CONTEMPORARY ECONOMIC POLICY

WEAI

AID, POLICIES, AND GROWTH: WHY SO MUCH CONFUSION?

SHAOMENG JIA and CLAUDIA R. WILLIAMSON*

We revisit the highly debated aid-policy-growth association. Our results overturn Burnside and Dollar's original findings by simply using new data over the same countries and years. Marginal effects from the extended sample (1962–2013) provide weak evidence that aid can promote growth in the presence of good policies. Post-Cold War (1990–2013) analysis, however, reveals that aid can decrease growth at any level of policy. The overwhelming majority of the results suggest aid conditional on policy is ineffective. This debate continues because the results are highly sensitive to country-year selection, choice of methodology, measurement of institutional quality, and growth rate measurement. (*JEL* F35, O23, O40)

I. INTRODUCTION

Foreign aid effectiveness is continuously debated in development economics with many scholars conceding that aid has not achieved its intended results. The United Nations Sustainable Development Goals (SDGs), however, call for an increase in development assistance to developing countries stating that foreign aid is "fundamental to equitable progress for all" (United Nations 2016, 44).

Given this tension between scholarly empirical evidence and the development community's call to increase foreign aid, donors are to be more selective in aid allocations to increase aid effectiveness (OECD 2005, 2008). Selectivity includes providing aid to countries that minimize corruption, maintain democracy over an autocratic regime, and support basic human rights such as civil, economic, and political liberties (Achta, Toman, and Rainer 2015; Dollar and Levin 2006; Easterly and Pfutze 2008; Easterly and Williamson 2011; Hagen 2015). In essence, aid can be made more effective if it is given to better-governed countries most in need.

We ask, does aid selectivity matter? Selectivity essentially argues that aid's effect is conditional on those margins in which donors are to be selective over, that is, the policy and institutional environment. In order to address this question, we reexamine the original selectivity paper: Burnside and Dollar's (2000) influential study on the aid-policy-growth connection. Burnside and Dollar (henceforth BD) argue that aid's effect is conditional on the economic environment in a recipient country. They empirically demonstrate that aid can positively influence

ABBREVIATIONS

BD: Burnside and Dollar
DFE: Dynamic Fixed Effects
EDA: Effective Development Assistance
EFW: Economic Freedom Index
ELR: Easterly, Levine and Roodman
FD: First-difference
FE: Fixed Effect
GDP: Gross Domestic Product
GMM: Generalized Method of Moments
HDI: Human Development Index
ICRG: International Country Risk Guide
IDA: International Development Association
ODA: Official Development Assistance
OLS: Ordinary Least Squares
PMG: Pooled Mean Group
PPP: GDP Per Capita
PWT: Penn World Tables
SDG: Sustainable Development Goal
2SLS: Two-Stage Least Squares
WDI: World Development Indicators

*We extend a special thank you to Craig Burnside for helpful comments and suggestions. We also thank two anonymous reviewers, William Easterly, Peter. T. Leeson, participants at the 2016 ASSA meetings and the 2015 Southern Economic Association Conference.

Jia: Visiting Assistant Professor, Department of Economics, Accounting, and Business, Muskingum University, New Concord, OH 43762, Phone 740-826-8270, Fax 740-826-8357, E-mail smjia@hotmail.com

Williamson: Associate Professor of Economics, Department of Finance and Economics, Mississippi State University, Mississippi State, MS 39762, Phone 724-255-8627, Fax 662-325-1977, E-mail claudia.williamson@msstate.edu

Contemporary Economic Policy (ISSN 1465-7287)
Vol. 37, No. 4, October 2019, 577–599
Online Early publication January 10, 2019

doi:10.1111/coep.12418

growth in healthy policy environments, sparking one of the most debated topics in development economics and among policymakers (Brumm 2003; Dalgaard 2004; Hansen and Tarp 2000; Leeson 2008). Easterly, Levine and Roodman (2004) (henceforth ELR), using the exact methodology over a larger data set, overturn BD's findings, weakening the significance of the aid-policy-growth association.

In this paper, we study whether aid's effectiveness depends on donor selectivity by empirically examining the aid-policy-growth debate. We do so by first revisiting the original works of BD and ELR with updated data. ELR overturn BD's findings with four additional years (1994–1997) and six additional countries, stating that this debate suffers from "a long and inconclusive literature that was hampered by limited data availability" (Easterly, Levine, and Roodman 2004, 774). A concern is that missing data may generate biased results (Breitwieser and Wick 2016). Our data set includes an additional 28 years (1962–1969 and 1994–2013) and six countries over BD, and 20 years (1962–1969 and 1998–2013) and nine countries over ELR, which almost doubles BD's number of observations and increases ELR's sample by up to 70%.[1] With the additional expansion of the data set, it is possible that ELR's results are overturned or new findings are discovered.

To test this possibility, we initially do not deviate from the specifications and methodology of BD and ELR, using ordinary least squares (OLS) and two-stage least squares (2SLS) as the main estimation techniques. We replicate the findings from both BD and ELR with updated data using multiple country and year specifications. We find that BD's results are not robust to the updated data. Simply using new data for BD's same countries and years (1970–1993), we do not find any significant aid/policy interaction terms. Furthermore, we show that BD's findings are associated with observations unique to their sample but that are unavailable to the updated sample. Our ELR replications for the 1970–1997 period, however, are highly consistent with ELR's results—no significant interactions of aid and policy are found.

We further test BD and ELR's specifications with two additional samples: an extended sample from 1962 to 2013 and a post-Cold War subsample from 1990 to 2013. In the extended sample, OLS regressions provide some support for BD's conclusion that aid may increase growth when a recipient country has top policy scores. The insignificant results from the 2SLS specifications and the overwhelming insignificant marginal effects make it difficult to conclude that aid is effective at increasing growth. The post-Cold War results report mostly insignificant marginal effects, consistent with the results from the 1962–2013 period. We do not find one positive and significant marginal effect; however, we do find several *negative* and significant marginal effects from 2SLS specifications, occurring at all levels of policy and mostly in lower-income countries. The post-Cold War sample suggests that aid can decrease growth regardless of the policy environment. Collectively, the two additional samples reveal that if aid significantly affects growth that effect is negative, not positive.

Next, we move beyond BD's original framework to take advantage of improvements in estimation techniques and advances in measurement of data. First, we employ a more comprehensive measure of economic institutions, the economic freedom of the world index (Gwartney, Lawson, and Hall 2016). We find that in countries with low economic freedom scores, aid may decrease growth, but there is some evidence that aid may increase growth in economically free countries. We also find that using Penn World Tables (PWT) measure of growth instead of World Development Indicators (WDI) growth rates results in more significant aid/policy interaction term coefficients. These results illustrate another reason why BD, using PWT, found significance and ELR, using WDI, did not. The measurement of aid, however, does not appear to matter as similar results are found using alternatively constructed aid measures.

We also test the sensitivity of estimation selection by switching from OLS and 2SLS estimation techniques to using fixed effects (FE), first-difference (FD), system generalized method of moments (GMM), pooled mean group (PMG), and dynamic fixed effects (DFE). By using superior estimation techniques, in particular controlling for endogeneity with GMM and PMG, the evidence supporting aid contributing to growth in a good policy environment is minimal. This suggests that BD's original results are also contingent on estimation selection.

Lastly, we examine aid's impact on measures of social well-being, including income inequality, the human development index, poverty rate, and

1. For detailed differences in observations and countries between our dataset and BD and ELR's 2SLS samples, see Tables S1 and S2, Supporting Information.

unemployment rate. We do not find any evidence that aid can improve social well-being, even when conditioned on a healthy policy environment.

Overall, we find that under limited scenarios, aid may promote growth in the presence of good policies, but the majority of results suggest aid's effect on growth is not conditional on the policy environment. There is additional evidence that aid can harm growth in poor policy environments. Our tests illustrate that differences in results are mostly driven by country-year selection, choice of methodology, and use of PWT or WDI to measure growth.

These findings reflect the sensitive feature of the aid-growth literature, making it difficult to compare across studies even when keeping the methodology unchanged. This provides support to the importance of replicating major studies as new data and estimation techniques become available (Burman, Reed, and Alm 2010; Easley, Madden, and Gray 2013; Evanschitzky and Armstrong 2013; McCullough, McGeary, and Harrison 2008; Mekasha and Tarp 2018). Replication avoids measurement error and disagreement over method selection caused by "… usual limitation of choosing a specification without clear guidance from the theory" (Easterly, Levine, and Roodman 2004, 774).

Our work contributes to the long-standing aid-policy-growth academic debate and reminds policymakers that simply providing aid to countries identified as having "good" policies may not create a "quick" growth fix. Even if donors become more selective and give aid to better governed countries, our work suggests that aid is unlikely to achieve economic growth or other goals outlined by the SDGs. In fact, our analysis over the time period when donors became more vocal to utilize aid as a poverty reduction tool, the post-Cold War period, indicates that aid may be hindering economic and social development. This suggests that policymakers should hesitate when designing development policy based on the assumption that aid can achieve development goals.

