDP)	76	0.220	0.10	0.648	0.24	0.822	0.04**	0.42	0.68
Services, etc., value added (% GDP)	76	−0.273	0.04**	0.242	0.56	0.634	0.23	0.19	0.65
Poverty headcount at \$2 a day	64	0.018	0.84	−0.471	0.05*	−0.438	0.19	0.01**	0.42
Poverty headcount at \$1.25 a day	64	0.084	0.57	−0.487	0.09*	−0.285	0.42	0.01**	0.78
Gini index	65	−0.186	0.14	−0.035	0.91	−0.328	0.57	0.56	0.07*

Notes: Each cell of columns OLS, LIML, and IPWLS reports the standardized coefficient on Aid/GDP from individual reduced form regressions in which the row variable enters as the dependent variable; adjacent columns report the corresponding probability that the true parameter is equal to zero; estimation method is indicated by the column headings; aid treated as endogenous in LIML and IPWLS only; "Endog. test" reports the probability from a Durbin–Wu–Hausman test that aid can be treated as exogenous; "Hansen-*J*" gives the probability associated with the Hansen-*J* statistic from a LIML regression using three aggregated instruments; all regressions include the same set of control variables (see text) and employ robust standard errors.
Source: authors' calculations; see Appendix B for variable definitions and sources.
* Significance: 0.1.
** Significance: 0.05
*** Significance: 0.01.

to GDP ratio and, in the adjacent cell, the estimated probability that the true parameter estimate equals zero. These correspond to results from individual regressions in which the row variable is the outcome of interest. Thus, for the growth outcome, the results are extracted directly from the corresponding columns of Table 1.

Complementing the positive aid–growth result, we find that aid reduces poverty but leaves inequality unaffected on average. While there are some differences between the LIML and IPWLS estimators as regards the statistical significance of these results, in part reflecting the limitations of the data, the magnitude, and direction of the estimates are highly consistent. Moreover, this pattern is in keeping with the theoretical relation between growth, poverty, and inequality (Bourguignon, 2003). In addition, the results show that aid inflows are associated with a decline in the weight of agriculture in GDP, implying that aid stimulates relatively more rapid growth of non-agricultural sectors. Indeed, the IPWLS estimates (and OLS) indicate a corresponding increase in industry's GDP share; however, the impact on services is more ambiguous.[9]

The final two columns of Table 2 report additional test statistics. First are results from Durbin–Wu–Hausman χ^2 tests, where the null hypothesis is that the aid variable can be treated as exogenous.[10] For four of the seven outcome variables the test is rejected, suggesting that concerns surrounding the endogeneity of aid are significant. This is notably the case for both growth and poverty, one explanation being that these variables are used directly by donors to decide how much aid to provide. However, the same test is not rejected for the three other variables, implying the corresponding OLS results are both consistent and efficient. The final column gives the probability from Hansen-*J* over-identification tests, based on the same regression specification in the LIML column, but employing three aggregated instruments for aid instead of the single generated aid instrument (see Section 2). A significant result (<10%) is ground to reject the joint null hypothesis that the instruments are valid, meaning they are uncorrelated with the estimated regression residuals. This test is passed comfortably in all cases except for the Gini coefficient. On the one hand, this may be taken to imply that both the OLS and instrumental variables (LIML, IPWLS) point estimates for the Gini are biased. On the other hand, this may be spurious—assuming the tests are independent, the probability that the Hansen-*J* test is passed at the 10% level in all seven cases is less than one in two, even if the null hypothesis is always true.

Table 3 reports reduced form results for the effect of Aid/GDP on the chosen set of intermediate outcomes, adopting the same format as Table 2. Again, the pattern of results is broadly consistent with a view that aid has a positive developmental impact on average. For instance, aid is associated with a larger investment share as well as a higher share of government consumption and government revenues in GDP. The latter measure excludes income from grants and therefore suggests a positive impact of aid on tax income. Estimates for sub-components of government spending indicate that aid boosts expenditure in social sectors, especially education.

As discussed further in Section 4, the impact of aid on a number of key social outcomes corroborates positive results in previous studies. We find that aid has a positive causal effect on average years of schooling, and secondary schooling in particular, likely operating through the government expenditure channel discussed in the preceding paragraph. Further, the signs of the estimated coefficients on health outcomes clearly point to a positive developmental contribution of aid even though the estimated impact of aid on government health expenditure is insignificant. The LIML estimates on the health outcomes also slightly exceed conventional significance levels in most cases; nonetheless, the IPWLS estimates for both infant mortality and life expectancy are statistically significant.

Table 3. *Summary of reduced form results for relationships between Aid/GDP and intermediate outcomes*

	N	OLS	Pr.	LIML	Pr.	IPWLS	Pr.	Endog. test	Hansen-*J*
Investment (% GDP)	78	0.319	0.00***	0.795	0.03**	0.357	0.16	0.15	0.73
Consumption (% GDP)	78	0.174	0.24	−0.515	0.25	−0.779	0.05**	0.07*	0.70
Government (% GDP)	78	0.513	0.00***	0.758	0.06*	0.873	0.05*	0.51	0.03**
Revenue, excluding grants (% GDP)	69	0.470	0.03**	2.362	0.00***	1.188	0.00***	0.00***	0.39
Health expend., public (% GDP)	78	0.403	0.10	0.363	0.36	−0.610	0.23	0.91	0.21
Education expend., public (% GDP)	76	0.485	0.00***	1.423	0.00***	1.644	0.00***	0.01**	0.18
Military expenditure (% GDP)	77	0.387	0.11	0.361	0.32	0.255	0.65	0.94	0.37
Av. years total schooling, 15+	72	0.206	0.10	1.010	0.04**	0.511	0.08*	0.03**	0.96
Av. years primary schooling, 15+	72	0.267	0.04**	0.673	0.13	0.246	0.28	0.33	0.97
Av. years secondary schooling, 15+	72	−0.006	0.97	1.476	0.03**	0.818	0.04**	0.00***	0.68
Life expectancy at birth, total (years)	78	−0.087	0.09*	0.187	0.16	0.329	0.06*	0.02**	0.33
Infant mortality rate	75	0.055	0.52	−0.306	0.17	−0.434	0.06*	0.09*	0.46
Mortality rate, under-5 (per 1,000)	75	0.101	0.23	−0.320	0.15	−0.297	0.14	0.04**	0.44
Death rate, crude (per 1,000 people)	78	0.216	0.01***	−0.162	0.40	−0.116	0.57	0.02**	0.97
Fertility rate (births/woman)	77	−0.053	0.47	−0.344	0.09*	0.102	0.49	0.12	0.05*
Consumer price inflation (%)	77	0.114	0.34	−0.677	0.18	−0.793	0.21	0.07*	0.70
Real interest rate (%)	77	−0.017	0.95	−0.488	0.37	0.199	0.70	0.36	0.27
Domestic credit to private sector (% GDP)	78	−0.053	0.61	−0.009	0.98	0.360	0.51	0.90	0.69

Notes: Each cell of columns OLS, LIML, and IPWLS reports the standardized coefficient on Aid/GDP from individual reduced form regressions in which the row variable enters as the dependent variable; adjacent columns report the corresponding probability that the true parameter is equal to zero; estimation method is indicated by the column headings; aid treated as endogenous in LIML and IPWLS only; "Endog. test" reports the probability from a Durbin–Wu–Hausman test that aid can be treated as exogenous; "Hansen-*J*" gives the probability associated with the Hansen-*J* statistic from a LIML regression using three aggregated instruments; all regressions include the same set of control variables (see text) and employ robust standard errors.
Source: authors' calculations; see Appendix B for variable definitions and sources.
* Significance: 0.1
** Significance: 0.05.
*** Significance: 0.01.

It should be recalled, however, that the outcome data employed here are of mixed quality and coverage. This is likely to inflate the imprecision of our results. This concern is particularly present in relation to findings for the monetary and financial sector indicators. These results are ambiguous, suggesting no evidence of a systematic effect of aid on inflation, real interest rates, or credit to the private sector. However, since these outcomes are particularly noisy and poorly scaled, these results should not be given too strong an interpretation.

Finally, the test statistics in the final two columns of Table 3 broadly follow the pattern of Table 2. For nearly two thirds of the intermediate outcomes we must reject the null hypothesis that aid is exogenous (at conventional significance levels of <10%). Thus, although aid may not be endogenous for all possible outcomes, such endogeneity needs to be taken seriously when considering intermediate outcomes—i.e., *ex ante*, aid cannot be assumed to be exogenous. With respect to the validity of the generated aid instrument, we cannot reject the null of the Hansen-J test for the large majority of outcomes, the two exceptions being government size and fertility rates. Again, assuming the tests are independent, the probability that the Hansen-J test is passed at the 5% level in all 18 cases is less than 40%, even if the null hypothesis is always true. Overall, therefore, the instrument performs well and there are no clear grounds on which to reject its suitability.

(b) *Sensitivity and falsification tests*

How sensitive are these results to alternative assumptions? While a battery of tests is possible, we focus on three specific aspects. The first involves application of a quantile regression estimator, evaluated at the median of the conditional outcome distribution, which is less sensitive to outliers than OLS methods or its instrumental variables analogs. To address the endogeneity of aid, a two step approach is adopted. In the first stage a quantile regression is run of observed Aid/GDP against the generated instrument and all other covariates employed in the model. The second stage, also estimated via a quantile estimator, is the outcome regression of interest which uses predicted values from the first stage in place of raw Aid/GDP.[11] Second, due to the concern that the reduced form results may incorporate effects that occur through income growth (e.g., aid → growth → education), it is helpful to include growth on the RHS. This essentially "blocks" all paths to the outcome variable that arise via changes in income. Third, we include raw population size on the RHS of the specification. The rationale for this is that the generated aid instrument depends on the relative population sizes of the donors to recipients. As Clemens and Bazzi (2009) argue, there are reasons to believe that population is a direct determinant of growth (and perhaps other key outcomes), which might invalidate the instrument. Thus, by including population on the RHS of the outcome regression we allow any such direct effects to be incorporated.[12]