II. AID-POLICY-GROWTH LITERATURE REVIEW

Why should aid's impact on growth depend on the policy environment? BD observed the importance of sound economic policies for growth in developing countries. In addition, empirical papers emerged documenting the lack of association between aid and growth in recipient countries (Boone 1996). BD argued that aid could boost growth by working through the recipient countries' policies in the same manner that policies influence growth in absence of aid (Burnside and Dollar 2000, 847). As summarized by Leeson (2008, 48), BD's "highly influential analysis has an intuitively appealing bottom line: aid can encourage economic growth in countries that pursue 'good' economic policies. Elsewhere, aid is essentially wasted dollars and has no impact on economic growth."

Since the release of both seminal articles, many scholars contribute to the aid-growth debate. For example, BD have accumulated over 5,300 citations and ELR are cited nearly 1,500 times.[2] After many studies, unfortunately, the impact of foreign aid remains inconclusive and the debate continues.

Most studies are in line with pro-ELR conclusions. For example, both Hansen and Tarp (2000) and Dalgaard and Hansen (2001) illustrate that BD's result relies on the exclusion of five outliers; thus, if the outliers are included, then BD's results do not hold. In addition, Brumm (2003) and Dalgaard, Hansen, and Tarp (2004) find that aid's impact on growth does not depend on a recipient's policy quality. In fact, aid can negatively impact growth under good policies or promote growth with bad policies. Furthermore, Dalgaard and Hansen (2001) conclude that aid and policies are substitutes where a healthy policy environment might reduce the effectiveness of foreign aid. Tan (2009, 5) also concludes that the aid and policy interaction "leads to lower growth in the long-run."

Moreover, additional works including Guillaumont and Chauvet (2001), Hudson and Mosley (2001), Hansen and Tarp (2001), Lensink and White (2001), Lu and Ram (2001), Easterly (2003), Islam (2005), Rajan and Subramanian (2008), Doucouliagos and Paldam (2010), Tashrifov (2012), Chatelain and Ralf (2014), and Dreher and Langlotz (2015) find no evidence supporting aid's positive impact on growth.

On the contrary, a number of articles support BD, concluding that aid does work effectively in a good policy environment (Ali and Isse 2005; Alvi, Mukherjee, and Shukralla 2008; Burnside and Dollar 2004; Collier and Dehn 2001; Verschoor and Kalwij 2006).[3] Contributing to the ambiguity of this debate, Dayton-Johnson

2. The citation count numbers are collected as of October 12, 2018 from Google scholar.

3. Burnside and Dollar (2004) and Dalgaard, Hansen, and Tarp (2004) switch from policy and examine institutional quality.

and Hoddinott (2003) and Kohama, Sawada, and Kono (2003) find mixed results. Ram (2003) finds positive and significant interactions of policy and bilateral aid but offsetting negative interactions with multilateral aid and policy.

Roodman (2007a) argues these inconsistencies and fragile results are driven by arbitrariness in specification choices and samples. Similarly, Roodman (2007b) concludes that effects of aid on growth cannot be detected with limited and noisy data. Qian (2015) further discusses differences in results associated with measurement issues due partly to the heterogeneous nature of aid.

With the exception of ELR, these follow-up studies carry out variations of BD's original framework using alternative approaches, including different measures of foreign aid and policies, alternative model specifications, additional control variables and instruments, as well as different country samples and time periods. These changes in methodology may partly explain the ambiguity of the findings. For instance, Lu and Ram (2001) find that policy has no significant influence on aid's effect on growth once country-FEs are included. Hansen and Tarp (2001) switch to a GMM method and find that aid increases growth via an investment channel but not through a policy channel. Rajan and Subramanian (2008) introduce a measure of bilateral aid and test for conditionality of both policy and geographical environments concluding that aid is ineffective.

Several recent and innovative papers analyzing the aid-growth association acknowledge the critical importance of BD and continue using the original BD specification (Clemens et al. 2012; Dreher, Eichenauer, and Gehring 2016; Dreher and Langlotz 2015; Juselius, Møller, and Tarp 2014; Tan 2009).[4] For example, Tan (2009) introduces a PMG estimator to capture the long-run homogeneous relation and finds that aid*policy has no significant impact on aid. However, using a well-specified cointegrated vector autoregressive method, Juselius, Møller, and Tarp (2014) find a positive long-run impact of aid on investment and gross domestic product (GDP). Dreher and Langlotz (2015) and Dreher, Eichenauer, and Gehring (2016) examine aid's effect conditional on donor factors, such as political motivations and donor government fractionalization. These two studies find no significant effect of aid on growth, except in good policy countries where aid can reduce growth.

4. These recent studies utilize updated data from Minasyan (2016).

The current paper distinguishes itself by focusing on aid's dependency on policy within the original BD framework. We expand the sample to include up to 52 years and utilize multiple advanced estimation techniques within the same data set. We contribute to this body of work by: (1) demonstrating that BD's results are contingent on country-year sample, use of PWT growth measure, and choice of estimation techniques; (2) providing evidence that aid can be harmful at any level of policy and in economically unfree countries; and (3) illustrating why the aid-growth debate continues. These results are robust to the use of sophisticated estimation methods. Combined, our results suggest that differences in the aid-policy-growth debate stem from sample biases, measurement issues, and choice of estimation techniques.

III. EMPIRICAL METHODOLOGY

Initially, we follow both BD's and ELR's methodology and data sources to investigate the relation between aid, policy, and growth.[5] BD employ methods of pooled OLS and 2SLS using a panel data set with 4-year averages. As shown in the following equation, BD's growth regression includes aid/GDP, a policy index, an aid*policy interaction term, log initial GDP, ethnic fractionalization, political assassinations, a fractionalization*assassinations interaction term, institutional quality, financial depth (M2/GDP lagged), regional dummies, and time dummies. In some specifications, an aid^2*policy term is included.

The 2SLS expresses as the following equations:

$$(1)\qquad a_{it} = y_{it}\gamma_y + p'_{it}\gamma_p + z'_{it}\gamma_z + a_t + \varepsilon^a_{it}$$

$$g_{it} = y_{it}\beta_y + \alpha_{it}\beta_a + p'_{it}\beta_p + \alpha_{it}p'_{it}\beta_1$$

$$(2)\qquad + z'_{it}\beta_z + g_t + \varepsilon^g_{it}$$

where i denotes countries, t denotes period, g_{it} is per capita real GDP growth, y_{it} is natural logarithm of per capita real GDP, a_{it} is international aid received relative to its total GDP, g_t and a_t are fixed-time effects, z'_{it} is a vector of other exogenous variables, p_{it} is the policy index vector constructed by BD.

The second stage is represented by Equation (2). The instruments, denoted by z'_{it}, include

5. BD and ELR have slightly different model specifications as they define regional country dummies and low income countries differently. See Tables S1 and S6.

dummies for Egypt, Franc Zone, and Central American countries, lagged arms imports and its interaction term with the policy index, population, two interaction terms using population and squared population with policy index, initial GDP per capita and its interaction with the policy index.

The construction of the policy index, p_{it}, calculates the weights of different policies to the growth regression. To reconstruct the policy index, we run the growth regression (1) excluding aid and aid*policy but include inflation, budget surplus, and the Sachs-Warner openness index (Sachs and Warner 1995). These coefficients create a beta policy index, policy^0. The constant is calculated by differencing the mean of the GDP growth rate and the mean of the beta policy index. Lastly, the constant is added to the beta policy index creating the new policy index.

This process can be summarized by the following steps.

(i) Run Equation (1) above without aid and aid*policy terms, and collect the policy coefficients:

$$g_{it} = y_{it}\beta_y + p'_{it}\beta_p + z'_{it}\beta_z + g_t + \varepsilon^g_{it}. \tag{3}$$

(ii) Construct a variable Policy^0 (denoted by p^0_{it}) with coefficients collected from step (i) and calculate the mean of Policy^0 (denoted by $\overline{p}$):

$$p^0_{it} = \beta_{bs}BudSurplus + \beta_{in}Inflation + \beta_{op}Openess; \text{ and get } \overline{p}. \tag{4}$$

(iii) Calculate the constant of the policy index (*Constant*), which is the difference between the mean of GDP growth rate and the mean of Policy^0:

$$Constant = \overline{g} - \overline{p}. \tag{5}$$

(iv) Add the constant term to p^0_{it}, and get the policy index:

$$p_{it} = p^0_{it} + Constant. \tag{6}$$

Our newly constructed policy index is highly correlated with BD/ELR, with pair-wise correlation coefficients up to .92.

In order to reconstruct the database, we gather variables from original sources in BD and ELR and expand the data set from 1962 to 2013 and up to 65 countries (based on ELR, OLS specification, before excluding outliers). Table S3 contains the specific source and method of construction for each variable, as well as the correlations between the new data, BD and ELR. Summary statistics are provided in Table S4.

Given the length of time between our study and BD and ELR, some of the variables are discontinued. For those variables, we extrapolate based on ELR's data and methodology by filling in the missing data with the closest observation. In a later section, we test the sensitivity of ELR's methodology for missing data by using alternative measurements of growth, aid, and institutions.