The results from the sensitivity tests are summarized in Table 4. As in previous tables, only the estimated coefficient on the Aid/GDP variable is reported alongside standard errors (in parentheses). Results for different estimators are indicated by abbreviations; the different specifications are in the columns, where the baseline specification is unchanged from earlier models. The broad finding from these tests is that both the direction and magnitude of the effects of aid are unchanged. For instance, the estimated aid–growth effect is 0.58 using a quantile instrumental variables regression and 0.68 using both

Table 4. *Summary of sensitivity tests for effects of Aid/GDP on various outcomes*

Outcome	Estim.	Specification					
		Baseline		Growth control		Pop. control	
		Beta	s.e.	Beta	s.e.	Beta	s.e.
Growth (per capita)	Q-reg	0.58*	(0.34)	.	.	0.82**	(0.35)
	LIML	0.64*	(0.38)	.	.	0.81*	(0.48)
	IPWLS	0.61**	(0.29)	.	.	0.68**	(0.33)
Poverty count ($2 a day)	Q-reg	−0.54*	(0.28)	−0.55**	(0.26)	−0.86***	(0.31)
	LIML	−0.47*	(0.24)	−0.34*	(0.19)	−0.55**	(0.26)
	IPWLS	−0.44	(0.33)	−0.19	(0.28)	−0.45	(0.34)
Revenue, excl. (% GDP)	Q-reg	1.73***	(0.37)	1.65***	(0.40)	2.08***	(0.48)
	LIML	2.36***	(0.71)	1.82***	(0.41)	2.70***	(0.96)
	IPWLS	1.19***	(0.39)	0.87***	(0.33)	1.16***	(0.42)
Investment (% GDP)	Q-reg	0.83**	(0.38)	0.96***	(0.33)	1.47***	(0.42)
	LIML	0.80**	(0.36)	0.64**	(0.30)	0.93**	(0.43)
	IPWLS	0.36	(0.25)	0.25	(0.24)	0.40	(0.25)
Life expectancy (years)	Q-reg	0.08	(0.16)	0.16	(0.15)	0.37*	(0.21)
	LIML	0.19	(0.13)	0.10	(0.11)	0.23	(0.15)
	IPWLS	0.33*	(0.18)	0.24	(0.15)	0.35*	(0.19)
Infant mortality rate	Q-reg	−0.35*	(0.18)	−0.48***	(0.17)	−0.42*	(0.22)
	LIML	−0.31	(0.22)	−0.26	(0.19)	−0.32	(0.26)
	IPWLS	−0.43*	(0.23)	−0.39*	(0.22)	−0.43*	(0.25)
Av. years total schooling	Q-reg	0.52**	(0.22)	0.65**	(0.28)	0.56**	(0.28)
	LIML	1.01**	(0.48)	0.86**	(0.41)	1.25*	(0.68)
	IPWLS	0.51*	(0.29)	0.47	(0.29)	0.53*	(0.32)

Notes: "Outcome" is the dependent variable; "Estim." indicates the regression estimator employed—Q-reg is a two-step instrumental variables quantile regression (50th percentile). "Baseline" specification is as per that used in Tables 2 and 3. "Growth control" specification includes real GDP per capita growth on the RHS; "Pop. control" includes population (in millions) on the RHS. "Beta" reports the estimated coefficient on Aid/GDP and "s.e." gives robust standard errors in parentheses; all variables are standardized.
Source: Authors' calculations; see Appendix B for variable definitions and sources.
* Significance: 0.1.
** Significance: 0.05.
*** Significance: 0.01.

the latter estimator and controlling for population size (also statistically significant at the 10% level in both cases). While there is some variation in the precise point estimates and standard errors, this is to be expected given the small sample size and noisy underlying data. Two further points emerge. First, when growth is included as a control variable on the RHS, many of the point estimates for non-growth outcomes decline slightly in magnitude relative to the baseline specification. For example, the estimated effect of aid on poverty using the LIML is -0.47 in the baseline specification but -0.34 when growth is controlled for. This implies that some of the impact of aid on poverty reduction is occurring through (aggregate) income growth. Second, the opposite tendency is found when we control for population size. This may be due to a negative correlation between average population and the generated aid instrument (i.e., larger countries are expected to receive less aid).

As noted in Section 2, the hypothesis that aid (exclusively) affects final outcomes through proximate determinants can be tested by including the latter covariates as controls on the RHS. Again, this amounts to blocking the effects of aid through these channels as their effect is partialled out of the corresponding effect of Aid/GDP on the chosen outcome. Both LIML and IPWLS estimates of such augmented regressions show no statistically significant effect of aid on growth when life expectancy, education and investment outcomes are included as additional control variables. This holds regardless of the other covariates included in the model. [13] The interpretation is that a large share of estimated effect of aid on growth is likely to come through these key channels.