To measure aid, many current studies use OECD's official development assistance (ODA); however, BD measures aid in terms of effective development assistance (EDA)[6] over GDP. To calculate EDA, BD regress EDA on ODA, retain the regression coefficient, and multiply it with new ODA data. To update this measure, we extrapolate EDA with the same methodology. The pair-wise correlations between our newly extrapolated EDA and BD/ELR's EDA are 0.71/0.74. The list-wise correlation with BD increases to 0.84.

According to BD, aid positively impacts growth in countries with good policy environments. Thus, to support BD, we should observe positive and significant aid*policy interaction terms. Furthermore, the marginal effects of aid should be positive and significant at least in high policy countries. If we find, however, insignificant aid*policy interaction terms and marginal effects, the results lend support to ELR. In addition, the results support ELR if aid*policy is significant but the marginal effects are insignificant or negative and significant in high policy countries. Lastly, it is possible that the marginal effects at low policy scores are negative and significant, indicating that aid is detrimental to growth in poor policy countries. This would indirectly support both BD and ELR, as BD and ELR's debate assumes aid is either helpful or neutral; hence, only the positive influence of aid is hypothesized with the negative possibility never being explicitly stated.

IV. REPLICATION AND EXTENSION RESULTS

First, we test the findings from both BD and ELR with updated data using multiple specifications: (1) BD countries and BD years (1970–1993), full country sample and BD years; (2) ELR countries and ELR years (1970–1997),

6. EDA is originally from Chang, Fernandez-Arias, and Serven (1999). This paper does not intend to recalculate Chang et al.'s work, given that data availability has changed; instead, we follow ELR's indirect extrapolation method.

full country sample and ELR years; (3) extended years (1962–2013) with BD countries, ELR countries, and full country sample; (4) post-Cold War (1990–2013) with BD countries, ELR countries, and full country sample. We also report the original findings from BD and ELR,[7] marginal effects at different policy levels, and test different subsets of BD. Bootstrapped standard errors are reported in all models.

A. Replicating BD and ELR, 1970–1993/97

BD and ELR test their specifications including and excluding outliers. We follow ELR and use the Hadi method to test for outliers and exclude those observations indicated as outliers. Table S5 reports outliers for each sample. Similar to ELR, we report the results corresponding to the OLS and 2SLS specifications from BD regressions 4 (all countries) and 7 (lower-income countries), which includes the outliers and an aid^2*policy term. We also report the findings for OLS and 2SLS for BD regressions 5 (all countries) and 8 (lower-income countries) excluding the outliers and dropping the aid^2*policy term.[8]

In Table 1, Panel A, we test the model under the same time period as BD with newly collected data. BD's original results show positive and significant coefficients on the aid*policy interaction term in six of eight specifications.[9] Once we replicate BD's exact specification with updated data, with only BD countries or all countries in the new data, none of the interactions are significant. The update of ELR's specifications also finds no significant interaction terms, as shown in Table 1, Panel B, supporting ELR's original results. Overall, we find that BD's results are not robust to the updated data, whereas ELR's replications are highly consistent with their original findings.

The most striking finding from this replication, however, is that BD's result disappears by simply updating the data. While ELR show the results disappear by extending the sample, we illustrate it by utilizing updated data over the same country-year sample. In order to understand what is driving these differences, we compare our sample to BD's and find that there are quite a few unique observations belonging to each dataset. The country and year selection remains the same; however, there are country-year pairings unique to each sample.[10] Thus, we conjecture these different observations might explain the differences in findings.

To test the sensitivity of these differences, in Table 2, we re-estimate the main tests with a number of BD country subsamples from 1970 to 1993. As shown in row 1, "Intersection of datasets," we exclude all unique observations from both the new data set and BD's original data set, creating an intersection subsample. We observe only two significant aid*policy interaction term coefficients at the 10% level: one negative and one positive. In addition, we find one positive and weakly significant marginal effect in high policy countries. Compared to BD's original data, the intersection set has 65 fewer observations. This suggests that BD's result is dependent on observations unique to their sample but unavailable with updated data.

Consistent with this conjecture, Hansen and Tarp (2000, 393) show that BD's results depend on the exclusion of five "outliers."[11] Dalgaard and Hansen (2001, 32–33) identify these five observations as "leverage points" due to their above-average influence on the fitted values but not classifying as econometric outliers.[12] To test this argument, we present a second specification in Table 2, row 2, using the updated sample and BD countries but dropping the five "outliers."[13]

7. We replicate both BD and ELR's works with their original data sets. Our replication matches their original results.

8. We follow BD to define lower-income countries as a country with real GDP per capita below $1,900 constant (1985) U.S. dollars in year 1970.

9. BD and ELR use different significance levels. For coefficients with p values greater than .05 but less than .10, it is considered significant under BD but not significant under ELR. This is one potential reason why ELR found fewer significant interaction terms than BD.

10. Comparing our "New data, BD countries" sample with BD's sample: under OLS, we find 32 and 24 unique observations, respectively; under 2SLS, there are 21 and 65, respectively. This also occurs in ELR's work. There are 19 and 47 unique observations across BD and ELR's 1970–1993 samples, respectively. Some observations were available in the 1990s but are no longer reported. In addition, data have become available that were not previously reported. See Table A2 in ELR for more information on their sample differences. Refer to Table S1 for more detail on the observation comparisons of BD, ELR and our new data, full sample.

11. The five "outliers" are Nicaragua (1986–1989, 1990–1993), Gambia (1986–1989, 1990–1993), and Guyana (1990–1993).

12. Furthermore, Chatelain and Ralf (2014, 93) point out that these observations affect the validity of White standard errors used in BD's model. White heteroscedasticity consistent standard errors are not useful when heteroscedasticity is driven by large outliers, such as those present in BD's work.

13. Hansen and Tarp (2000) argue that the five BD "outliers" are not beyond the three standard errors band; hence, they should not be excluded as outliers. Dalgaard and Hansen (2001) also indicate that these are not outliers. Our results also support this argument. Refer to Table S5 for details of outliers dropped for each of the models.

TABLE 1
Replication with New Data 1970–93/97, BD and ELR Regressions 4, 7, 5, and 8

		Outliers Included				Hadi Method, Outliers Excluded			
		All Countries		Lower Income Countries		All Countries		Lower Income Countries	
		4/OLS	4/2SLS	7/OLS	7/2SLS	5/OLS	5/2SLS	8/OLS	8/2SLS
Panel A: BD 1970–1993, coefficients for aid*policy and aid^2*policy term									
Aid*policy	BD original	0.20**	0.37	0.27**	0.43	0.19**	0.18*	0.26**	0.25**
		(0.09)	(0.33)	(0.12)	(0.49)	(0.07)	(0.10)	(0.08)	(0.01)
	New data, BD countries	0.09	−0.11	0.08	−0.38	0.11	0.13	0.03	0.02
		(0.10)	(0.31)	(0.13)	(0.48)	(0.07)	(0.12)	(0.09)	(0.15)
	New data, full sample	0.09	−0.02	0.08	−0.34	0.11	0.13	0.02	0.01
		(0.10)	(0.31)	(0.12)	(0.52)	(0.07)	(0.11)	(0.10)	(0.15)
Aid^2*policy	BD original	−0.02*	−0.04	−0.02**	−0.04				
		(0.01)	(0.04)	(0.01)	(0.05)				
	New data, BD countries	−0.01	0.03	−0.01	0.06				
		(0.01)	(0.05)	(0.01)	(0.06)				
	New data, full sample	−0.01	0.02	−0.01	0.06				
		(0.01)	(0.05)	(0.01)	(0.07)				
Number of observations	BD original	275	275	189	189	270	270	182	184
	New data, BD countries	283	231	188	152	277	227	183	149
	New data, full sample	300	243	192	156	294	239	187	153
Panel B: ELR 1970–1997, coefficients for aid*policy and aid^2*policy term									
Aid*policy	ELR original	−0.14		−0.27		−0.15	0.01	−0.20	−0.20
		(1.31)		(1.89)		(1.09)	(0.05)	(1.26)	(0.65)
	New data, ELR countries	0.04	−0.19	0.09	−0.32	0.02	0.06	0.01	0.08
		(0.10)	(0.25)	(0.12)	(0.42)	(0.06)	(0.10)	(0.08)	(0.14)
	New data, full sample	0.05	0.12	0.09	0.05	0.08	0.15	0.12	0.17
		(0.08)	(0.28)	(0.10)	(0.40)	(0.07)	(0.12)	(0.08)	(0.17)
Aid^2*policy	ELR original	0.03**		0.03**					
		(2.25)		(2.35)					
	New data, ELR countries	−0.01	0.04	−0.01	0.05				
		(0.01)	(0.04)	(0.01)	(0.06)				
	New data, full sample	−0.00	0.00	−0.01	0.02				
		(0.01)	(0.05)	(0.01)	(0.05)				
Number of observations	ELR original	356	356	244	244	345	345	236	236
	New data, ELR countries	358	296	239	195	352	292	234	192
	New data, full sample	390	315	257	205	385	312	253	203

Notes: Bootstrap standard errors are reported in parentheses. Time fixed effects are included in all regressions, but not reported in the table. Each specification includes a constant term, measure of aid/GDP, a policy index, an aid*policy interaction term, log initial GDP, ethnic fractionalization, political assassinations, a fractionalization*assassinations interaction term, a measure of institutional quality, and a measure of financial depth (M2/GDP lagged), regional dummies for Sub-Saharan Africa and fast-growing East Asian countries. BD and ELR specifications differ in their definitions of regional dummies and low income (see Table S6). See Table S3 for detailed variable description. Regression numbers are matched with BD/ ELR original works.