4. INTERPRETATION

Thus far, discussion of results has concentrated on the sign and domain of parameter estimates. It is helpful to reflect on whether these ranges are plausible. This is sometimes difficult to ascertain directly from the previous tables as results were given in standardized form. Consequently, for a selected number of final and intermediate outcomes, Table 5 presents the reduced form point estimates and 90% confidence intervals for the expected return to an average annual aid inflow equal to 5% of GDP over the period 1970–2007 (which is slightly greater than double the observed median Aid/GDP for all countries in the sample; see Appendix B). We find the long-run impacts of aid are both plausible and material. According to the LIML point estimates, such an aid inflow is expected to increase the average annual rate of economic growth by around 1.5 percentage points, reduce poverty by around 15 percentage points, raise the investment share of GDP by around 11 percentage points, augment average schooling by 2.8 years, boost life expectancy at birth by 2.35 years and reduce infant mortality by 14 in every 1,000 births. [14]

Viewed together, the results show a consistent pattern that aid has made a positive long-run developmental contribution on average. The results concerning the impact of aid on growth are also consistent with other research. The present study applies the same methods developed in AJT10 to an extended dataset, yielding highly consistent point estimates for the aid–growth coefficient. Specifically, for the equivalent specification and estimator, Arndt *et al.* (2010a) report a coefficient of 0.42 on aid and a standard error of 0.19 for the period 1970–2000, implying a 95% confidence interval ranging from 0.05 to 0.79. The comparable (unstandardized) coefficient corresponding to column III of Table 1 is 0.29, with a standard error of 0.18. Thus, for the periods 1960–2000, 1970–2000 (estimated in AJT10) and 1970–2007 (here), the estimated impact of aid on growth lies in a highly consistent domain.

The reduced form results for other outcomes also are consistent with previous studies. Investment is frequently identified as a principal growth determinant (Mankiw *et al.*, 1992; Sala-i-Martin, Doppelhofer, & Miller, 2004), and evidence points to (very) long-run growth effects from improvements in aggregate health (Jack & Lewis, 2009). Gomanee, Girma, and Morrissey (2005), Masud and Yontcheva (2007) and Mishra and Newhouse (2009) all find positive effects of certain kinds of aid on health outcomes; while Birchler and Michaelowa (2013) reports positive effects of aid on education enrollment rates (also Michaelowa, 2004). Similar to our falsification test, Hansen and Tarp (2001) find that aid is not significant in a growth regression which controls for investment and human capital, but that aid remains a significant determinant of investment. Furthermore, our results provide a basis to reject the (largely) theoretical concerns that aid undermines domestic revenue mobilization (e.g., Moss, Pettersson, & van de Walle, 2006). Rather, our results are closer to those of Pivovarsky, Clements, Gupta, and Tiongson (2003), who find a positive revenue impact from concessional loans (but a small negative effect from grants). Similarly, and contrary to concerns that aid's positive developmental impact is muted due to its fungibility (Pack & Pack, 1993), our results corroborate van de Walle and Mu (2007) and show that some aid "sticks" to the social sectors and, thus, is not entirely fungible.

Table 5. *Estimated increment to the long-run average level of various outcomes expected from receiving a sustained aid inflow equal to 5% of GDP (over the period 1970–2007)*

Variable	OLS			LIML		
	Lower	Point	Upper	Lower	Point	Upper
GDP per capita growth	−0.56	−0.27	0.02	0.02	1.51	3.00
Agriculture, value added (% GDP)	−1.98	0.62	3.21	−22.38	−11.60	−0.83
Poverty headcount at $1.25 a day	−5.19	2.67	10.54	−30.70	−15.52	−0.35
Investment (% GDP)	2.46	4.35	6.25	2.86	10.85	18.85
Government (% GDP)	3.22	5.72	8.21	1.02	8.45	15.87
Revenue, excluding grants (% GDP)	1.34	5.49	9.64	13.93	27.60	41.27
Av. years total schooling, 15+	−0.01	0.57	1.15	0.59	2.81	5.02
Life expectancy at birth, total (years)	−2.18	−1.10	−0.02	−0.41	2.35	5.11
Infant mortality rate	−4.06	2.56	9.17	−31.61	−14.37	2.88

Notes: The table reports the raw estimated effect of a 5% Aid/GDP inflow on selected outcome variables based on the reduced form regressions summarized in Tables 2 and 3; "lower" and "upper" refer to 90% confidence limits; estimators indicated by column headings.
Source: Authors' calculations; see Appendix B for variable definitions and sources.

With respect to the link between poverty and growth, the reduced form results enable us to derive an estimate of the aid-induced growth semi-elasticity of poverty (GSEP). This is given by the estimated absolute change in the poverty (headcount) rate divided by the estimated percentage change in mean income due to aid over the period.[15] For both the US$1.25 and US$2 poverty measures, we find that the aid-induced GSEP is around 0.30 (or 0.26 and 0.31 respectively), meaning that a 1% increase in mean income tends to lead to a 0.30 percentage point fall in the headcount poverty rate. This is situated just below the average of the range of GSEP estimates calculated by Klasen and Misselhorn (2008, Table 7), suggesting there is no reason to conclude that aid is any less effective in reducing poverty than other growth drivers over the long-run.