*$p < .1$.
**$p < .05$.
***$p < .01$.

We do not find any significant coefficients for the interaction terms or the marginal effects, supporting the findings in Table 1.[14]

Another possible explanation driving the differences is the measurement of two key variables of interest, the updated measures of aid and the policy index. Not only are these the main variables of interest, but they are also less consistent across data sets.

We first retest the models substituting our updated BD aid with the original BD aid measure. To minimize influence from unique observations, we use the intersection sample set as in row 1. As shown in row 3, we find two coefficients with positive and significant interactions at the 10% level. There are three positive and significant marginal effects from low-income countries with policy at the 90th percentile. This indicates that differences in our updated work and BD's original results are partly driven by changes in aid measurement.[15]

Addressing this concern, Dalgaard and Hansen (2001) provide support for the validity of ELR's method of extrapolating EDA by arguing that the difference between EDA and ODA is a simple transformation, with a correlation as high

14. Our "New data, BD countries" sample includes two of the five "outliers"—Gambia (1986–1989, 1990–1993).

15. Rows (1) and (3) have slight observation differences in some of the models. We retest dropping the extra observations, and the result holds. This is not tabulated to save space.

TABLE 2
BD Subsets with New Data, 1970–1993, BD Regressions 4, 7, 5, and 8

			Outliers Included				Hadi Method, Outliers Excluded			
			All Countries		Lower Income		All Countries		Lower Income	
			4/OLS	4/2SLS	7/OLS	7/2SLS	5/OLS	5/2SLS	8/OLS	8/2SLS
Panel A: Coefficients for aid*policy term										
Aid*policy	1	Intersection of data sets	0.06	0.08	0.20	−0.20	−0.20	−0.32*	0.26*	0.26
			(0.16)	(0.37)	(0.21)	(0.47)	(0.14)	(0.19)	(0.14)	(0.22)
	Number of observations		210	210	133	133	201	201	131	131
	2	BD countries, drop BD outliers	0.05	−0.06	0.05	−0.33	−0.00	−0.03	0.04	0.02
			(0.12)	(0.32)	(0.15)	(0.46)	(0.09)	(0.17)	(0.11)	(0.20)
	Number of observations		284	230	186	151	272	222	180	146
	3	Intersection of data sets, BD aid	0.24	−0.43	0.49*	0.10	−0.19	−0.27	0.31**	0.31
			(0.22)	(0.46)	(0.29)	(0.55)	(0.16)	(0.23)	(0.14)	(0.23)
	Number of observations		210	210	133	133	204	204	132	132
	4	Intersection of data sets, BD policy	−0.03	−0.03	0.11	−0.09	−0.11	−0.18	0.16	0.13
			(0.18)	(0.46)	(0.21)	(0.48)	(0.12)	(0.17)	(0.10)	(0.18)
	Number of observations		210	210	133	133	203	203	130	130
Panel B: Marginal effects for aid										
Marginal effects	1	Policy at 10th percentile	−0.06	0.02	0.01	−0.10	0.18	0.21	−0.09	−0.04
			(0.16)	(0.38)	(0.15)	(0.31)	(0.16)	(0.29)	(0.16)	(0.29)
		Policy at means	0.00	0.08	0.13	−0.10	0.03	−0.03	0.10	0.15
			(0.17)	(0.42)	(0.18)	(0.32)	(0.12)	(0.25)	(0.13)	(0.22)
		Policy at 90th percentile	0.13	0.21	0.42	−0.07	−0.33	−0.61	0.54**	0.59
			(0.32)	(0.64)	(0.36)	(0.62)	(0.28)	(0.40)	(0.26)	(0.39)
	2	Policy at 10th percentile	0.03	−0.06	0.14	−0.09	0.05	0.03	0.12	0.06
			(0.11)	(0.36)	(0.12)	(0.32)	(0.11)	(0.28)	(0.12)	(0.27)
		Policy at means	0.09	−0.02	0.18	−0.09	0.05	−0.01	0.16	0.08
			(0.13)	(0.38)	(0.13)	(0.33)	(0.10)	(0.27)	(0.12)	(0.24)
		Policy at 90th percentile	0.22	0.08	0.28	−0.09	0.04	−0.09	0.25	0.13
			(0.26)	(0.56)	(0.29)	(0.62)	(0.28)	(0.54)	(0.35)	(0.58)
	3	Policy at 10th percentile	0.21	0.29	0.19	0.07	0.32	0.38	0.03	0.05
			(0.25)	(0.61)	(0.24)	(0.56)	(0.49)	(0.23)	(0.41)	(0.22)
		Policy at means	0.35	0.07	0.44	0.15	0.10	0.23	0.25	0.27
			(0.27)	(0.60)	(0.28)	(0.39)	(0.42)	(0.18)	(0.31)	(0.19)
		Policy at 90th percentile	0.66	−0.43	1.02*	0.32	−0.38	−0.11	0.78**	0.80**
			(0.48)	(0.88)	(0.55)	(1.01)	(0.53)	(0.31)	(0.38)	(0.32)
	4	Policy at 10th percentile	−0.01	0.02	0.02	0.31	0.11	0.04	−0.04	−0.00
			(0.14)	(0.36)	(0.13)	(0.96)	(0.15)	(0.30)	(0.15)	(0.29)
		Policy at means	−0.00	0.01	0.13	0.21	0.00	−0.15	0.14	0.14
			(0.16)	(0.43)	(0.16)	(0.51)	(0.13)	(0.26)	(0.12)	(0.22)
		Policy at 90th percentile	0.01	−0.00	0.31	0.06	−0.20	−0.49	0.42*	0.36
			(0.31)	(0.70)	(0.31)	(0.54)	(0.27)	(0.42)	(0.23)	(0.38)

Notes: Bootstrap standard errors are reported in parentheses. See Table S3 for detailed variable description. In 8/OLS, row (2) has an additional observation over rows (1) and (3). If we drop the extra observations, the results are unchanged.

*$p < .1$.
**$p < .05$.
***$p < .01$.

as 0.94. In the next section, we further address this concern by employing two alternative aid measures.

Lastly, in row 4, we substitute BD's original policy index for our updated policy index given the plausibility that differences are related to measuring a country's policy. We find one positive and significant marginal effect at the 10% level, occurring in the high policy, low-income specification, 8/OLS.

Collectively, these sensitivity checks suggest that the change in significance from BD's work to our updated findings is driven by a change in observations, thus generalizing the

TABLE 3
Extended Sample, 1962–2013, BD and ELR Regressions 4, 7, 5, and 8

		Outliers Included				Hadi Method, Outliers Excluded			
		All Countries		Lower Income		All Countries		Lower Income	
		4/OLS	4/2SLS	7/OLS	7/2SLS	5/OLS	5/2SLS	8/OLS	8/2SLS
Panel A: BD 1962–2013, Coefficients for aid, policy, aid*policy and aid^2*policy									
Aid	New data, BD countries	−0.06	−0.57*	0.03	−0.46	−0.09	−0.44	−0.02	−0.33
		(0.09)	(0.34)	(0.10)	(0.32)	(0.12)	(0.33)	(0.14)	(0.30)
	New data, full sample	−0.06	−0.75**	0.03	−0.49	−0.10	−0.56*	−0.03	−0.34
		(0.10)	(0.38)	(0.12)	(0.34)	(0.11)	(0.32)	(0.14)	(0.32)
Policy	New data, BD countries	0.77***	0.82***	0.87***	1.29**	0.84***	0.65***	0.96***	0.62
		(0.15)	(0.22)	(0.27)	(0.53)	(0.15)	(0.18)	(0.24)	(0.44)
	New data, full sample	0.77***	0.81***	0.87***	1.24**	0.83***	0.62***	0.95***	0.66
		(0.14)	(0.22)	(0.23)	(0.49)	(0.15)	(0.17)	(0.24)	(0.40)
Aid*policy	New data, BD countries	0.15**	−0.11	0.13*	−0.38	0.09	0.24**	0.06	0.22
		(0.07)	(0.20)	(0.07)	(0.29)	(0.05)	(0.10)	(0.07)	(0.14)
	New data, full sample	0.13**	−0.11	0.11	−0.32	0.09*	0.25***	0.06	0.19
		(0.06)	(0.20)	(0.07)	(0.25)	(0.05)	(0.09)	(0.06)	(0.13)
Aid^2*policy	New data, BD countries	−0.01	0.06*	−0.01	0.09**				
		(0.01)	(0.04)	(0.01)	(0.04)				
	New data, full sample	−0.01	0.07*	−0.01	0.08**				
		(0.01)	(0.04)	(0.01)	(0.03)				
Number of observations	New data, BD countries	506	419	337	277	499	416	332	275
	New data, full sample	538	443	343	283	530	439	338	281
Panel B: ELR 1962–2013, Coefficients for aid, policy, aid*policy and aid^2*policy									
Aid	New data, ELR countries	−0.09	−0.50	−0.00	−0.44	−0.13	−0.41	−0.04	−0.14
		(0.10)	(0.33)	(0.12)	(0.37)	(0.11)	(0.28)	(0.13)	(0.35)
	New data, full sample	−0.30**	−1.03**	−0.22	−1.25**	−0.42***	−0.90**	−0.43**	−0.88*
		(0.13)	(0.41)	(0.15)	(0.53)	(0.16)	(0.37)	(0.19)	(0.48)
Policy	New data, ELR countries	0.73***	0.77***	0.94***	1.56***	0.80***	0.57***	1.03***	0.97**
		(0.15)	(0.23)	(0.25)	(0.54)	(0.14)	(0.18)	(0.25)	(0.45)
	New data, full sample	0.58***	0.51**	0.73**	0.61	0.61***	0.40**	0.63**	0.20
		(0.18)	(0.24)	(0.30)	(0.58)	(0.15)	(0.20)	(0.30)	(0.53)
Aid*policy	New data, ELR countries	0.16**	−0.13	0.14*	−0.56*	0.11**	0.24**	0.07	0.10
		(0.06)	(0.21)	(0.08)	(0.33)	(0.05)	(0.09)	(0.06)	(0.16)
	New data, full sample	0.23***	0.16	0.20**	−0.04	0.23***	0.41***	0.25***	0.44**
		(0.07)	(0.23)	(0.09)	(0.30)	(0.07)	(0.13)	(0.08)	(0.21)
Aid^2*policy	New data, ELR countries	−0.01	0.07*	−0.01	0.11**				
		(0.01)	(0.04)	(0.01)	(0.05)				
	New data, full sample	−0.01	0.05	−0.00	0.09*				
		(0.00)	(0.04)	(0.00)	(0.05)				
Number of observations	New data, ELR countries	551	462	365	303	545	458	361	301
	New data, full sample	600	493	393	321	591	488	388	320

Notes: Bootstrap standard errors are reported in parentheses. Time fixed effects are included in all regressions. Each specification includes a constant term, measure of aid/GDP, a policy index, an aid*policy interaction term, log initial GDP, ethnic fractionalization, political assassinations, a fractionalization*assassinations interaction term, a measure of institutional quality, and a measure of financial depth (M2/GDP lagged), regional dummies for Sub-Saharan Africa and fast-growing East Asian countries. BD and ELR specifications differ in their definitions of regional dummies and low-income (see Table S6). See Table S3 for detailed variable description.

*$p < .1$.
**$p < .05$.
***$p < .01$.

findings of Hansen and Tarp (2000). Our results reflect the observation sensitive feature of the aid-growth literature, making it difficult to compare across studies even when keeping the methodology unchanged.

B. Extended Sample Analysis

In Table 3, we extend the sample with more periods, averaged from 1962 to 2013. Panel A replicates the BD specifications, and Panel B replicates the ELR specifications. Reporting on coefficients with a 5% or lower p value, in the BD specifications we find four positive and significant interaction term coefficients out of a possible 16. Two are from OLS regressions with outliers; once the outliers are excluded, however, the coefficients are no longer significant. The other two significant coefficients are from 2SLS specifications excluding outliers. For 16 ELR specifications, we report six of eight OLS specifications with positive and significant interactions. When excluding outliers, we also find three of eight 2SLS regressions with positive and significant interaction term coefficients. Interestingly, with the extended sample, the ELR specifications are more supportive of BD's conclusion than the BD specifications.

TABLE 4
Marginal Effects, 1962–2013, BD and ELR Regressions 4, 7, 5, and 8

		Outliers Included				Hadi Method, Outliers Excluded			
		All Countries		Lower Income		All Countries		Lower Income	
		4/OLS	4/2SLS	7/OLS	7/2SLS	5/OLS	5/2SLS	8/OLS	8/2SLS
Panel A: BD 1962–2013, coefficients for marginal effects of aid									
Policy at 10th percentile	New data, BD countries	0.02	−0.50	0.10	−0.45*	−0.01	−0.29	0.09	−0.14
		(0.08)	(0.33)	(0.09)	(0.26)	(0.07)	(0.27)	(0.08)	(0.22)
	New data, full sample	0.04	−0.66*	0.10	−0.47*	0.01	−0.35	0.09	−0.13
		(0.08)	(0.37)	(0.09)	(0.28)	(0.07)	(0.26)	(0.07)	(0.23)
Policy at mean	New data, BD countries	0.15	−0.38	0.18*	−0.43	0.10	−0.00	0.17**	0.12
		(0.10)	(0.35)	(0.09)	(0.29)	(0.08)	(0.24)	(0.07)	(0.20)
	New data, full sample	0.15	−0.54	0.17*	−0.44	0.11	−0.04	0.16**	0.10
		(0.09)	(0.40)	(0.09)	(0.30)	(0.07)	(0.23)	(0.07)	(0.19)
Policy at 90th percentile	New data, BD countries	0.30**	−0.25	0.28*	−0.41	0.25**	0.32	0.27**	0.42
		(0.15)	(0.42)	(0.14)	(0.42)	(0.12)	(0.26)	(0.12)	(0.30)
	New data, full sample	0.29**	−0.40	0.27*	−0.41	0.24**	0.30	0.24**	0.38
		(0.14)	(0.48)	(0.14)	(0.42)	(0.12)	(0.25)	(0.11)	(0.27)
Number of observations	New data, BD countries	506	419	337	277	499	416	332	275
	New data, full sample	538	443	343	283	530	439	338	281
Panel B: ELR 1962–2013, coefficients for marginal effects of aid									
Policy at 10th percentile	New data, ELR countries	0.01	−0.42	0.08	−0.46	−0.02	−0.24	0.05	−0.06
		(0.08)	(0.31)	(0.09)	(0.30)	(0.08)	(0.23)	(0.08)	(0.21)
	New data, full sample	−0.02	−0.60*	0.03	−0.70**	−0.06	−0.42*	−0.01	−0.25
		(0.09)	(0.36)	(0.08)	(0.34)	(0.07)	(0.24)	(0.08)	(0.21)
Policy at mean	New data, ELR countries	0.15	−0.31	0.17*	−0.48	0.13*	0.04	0.11	0.05
		(0.09)	(0.34)	(0.09)	(0.32)	(0.07)	(0.20)	(0.08)	(0.17)
	New data, full sample	0.14	−0.34	0.15*	−0.43	0.09	−0.07	0.12	0.06
		(0.09)	(0.37)	(0.09)	(0.32)	(0.07)	(0.20)	(0.08)	(0.18)
Policy at 90th percentile	New data, ELR countries	0.35***	−0.18	0.28*	−0.50	0.31***	0.37	0.18	0.17
		(0.13)	(0.41)	(0.14)	(0.43)	(0.11)	(0.22)	(0.13)	(0.27)
	New data, full sample	0.35***	−0.03	0.31**	−0.09	0.29***	0.34	0.28**	0.45
		(0.13)	(0.42)	(0.13)	(0.39)	(0.11)	(0.21)	(0.12)	(0.28)
Number of observations	New data, ELR countries	551	462	365	303	545	458	361	301
	New data, full sample	600	493	393	321	591	488	388	320

Notes: Bootstrap standard errors are reported in parentheses. Marginal effects from regressions in Table 3.
*$p < .1$.
**$p < .05$.
***$p < .01$.

Although these results are more supportive of BD than previous tests, they do not provide a conclusive answer as to whether aid can support growth in the presence of good policies. BD and ELR analyze the sign and significance of the coefficient of the interaction terms without reporting the marginal effects. In order to provide further insight, however, we calculate marginal effects at different policy levels.

Table 4 reports marginal effects of aid for the policy index at the 10th percentile (poor policy), the mean, and the 90th percentile (good policy) for all the specifications from Table 3. The evidence suggests that the aid-growth relation is not conditional on policies adopted in recipient countries, as most marginal effects are insignificant. Any positive and significant marginal effect reported occurs only in OLS estimations. For example, in both BD and ELR specifications, we find six of eight OLS regressions with positive and significant marginal effects for good policy countries. However, this result does not hold for the 2SLS estimations. In fact, in one 2SLS estimation, we find a negative and significant marginal effect for low-income, poor policy countries, suggesting that in countries with poor policies aid can decrease economic growth. These results highlight the need to properly account for endogeneity—an issue we address in a later section.

Combined, the extended sample OLS regressions provide some support for BD's conclusion that aid may increase growth when a recipient country has top policy scores. The insignificant results from the 2SLS specifications and the overwhelming insignificant marginal effects make it difficult to conclude that aid is effective at increasing growth.