5. CONCLUSIONS

This study aimed to answer the question: "What has aid accomplished over the past four decades?" Evidence of this kind speaks to the first order policy problem facing donors (recipients)—namely, whether they should continue to provide (accept) foreign aid. Drawing on a simple graphical illustration, we began by presenting a general structural model of the relationship between aid and aggregate outcomes, which is consistent with the framework employed in the literature on growth empirics. To estimate this model, we first calculated reduced form estimates of the relationship between aid and final outcomes. The results confirm a robust positive impact of aid on growth for the 1970–2007 period, thereby replicating the findings of AJT10 (and other recent studies) using an extended dataset. The aggregate effects of aid are also coherent. On average and over the long-run, foreign aid reduces poverty with no significant impacts on inequality. Aid also contributes to more rapid expansion of "modern" sectors (industry) and a relative decline of agriculture's share in GDP.

To gain insight into relevant transmission channels, we applied the same reduced form approach to a set of intermediate outcomes. These revealed a range of positive and significant effects due to aid, including on investment, government revenue, government spending, and social outcomes. Lastly, we verified the sensitivity of the results to alternative estimators and specifications, such as when income growth and population are included as regression controls. These gave no reason to question our overall findings. Moreover, the results suggest that the effect of aid on a range of non-growth outcomes cannot be attributed solely to the impact occurring via income growth. These results were substantiated by a series of falsification tests. These suggest that investments in physical capital and improvements in human capital are likely to be key transmission channels through which aid promotes growth.

In summary, based on results covering a wide range of outcomes, aid can point to a series of accomplishments with a positive impact on the growth and development process. There is no evidence that aid is detrimental. Aid has contributed to economic growth by stimulating its proximate determinants—e.g., physical capital accumulation and improving human capital, particularly education and health. Overall, the experience of the past four decades or so provide no support to the argument that aid flows should cease. Moreover, the present analysis provides some guidance on the form of assistance by highlighting both the non-growth effects of aid as well as the importance of physical and human capital accumulation. Finally, considered in light of the the great expectations associated with aid in the 1960s and early 1970s, the magnitude of the estimated effects of aid are generally moderate but become material over the long-run. It follows that aid should not be considered a panacea or silver-bullet for stimulating growth and development.

NOTES

1. For discussion and application of this approach to the analysis of foreign aid see Arndt et al. (2011).

2. The convention adopted here is that lower case Latin letters represent individual random variables, while upper case letters refer to vectors of random variables.

3. See for example: http://www.oecd.org/dac/stats/crsguide.

4. The practice of using the endpoint level is encountered in the cross-country growth regression literature where final income can be used in place of the growth rate, and initial income is dropped from the RHS (e.g., Mankiw *et al.*, 1992).

5. Purchasing power parity corrections go some way to address this, but these face acute challenges in accurately adjusting for differences in the cost of public service provision. For discussion see http://go.worldbank.org/I0AHGSYF80.

6. The distinction between these types of outcomes is not important from a technical point of view and there may be some debate as to classifications. Nonetheless, corresponding to the logic of production functions, this terminology is retained for clarity of exposition.

7. For reference, detailed results pertaining to each of the reduced form models can be found in Arndt, Jones, and Tarp (2013, Appendix C).

8. Some of the outcomes considered in Tables 2 and 3 are bound in certain ways (e.g., above zero and/or below 100 for percentage of GDP outcomes). However, none of the observations in the data set lie on or even particularly close to these bounds. Consequently, no specific techniques are employed to take censoring into account.

9. It is not possible to say whether this result is driven by aid specifically targeted at industry or whether other mechanisms are at play. Further research will be of use to shed light here.

10. Implemented in Stata via the `endog ()` option of the `ivreg2` command used to estimate the LIML results.

11. This follows the logic of a TSLS estimator. The quantile regression estimator used is the native `qreg` command available in Stata v11 estimated at the 50th percentile, with variance–covariance matrix estimated using a bi-weight kernel. See Koenker and Hallock (2001) for an overview of this estimator.

12. As shown by Clemens and Bazzi (2009), logged population is critical for the strength of the generated instrument (see Eqn. (1)). To avoid this concern raw population size is employed, thereby controlling for some (not all) population-related effects while also maintaining instrument strength.

13. Detailed results are available on request from the authors.

14. These effects refer to the expected change in the average of the outcome variable over the full period—i.e., the difference in the average for that variable *versus* its counterfactual average.

15. Calculated as $\text{GSEP} = -\beta_p/[(1+\beta_g)^{37} - 1]$, where β_p is the estimated coefficient on Aid/GDP in the reduced form aid-poverty regression and β_g is the estimated coefficient on Aid/GDP in the reduced form aid–growth regression (appropriately scaled). Note that the latter coefficient estimates the expected increase in the average annual growth rate over the period 1970–2007, while the former estimates the expected overall change in the poverty rate due to aid over the same timeframe. Consequently, to compare like with like, we need to calculate the expected overall percentage change in mean income, which is given by the denominator of the GSEP equation.

REFERENCES

Acemoglu, D., & Johnson, S. (2007). Disease and development: The effect of life expectancy on economic growth. *Journal of Political Economy, 115*, 925–985.

Angrist, J. D., & Pischke, J. S. (2008). Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

Arndt, C., Jones, S., & Tarp, F. (2010a). *Aid, growth, and development: Have we come full circle?* WIDER working paper 2010-96. World Institute for Development Economic Research (UNU-WIDER).

Arndt, C., Jones, S., & Tarp, F. (2010b). Aid, growth, and development: Have we come full circle?. *Journal of Globalization and Development, 1* (Article 5).

Arndt, C., Jones, S., & Tarp, F. (2011). *Aid effectiveness: Opening the black box*. In WIDER working paper 2011-44. World Institute for Development Economic Research (UNU-WIDER).

Arndt, C., Jones, S., & Tarp, F. (2013). *Assessing foreign aid's long-run contribution to growth and development*. In WIDER working paper 2013/072. Helsinki: UNU-WIDER.

Balke, A., & Pearl, J. (1997). Bounds on treatment effects from studies with imperfect compliance. *Journal of the American Statistical Association, 92*, 1171–1176.

Birchler, K., & Michaelowa, K. (2013). Making aid work for education in developing countries: An analysis of aid effectiveness for primary education coverage and quality. WIDER Working Paper 2013/021. World Institute for Development Economic Research (UNU-WIDER). URL: http://ideas.repec.org/p/unu/wpaper/wp2013-021.html.

Bloom, D. E., Canning, D., & Sevilla, J. (2004). The effect of health on economic growth: A production function approach. *World Development, 32*, 1–13.

Bourguignon, F. (2003). The growth elasticity of poverty reduction: Explaining heterogeneity across countries and time periods. In T. S. Eicher, & S. J. Turnovsky (Eds.), *Inequality and growth: Theory and policy implications* (pp. 3–27). MIT Press.

Bourguignon, F., & Sundberg, M. (2007). Aid effectiveness – Opening the black box. *American Economic Review – Papers and Proceedings, 97*, 316–321.

Clemens, M., & Bazzi, S. (2009). *Blunt instruments: On establishing the causes of economic growth.* Working paper 171. Center for Global Development.

Clemens, M. A., Radelet, S., Bhavnani, R. R., & Bazzi, S. (2012). Counting chickens when they hatch: Timing and the effects of aid on growth. *The Economic Journal, 122*, 590–617.

Cohen, D., & Soto, M. (2007). Growth and human capital: Good data, good results. *Journal of Economic Growth, 12*, 51–76.

Frot, E., & Perrotta, M. (2012). *Aid effectiveness: New instrument, new results*. Mimeo. Stockholm Institute of Transition Economics. URL: <www.hhs.se/SITE/Staff/Documents/MPdraft_2012.pdf>.

Gomanee, K., Girma, S., & Morrissey, O. (2005). Aid, public spending and human welfare: Evidence from quantile regressions. *Journal of International Development, 17*, 299–309.

Hansen, H., & Tarp, F. (2001). Aid and growth regressions. *Journal of Development Economics, 64*, 547–570.

Jack, W., & Lewis, M. (2009). Health investments and economic growth: Macroeconomic evidence and microeconomic foundations. In M. Spence, & M. Lewis (Eds.), *Health and growth* (pp. 1–41). World Bank.

Juselius, K., Møller, N.F., & Tarp, F. (2013). The long-run impact of foreign aid in 36 African countries: Insights from multivariate time series analysis. *Oxford Bulletin of Economics and Statistics* (in press). http://dx.doi.org/10.1111/obes.12012.

Klasen, S., & Misselhorn, M. (2008). *Determinants of the growth semi-elasticity of poverty reduction*. IAI Discussion Papers 176. Ibero-America Institute (IAI) for Economic Research.

Koenker, R., & Hallock, K. (2001). Quantile regression: An introduction. *Journal of Economic Perspectives, 15*, 43–56.

Mankiw, N. G., Romer, D., & Weil, D. N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics, 107*, 407–437.

Masud, N., & Yontcheva, B. (2007). *Does foreign aid reduce poverty? Empirical evidence from nongovernmental and bilateral aid*. IMF working papers 07/100. International Monetary Fund.

Mekasha, T. J., & Tarp, F. (2013). Aid and growth: What meta-analysis reveals. *Journal of Development Studies, 49*, 564–583.

Michaelowa, K. (2004). Aid effectiveness reconsidered – Panel data evidence for the education sector. Discussion Paper Series 26374. Hamburg Institute of International Economics.

Mishra, P., & Newhouse, D. (2009). Does health aid matter?. *Journal of Health Economics, 28*, 855–872.

Mosley, P. (1987). Foreign aid, its defense and reform. University Press of Kentucky.

Moss, T., Pettersson, G., & van de Walle, N. (2006). *An aid-institutions paradox? A review essay on aid dependency and state building in sub-Saharan Africa.* Working Papers 74. Center for Global Development.

Odedokun, M. (2003). Analysis of deviations and delays in aid disbursements. *Journal of Economic Development, 28*, 137–169.

Pack, H., & Pack, J. R. (1993). Foreign aid and the question of fungibility. *The Review of Economics and Statistics, 75*, 258–265.

Pearl, J. (2009). Causality: Models, reasoning and inference. *2nd ed.* Cambridge University Press.