C. Post-Cold War Analysis

The post-Cold War period (1990–2013) is marked with significant changes to the aid landscape (Dunning 2004; Frot, Olofsgard, and

TABLE 5
Post-Cold War Sample, 1990–2013, BD and ELR Regressions 4, 7, 5, and 8

		Outliers Included				Hadi Method, Outliers Excluded			
		All Countries		Lower Income		All Countries		Lower Income	
		4/OLS	4/2SLS	7/OLS	7/2SLS	5/OLS	5/2SLS	8/OLS	8/2SLS
Panel A: BD 1990–2013, coefficients for aid, policy, aid*policy and aid^2*policy									
Aid	New data, BD countries	−0.33*	−0.80	−0.16	−0.50	−0.73***	−1.51***	−0.33	−0.73
		(0.18)	(0.57)	(0.18)	(0.47)	(0.25)	(0.55)	(0.24)	(0.46)
	New data, full sample	−0.71	−1.71	−0.11	−0.31	−0.76	−0.49	0.15	1.54
		(0.52)	(1.06)	(0.59)	(2.14)	(0.75)	(1.65)	(0.70)	(1.77)
Policy	New data, BD countries	0.71**	0.64	1.23**	2.56**	0.28	−0.27	0.93	0.64
		(0.29)	(0.48)	(0.52)	(1.12)	(0.26)	(0.42)	(0.59)	(0.88)
	New data, full sample	0.55	0.44	1.40	2.39	0.46	0.75	2.40**	4.25**
		(0.34)	(0.55)	(0.92)	(1.81)	(0.60)	(1.02)	(1.13)	(2.07)
Aid*policy	New data, BD countries	0.13	−0.01	−0.06	−0.92*	0.37***	0.75***	0.16	0.34
		(0.13)	(0.36)	(0.15)	(0.55)	(0.12)	(0.25)	(0.13)	(0.26)
	New data, full sample	0.29	0.31	0.01	−0.61	0.34	0.16	−0.07	−0.79
		(0.22)	(0.42)	(0.25)	(0.83)	(0.34)	(0.74)	(0.32)	(0.78)
Aid^2*policy	New data, BD countries	0.01	0.07	0.02	0.14**				
		(0.01)	(0.05)	(0.02)	(0.06)				
	New data, full sample	0.01	0.06	0.01	0.09*				
		(0.01)	(0.05)	(0.01)	(0.05)				
Number of observations	New data, BD countries	245	211	165	141	237	205	159	137
	New data, full sample	324	275	214	181	315	268	211	180
Panel B: ELR 1990–2013, coefficients for aid, policy, aid*policy and aid^2*policy									
Aid	New data, ELR countries	−0.44**	−0.96**	−0.20	−0.57	−0.70***	−1.22***	−0.30	−0.65*
		(0.19)	(0.41)	(0.20)	(0.43)	(0.24)	(0.40)	(0.22)	(0.34)
	New data, full sample	−0.80	−1.59*	−0.31	0.21	−1.17**	−1.12	−0.35	2.80
		(0.50)	(0.95)	(0.58)	(1.71)	(0.59)	(1.52)	(0.82)	(2.56)
Policy	New data, ELR countries	0.69**	0.61	1.30**	2.38**	0.32	−0.19	0.97	0.62
		(0.28)	(0.38)	(0.54)	(0.93)	(0.29)	(0.34)	(0.59)	(0.84)
	New data, full sample	0.45	0.40	1.29*	2.62*	0.16	0.43	1.81	5.47*
		(0.33)	(0.50)	(0.78)	(1.49)	(0.52)	(0.91)	(1.14)	(2.90)
Aid*policy	New data, ELR countries	0.13	−0.04	−0.06	−0.80**	0.34***	0.63***	0.18	0.33*
		(0.12)	(0.25)	(0.14)	(0.39)	(0.11)	(0.17)	(0.11)	(0.19)
	New data, full sample	0.34	0.35	0.05	−0.63	0.53*	0.39	0.15	−1.39
		(0.21)	(0.37)	(0.24)	(0.67)	(0.27)	(0.65)	(0.37)	(1.13)
Aid^2*policy	New data, ELR countries	0.01	0.08**	0.02	0.13***				
		(0.01)	(0.04)	(0.02)	(0.04)				
	New data, full sample	0.01	0.05	0.01	0.05				
		(0.01)	(0.06)	(0.01)	(0.04)				
Number of observations	New data, ELR countries	272	238	180	156	265	232	176	153
	New data, full sample	324	275	199	169	315	268	195	167

Notes: Bootstrap standard errors are reported in parentheses. Time fixed effects are included in all regressions. BD and ELR specifications differ in their definitions of regional dummies and low-income (see Table S6). See Table S3 for detailed variable description.
*$p < .1$.
**$p < .05$.
***$p < .01$.

Perrota 2014; Greenhill, Prizzon, and Rogerson 2016; Griffin 2000). In addition, the average quality of economic policy and economic institutions has increased and the institutional gap between developed and developing countries has decreased (Gwartney, Lawson, and Hall 2017, 19). Thus, we examine BD and ELR specifications post-Cold War.[16] To account for improvements in institutional quality, we replace BD and ELR's 1980 international country risk guide (ICRG) static measure with varying values collected from ICRG (PRS Group 2016).[17] Otherwise, the model is left unchanged.

The post-Cold War evidence is more in favor of ELR's nonsignificant interaction term coefficients. As shown in Table 5, across both BD and ELR specifications most interactions are insignificant. Among the BD regressions, we report a

16. See Table S2 for country differences among BD, ELR, and the post-1990 sample.

17. Following Rajan and Subramanian (2008), we take the sum of bureaucratic quality, rule of law, and corruption. Scores range 0–16. Data are available from 1984 to 2015, so we only test the post-1990 period with updated ICRG data.

TABLE 6
Marginal Effects, 1990–2013, BD and ELR Regressions 4, 7, 5, and 8

		Outliers Included				Hadi Method, Outliers Excluded			
		All Countries		Lower Income		All Countries		Lower Income	
		4/OLS	4/2SLS	7/OLS	7/2SLS	5/OLS	5/2SLS	8/OLS	8/2SLS
Panel A: BD 1990–2013, coefficients for marginal effects of aid									
Policy at 10th percentile	New data, BD countries	−0.21	−0.63	−0.14	−0.76**	−0.21*	−0.56	−0.10	−0.56**
		(0.14)	(0.47)	(0.13)	(0.34)	(0.11)	(0.37)	(0.11)	(0.27)
	New data, full sample	−0.12	−0.71	−0.04	−0.75	−0.11	−0.32	0.00	−0.12
		(0.16)	(0.47)	(0.15)	(0.46)	(0.13)	(0.27)	(0.11)	(0.29)
Policy at mean	New data, BD countries	−0.05	−0.40	−0.11	−1.11**	−0.01	−0.01	0.01	−0.34
		(0.16)	(0.15)	(0.46)	(0.49)	(0.09)	(0.49)	(0.09)	(0.27)
	New data, full sample	−0.05	−0.62	−0.03	−0.79**	−0.04	−0.23	0.01	−0.27
		(0.14)	(0.45)	(0.13)	(0.37)	(0.11)	(0.25)	(0.09)	(0.22)
Policy at 90th percentile	New data, BD countries	0.03	−0.30	−0.10	−1.27**	0.08	0.10	0.05	−0.24
		(0.19)	(0.51)	(0.18)	(0.63)	(0.11)	(0.29)	(0.11)	(0.35)
	New data, full sample	0.02	−0.49	−0.02	−0.84**	0.03	−0.11	0.02	−0.44*
		(0.14)	(0.45)	(0.14)	(0.35)	(0.12)	(0.26)	(0.12)	(0.25)
Panel B: ELR 1990–2013, coefficients for marginal effects of aid									
Policy at 10th percentile	New data, ELR countries	−0.25*	−0.71**	−0.15	−0.75***	−0.21*	−0.39	−0.08	−0.39**
		(0.14)	(0.33)	(0.12)	(0.28)	(0.11)	(0.24)	(0.11)	(0.18)
	New data, full sample	−0.10	−0.60	−0.11	−0.58	−0.07	−0.36	0.00	−0.19
		(0.16)	(0.45)	(0.14)	(0.42)	(0.13)	(0.30)	(0.09)	(0.28)
Policy at mean	New data, ELR countries	−0.09	−0.48	−0.11	−0.90***	−0.01	−0.00	0.03	−0.23
		(0.14)	(0.34)	(0.12)	(0.34)	(0.08)	(0.22)	(0.09)	(0.20)
	New data, full sample	−0.02	−0.50	−0.08	−0.68**	0.01	−0.26	0.03	−0.33
		(0.14)	(0.44)	(0.11)	(0.34)	(0.11)	(0.29)	(0.09)	(0.22)
Policy at 90th percentile	New data, ELR countries	−0.00	−0.37	−0.09	−0.98**	0.09	0.20	0.08	−0.15
		(0.16)	(0.38)	(0.16)	(0.43)	(0.10)	(0.25)	(0.11)	(0.26)
	New data, full sample	0.08	−0.37	−0.05	−0.77**	0.10	−0.13	0.06	−0.47*
		(0.15)	(0.43)	(0.13)	(0.33)	(0.12)	(0.30)	(0.11)	(0.27)

Notes: Bootstrap standard errors are reported in parentheses. Marginal effects from Table 5.
*$p < .1$.
**$p < .05$.
***$p < .01$.

positive and significant (5% or lower p value) coefficient in one of eight for both the OLS and 2SLS specifications. In the ELR regressions, we also find one OLS and one 2SLS specification with a positive and significant interaction term coefficient. However, for low-income countries (including outliers) we also find one negative and significant coefficient from the 2SLS estimation.[18]

Perhaps more interesting are the marginal effects presented in Table 6. There is not one positive and significant marginal effect. We do find several *negative* and significant marginal effects from 2SLS specifications, occurring at all levels of the policy index and mostly in lower-income countries. For poor policy countries, aid's conditional impact is negative and significant in 5 of 16 2SLS estimations. Mean level policy countries report 4 of 16 2SLS specifications with negative and significant marginal effects. For good policy countries, the critical test for BD, from the 2SLS estimations, we also find 4 of 16 with negative and significant marginal effects. No OLS estimation reports a significant marginal effect (at 5% level or lower p value), suggesting that controlling for endogeneity is especially important post-Cold War.

Collectively, the post-Cold War results indicate that the majority of the marginal effects are insignificant, consistent with the results from the 1962–2013 period; however, the significant marginal effects we do find are *negative*. This finding provides evidence that aid can decrease growth regardless of the type of policy environment, a possibility not explored in either BD or

18. This may indicate that the post-Cold War data exhibit more nonlinear associations. BD argue the quadratic interaction terms control for the nonlinear relation caused by outliers; hence, once the outliers are dropped the quadratic terms are excluded as well. The significant quadratic interaction terms are found in both the 1962–2013 and post-1990 samples. Previous findings also report significant quadratic terms (Dayton-Johnson and Hoddinott 2003; Hansen and Tarp 2000; Kohama, Sawada, and Kono 2003). Also, Chatelain and Ralf (2014, 94) find the quadratic interaction term can be a spurious effect.

ELR.[19] For example, according to the 7/2SLS, ELR country specification, a 1% increase in aid may *decrease* growth by 0.75 (10th percentile policy), 0.90 (mean policy), or 0.98 (90th percentile policy) percentage points. The negative results are also strongest in low-income countries, those countries policymakers are specifically attempting to help.

Before moving on, we examine the possible importance of improvements in policy and aid allocation patterns based on policy to explain the differences in BD's results and the updated findings. Since BD and ELR's analysis, countries have improved their economic policies, including reducing trade barriers. For example, the updated trade openness variable indicates that 26 BD countries and 25 ELR countries now have an open trade status. Under the 2SLS full country sample, comparing the 1970–1993/1997 sample with the post-1990 period, the mean policy score increases from 1.40 to 2.22 under the BD specification, and from 1.67 to 2.18 under the ELR specification. As ELR suggest, significant interactions may occur if the policy environment of recipient countries improves. Thus, these improvements could explain some of the differences between the original findings and the updated results.

To consider this possibility, we examine skewness in the policy index over time. Kurtosis in all samples is positive indicating heavily weighted tails. Skewness for 1970–1993/1997 is positive, but it is negative for the post-1990 sample (and the 1962–2013 extended sample). This suggests that policy scores increased, on average, after 1990. We illustrate this trend in Figure 1, plotting quartiles of the policy index. Figures 1A and 1B compare the policy trend using full country sample with BD's years (1970–1993) and the post-1990 sample, respectively. The majority of policy scores for the BD sample are below two but over two for the post-1990 sample. This suggests that policies are improving over time with better policies occurring post-1990.

If policy is improving over time, according to BD's argument, aid's conditional impact should also be increasing; however, we do not find evidence in support of this argument. This could be due to aid allocation patterns. Based on the post-1990 sample, we find, on average, that countries with bottom policy scores receive more aid than countries with top policy scores.[20] Donors tend to allocate disproportionately more aid to the poorest policy countries, which could explain the lack of a policy conditioning effect.

We test for this trade-off in the 1970–1993/1997 and post-1990 samples. In BD's 1970–1993 sample, good policy countries *positively* correlate with aid. In the other samples, the correlation between aid and policy is *negative*. This pattern of aid allocation suggests that during BD's time frame aid flowed to countries with a better policy environment, increasing the likelihood that aid could be effective. However, under different periods, aid tends to flow to poor policy countries, possibly rendering aid less effective. We view this as additional evidence that BD's result is contingent on the timing of their analysis.

Collectively, the majority of the estimations from the replication, extended sample, and post-Cold War period report insignificant coefficients and marginal effects. Thus, we are unable to support BD's conclusion that a good policy environment increases aid effectiveness. The results support ELR's finding that aid's impact on growth is not conditional on policy. However, we do find a few positive and significant interaction term coefficients and marginal effects that are both positive and negative. This makes it difficult to draw a strong definitive conclusion; however, the findings do suggest that calling for more aid as an attempt to meet the SDGs may have limited success with potential unintended consequences.

V. MOVING BEYOND THE ORIGINAL FRAMEWORK

A. Alternative Measure of Institutions

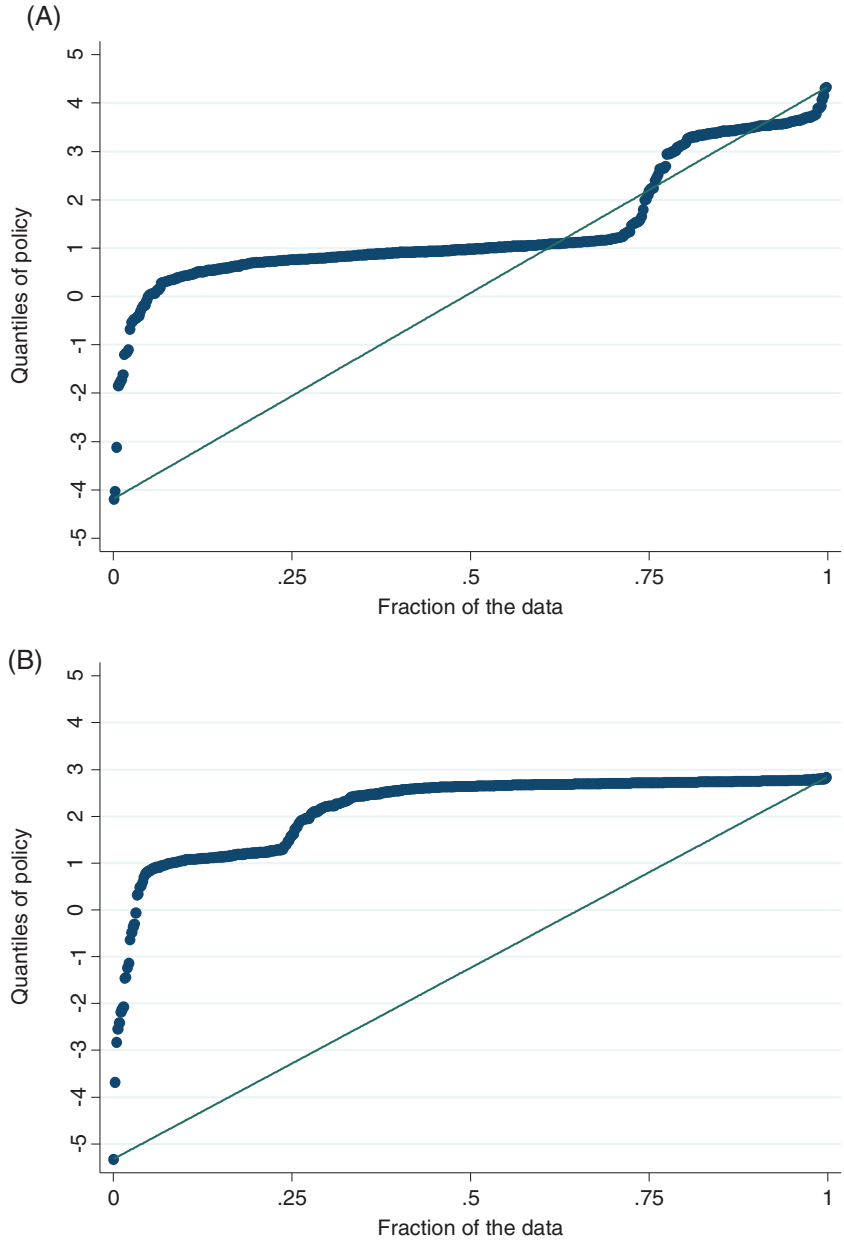
Along with the improvement of policy, institutional quality has also improved over time. At the time of BD's publication, the new institutional

19. It is possible that aid*policy interaction term coefficient is biased due to not controlling for aid^2 and $policy^2$. To address this concern, we include both Aid^2 and $Policy^2$ to the original framework using BD and ELR models over the extended and post-Cold War samples (outliers included). When using OLS for the extended sample, three marginal effects are positive and significant for policy at the 90th percentile. However, in one 2SLS estimation, we find one negative and significant marginal effect in low-income, good policy countries. Results are unchanged with excluding outliers. In general, including the two squared terms does not alter the prior findings, indicating that the interaction term is not picking up the effect from these terms. Results are not tabulated but available upon request. We thank an anonymous reviewer for this comment.