Pivovarsky, A., Clements, B.J., Gupta, S., & Tiongson, E. (2003). Foreign aid and revenue response: Does the composition of aid matter? IMF Working Papers 03/176. International Monetary Fund.

Rajan, R. G., & Subramanian, A. (2008). Aid and growth: What does the cross-country evidence really show?. *The Review of Economics and Statistics, 90*, 643–665.

Rajan, R. G., & Subramanian, A. (2011). Aid, Dutch disease, and manufacturing growth. *Journal of Development Economics, 94*, 106–118.

Riddell, R. (2007). *Does foreign aid really work?*. USA: Oxford University Press.

Sachs, J. (2006). *The end of poverty: Economic possibilities for our time*. USA: Penguin.

Sala-i-Martin, X., Doppelhofer, G., & Miller, R. I. (2004). Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach. *The American Economic Review, 94*, 813–835.

Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics, 70*, 65–94.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business & Economic Statistics, 20*, 518–529.

Tavares, J. (2003). Does foreign aid corrupt?. *Economics Letters, 79*, 99–106.

Temple, J. (2010). Aid and conditionality. *Handbook of Development Economics, 5*, 4416–4522.

van de Walle, D., & Mu, R. (2007). Fungibility and the flypaper effect of project aid: Micro-evidence for Vietnam. *Journal of Development Economics, 84*, 667–685.

Woolcock, M. (2009). Toward a plurality of methods in project evaluation: A contextualized approach to understanding impact trajectories and efficacy. *Journal of Development Effectiveness, 1*, 1–14.

Woolcock, M. (2011). The importance of time and trajectories in understanding project effectiveness. Development impact blog. URL: http://www.samea.org.za/News-52.phtml.

APPENDIX A. ADDITIONAL MATERIAL

This appendix provides a more technical exposition of the link between Figure 1 and feasible empirical approaches for identifying the relation between aid and either intermediate or final outcomes. The starting point, following Pearl (2009), is to note that as a DAG the same graph has a corresponding non-parametric structural equation representation. Assuming the functions are autonomous, the figure corresponds to Eqns. (A.1)–(A.4) below.

$$v = f_v(T, u_v) \tag{A.1}$$

$$a = f_a(v, T, u_a) \tag{A.2}$$

$$x_m = f_x(w, T, u_{x_m}) \tag{A.3}$$

$$y = f_y(T, X, u_y) \tag{A.4}$$

and where it is additionally assumed:

$$E[u_v u_j] = 0 \ \forall \ j \in J = \{a, x_1, x_2 \ldots, x_M, y\}$$

$$E[u_j u_k] \neq 0 \ \forall \ j, \ k \in J.$$

These are general expressions from which empirical specifications used in the applied growth literature can be derived as special cases (e.g., using additive errors). For instance, defining y in Eqn. (A.4) as aggregate GDP growth, this maps directly to the standard equation specifying growth as a function of initial and steady state income Mankiw et al. (1992). That is, T contains initial income as well as various fixed factors that affect long-run productivity, while $X = (x_1, x_2, \ldots, x_M)$ contains proximate time-varying factors, such as the rate of accumulation of human and physical capital that also affect steady state income.

The key assumption of the system is that $(u_v|T) \perp\!\!\!\perp (u_a, u_x, u_y|T)$. This means that the effect of aid on a chosen element x_m, or on y, can be recovered indirectly as the ratio of causal effects due to v. As noted by Balke and Pearl (1997), however, this is only feasible with the additional assumption that the underlying functional forms are linear, such that error terms are additive. This can be seen algebraically by taking the reduced form associated with Eqns. (A.3) and (A.4):

$$\begin{aligned} y &= f_y(T, X, u_y) \\ &= f_y([f_x(a, T, u_{x_1}), \ldots, f_x(a, T, u_{x_M})], T, u_y) \\ &= \sum_{m=1}^{M} \alpha[\beta_{1m} a + T'\beta_{2m} + u_{x_m}] + T'\gamma + u_y \\ &= \tilde{\lambda} a + T'\tilde{\mu} + \tilde{u} \end{aligned} \tag{A.5}$$

where the tilde superscripts denote aggregated parameters. Multiplying (A.5) through by v, taking expectations and rearranging, yields an instrumental variables estimand: $\delta y/\delta a = \mathrm{Cov}(v, y|T)/\mathrm{Cov}(v, a|T)$. The reduced form effect of aid on, say, x_1 can be estimated analogously: $\delta x_1/\delta a = \mathrm{Cov}(v, x_1|T)/\mathrm{Cov}(v, a|T)$.

APPENDIX B. SUMMARY STATISTICS AND VARIABLE SOURCES

The table below summarizes the variables used in the analysis, the measurement scale employed, and the original data sources (with source-variable reference code where available). Please see the notes at the end of the table for further details.

	N	Median	Mean	St. dev	Scale	Source	Reference code
Aid variables							
Aid/GDP	78	2.28	3.49	3.77	[A]	[1]	–
Aid per capita	78	26.64	36.78	44.74	[A]	[1]	–
Generated aid instrument	78	62.52	62.26	12.13	[A]	[2]	–
Final outcomes							
Real GDP growth per capita	78	1.68	1.73	1.79	[A]	[3]	rgdpchg
Agriculture, value added (% GDP)	76	20.65	22.59	13.42	[A]	[4]	NV.AGR.TOTL.ZS
Industry, value added (% GDP)	76	29.63	29.71	9.90	[A]	[4]	NV.IND.TOTL.ZS
Services, etc., value added (% GDP)	76	48.72	47.71	9.66	[A]	[4]	NV.SRV.TETC.ZS
Poverty headcount at $2 a day	64	43.30	45.39	29.72	[D]	[4]	SI.POV.2DAY
Poverty headcount at $1.25 a day	64	21.65	28.35	24.03	[D]	[4]	SI.POV.DDAY
Gini index	65	44.19	44.46	7.87	[D]	[4]	SI.POV.GINI
Intermediate outcomes							
Investment in real GDP	78	17.18	18.53	10.30	[A]	[3]	ki
Private consumption in real GDP	78	68.36	70.02	19.95	[A]	[3]	kc
Government consumption in real GDP	78	16.60	18.04	8.41	[A]	[3]	kg
Revenue, excluding grants (% GDP)	69	19.95	21.66	8.82	[A]	[4]	GC.REV.XGRT.GD.ZS
Health expend., public (% GDP)	78	2.67	2.89	1.39	[A]	[4]	SH.XPD.PUBL.ZS
Education expend., public (% GDP)	76	3.76	3.97	1.50	[A]	[4]	SE.XPD.TOTL.GD.ZS
Military expenditure (% GDP)	77	1.91	2.34	1.70	[A]	[4]	MS.MIL.XPND.GD.ZS
Life expectancy at birth, total (years)	78	60.96	59.00	9.50	[B]	[4]	SP.DYN.LE00.IN
Infant mortality rate	75	67.48	71.78	35.38	[B]	[4]	SP.DYN.IMRT.IN
Death rate, crude (per 1,000 people)	78	11.05	11.92	4.62	[B]	[4]	SP.DYN.CDRT.IN
Fertility rate (births/woman)	77	4.66	4.68	1.38	[B]	[4]	SP.DYN.TFRT.IN
Consumer price inflation (%)	77	10.31	52.69	150.30	[A]	[4]	FP.CPI.TOTL.ZG
Real interest rate (%)	77	6.58	7.00	8.78	[A]	[4]	FR.INR.RINR
Domestic credit to private sector (% GDP)	78	23.74	29.37	21.40	[A]	[4]	FS.AST.PRVT.GD.ZS
Av. years total schooling, 15+	72	4.89	5.01	2.10	[B]	[5]	BAR.SCHL.15UP
Av. years primary schooling, 15+	72	3.64	3.49	1.42	[B]	[5]	BAR.PRM.SCHL.15UP

(*continued on next page*)

APPENDIX B—(*continued*)

	N	Median	Mean	St. dev	Scale	Source	Reference code
Av. years secondary schooling, 15+	72	1.35	1.37	0.72	[B]	[5]	BAR.SEC.SCHL.15UP
Control variables							
Income per capita	78	7.88	7.84	0.79	[C]	[6]	–
Sachs–Warner trade policy index	78	0.32	0.32	0.29	[C]	[6]	–
Life expectancy	78	51.99	52.88	9.77	[C]	[6]	–
Geography	78	−1.00	−0.55	0.77	[C]	[6]	–
Ethnic fractionalization	78	0.54	0.47	0.29	[C]	[6]	–
Primary eduction enrollment rate	78	0.67	0.65	0.29	[C]	[7]	p60
Coastal population density	78	30.36	101.25	358.82	[C]	[7]	dens65c
Malaria prevalence index	78	0.54	0.51	0.43	[C]	[7]	sa_mr
Price of investment goods	78	85.83	93.68	62.20	[C]	[7]	iprice1
Civil liberties	78	0.33	0.41	0.27	[C]	[7]	civ72
Air distance to major cities	78	8.47	8.41	0.50	[C]	[7]	airdist
Oil producer	78	0.00	0.36	0.48	[C]	[7]	oildummy
Landlocked	78	0.00	0.21	0.41	[C]	[7]	landlock
Prevalence of HIV (% of pop 15–49)	68	0.80	3.12	5.70	[D]	[4]	SH.DYN.AIDS.ZS

Scales: [A] full period mean (1970–2007); [B] average of earliest start and latest end values; [C] initial value (1960s or early 1970s if the former unavailable); [D] latest end value only;

Sources: [1] Authors' calculations from OECD-DAC (www.oecd.org/dac/stats/idsonline; downloaded May 2009); [2] authors' estimates based on the method set out in Arndt *et al.* (2010a, 2010b), using updated and cleaned OECD-DAC dataset; [3] Penn World Tables v6.3 (http://pwt.econ.upenn.edu) [4] World Bank, World Development Indicators and Global Development Finance (http://data.worldbank.org/data-catalog; downloaded April 20, 2010); [5] World Bank, Education Statistics (http://data.worldbank.org/data-catalog; downloaded April 20, 2010); [6] Arndt *et al.* (2010a, 2010b); [7] Sala-i-Martin *et al.* (2004).