20. For example, countries in the bottom 10% of policy scores receive about 20% (BD) or 30% (ELR) more aid than top 10% policy scoring countries, according to the 2SLS full country samples. These numbers increase to about 7 times and 60 times when comparing the bottom 1% countries to the top 1% countries in policy scores, respectively.

FIGURE 1

Trend in Policy Scores. (A) Quantile Plot of Policy, BD Specification, 1970–1993 Full Sample and (B) Quantile Plot of Policy, BD Specification, 1990–2013 Full Sample



literature and data was in its infancy. BD used the available data on institutions from ICRG, holding the 1980 values constant throughout their sample.

To capture the changes in institution quality, we utilize an alternative institutional measure: the economic freedom index (EFW) (Gwartney, Lawson, and Hall 2016). The EFW is a widely cited measure of the quality of a country's economic environment including protection of property rights, size of government, business

TABLE 7

Alternative Institutional Measure, Full Sample, 1990–2013, BD and ELR Regressions 4, 7, 5, and 8

		Outliers Included				Hadi Method, Outliers Excluded			
		All Countries		Lower Income		All Countries		Lower Income	
		4/OLS	4/2SLS	7/OLS	7/2SLS	5/OLS	5/2SLS	8/OLS	8/2SLS
Panel A: EFW, 1990–2013 full sample, coefficients for aid, policy, aid*policy and aid^2*policy with marginal effects									
Aid	BD specification	−1.22***	−2.94***	−0.94**	−2.95***	−1.19***	−1.51**	−0.67*	−1.20*
		(0.34)	(0.80)	(0.41)	(1.03)	(0.30)	(0.63)	(0.38)	(0.68)
	ELR specification	−1.15***	−2.86***	−0.85**	−3.21***	−1.16***	−1.57**	−0.54*	−1.13*
		(0.31)	(0.81)	(0.38)	(1.04)	(0.27)	(0.66)	(0.29)	(0.66)
EFW	BD specification	0.50**	0.26	0.68	0.05	0.33	0.39	0.66	0.41
		(0.23)	(0.29)	(0.43)	(0.57)	(0.23)	(0.28)	(0.45)	(0.52)
	ELR specification	0.51**	0.27	0.80**	0.13	0.34*	0.40	0.69**	0.46
		(0.20)	(0.29)	(0.33)	(0.57)	(0.20)	(0.28)	(0.32)	(0.52)
Aid*EFW	BD specification	0.13***	0.30***	0.10*	0.31**	0.22***	0.20*	0.14*	0.19
		(0.05)	(0.10)	(0.06)	(0.14)	(0.06)	(0.12)	(0.07)	(0.13)
	ELR specification	0.12***	0.29***	0.08	0.31**	0.21***	0.20*	0.12**	0.18
		(0.04)	(0.11)	(0.05)	(0.14)	(0.05)	(0.12)	(0.06)	(0.13)
Aid^2*EFW	BD specification	0.01***	0.02**	0.01***	0.02**				
		(0.00)	(0.01)	(0.00)	(0.01)				
	ELR specification	0.01***	0.02**	0.01***	0.03***				
		(0.00)	(0.01)	(0.00)	(0.01)				
Panel B: EFW, 1990–2013 full sample, coefficients for marginal effects of aid									
EFW at 10th percentile	BD specification	−0.37***	−1.09***	−0.22*	−0.86***	−0.15**	−0.53***	−0.02	−0.27*
		(0.11)	(0.31)	(0.11)	(0.29)	(0.07)	(0.16)	(0.07)	(0.14)
	ELR specification	−0.36***	−1.08***	−0.23**	−0.99***	−0.16**	−0.58***	0.02	−0.24
		(0.11)	(0.32)	(0.12)	(0.31)	(0.07)	(0.18)	(0.07)	(0.15)
EFW at mean	BD specification	−0.13	−0.62**	−0.03	−0.38*	0.14	−0.28	0.14	−0.06
		(0.10)	(0.26)	(0.10)	(0.23)	(0.08)	(0.18)	(0.09)	(0.18)
	ELR specification	−0.13	−0.63**	−0.06	−0.50**	0.12	−0.33*	0.16*	−0.03
		(0.10)	(0.27)	(0.10)	(0.25)	(0.09)	(0.20)	(0.09)	(0.18)
EFW at 90th percentile	BD specification	0.10	−0.16	0.15	0.10	0.42***	−0.04	0.29**	0.15
		(0.13)	(0.29)	(0.14)	(0.31)	(0.14)	(0.28)	(0.14)	(0.30)
	ELR specification	0.09	−0.19	0.10	0.00	0.39***	−0.08	0.30**	0.18
		(0.13)	(0.30)	(0.14)	(0.32)	(0.14)	(0.29)	(0.14)	(0.30)
Number of observations	BD specification	435	320	285	208	428	315	282	207
	ELR specification	435	320	265	201	429	316	261	199

Notes: Robust standard errors are reported in parentheses for Panel A. Panel A replaces ICRG and policy index with EFW measures of institutions. Panel B reports the marginal effects based on Panel A. Time fixed effects are included in all regressions. See Table S3 for detailed variable description.

*$p < .1$.

**$p < .05$.

***$p < .01$.

regulations, stable money, and trade barriers (see Hall and Lawson 2014, for a summary). EFW is a more comprehensive measure of economic institutions and policy. Thus, we drop ICRG values and replace them with economic freedom scores, but we also drop the policy index due to the overlap between economic freedom and the policy index. Data are available starting in the 1990s; thus, we use the post-Cold War sample.

To test aid's conditional impact, we create an interaction term between aid and EFW. As such, we broaden BD's claim that aid in a good policy country can increase growth to now consider that aid can increase growth in more economically free countries, which protect property, enforce contracts, reduce trade barriers, and so on.

As shown in Table 7, Panel A, the majority of the aid*EFW coefficients from the OLS estimations are positive and significant (six of eight). We also find four of eight positive and significant interactions from the 2SLS specifications if outliers are included in the model. The marginal effects, however, do not support the idea that aid can increase growth in high-quality institution countries. Instead, for countries with EFW scores in the 10th percentile, the overwhelming majority of marginal effects are negative and significant. This suggests that in economically unfree countries aid can decrease economic growth. At a mean level of economic freedom, most marginal effects are insignificant. The three significant marginal effects reported are negative and from the 2SLS estimation. As economic freedom increases, aid's conditional impact switches from negative to positive but the effects are usually insignificant.

B. Alternative Measurement of Growth

In this section, we further explore potential reasons why results differ across aid studies. Ram and Ural (2014) compare real GDP per capita (PPP) in WDI and PWT and find large measurement differences. They suggest using both data sources for GDP as robustness. When measuring growth, BD uses PWT 5.6 and ELR utilizes WDI (World Bank Group 2018). In our previous tests, we also use data on growth from WDI.

To test the sensitivity of the results, we switch our dependent variable to GDP growth rates collected from PWT 8.1 (see Feenstra, Inklaar, and Timmer 2015) instead of WDI.[21,22] We do indeed find more significant interaction term coefficients compared to our prior results. Out of the 16 OLS specifications, 13 have positive and significant interaction term coefficients. From the 2SLS estimations, 4 of 16 are also positive and significant (when excluding outliers). Examining the marginal effects suggest that aid does not influence growth in poor policy or mean-level policy countries, as most marginal effects are insignificant. In countries at the 90th percentile of policy, half of the OLS estimations report positive and significant marginal effects. However, only 4 of 16 2SLS estimations have positive and significant marginal effects (outliers excluded). We do not tabulate these findings, but they are available upon request.

Together, this provides another potential reason why BD found significance and ELR did not—the measurement of growth is sensitive to methodology differences between PWT and WDI. The extended sample specifications utilizing PWT growth measures are the strongest in favor of BD. When analyzing the post-Cold War sample, however, the majority of interaction term coefficients remain insignificant, as do the marginal effects from both the extended and post-1990 period.

C. Alternative Measurements of Aid

Next, we examine how measurement of foreign aid may cause differences in the results. We use two alternative measurements, constant dollar EDA and ODA. Previously, we adopt ELR's aid measurement that divides current dollar EDA by current dollar GDP. BD, however, use nominal aid flows and nominal GDP deflated with the unit-value of imports price index from International Financial Statistics. This difference in the construction of aid provides another potential reason for dissimilar findings across aid studies, including BD and ELR.

In order to exhaust this possibility, we reconstruct a constant dollar EDA measure by deflating nominal EDA using the investment deflator from PWT 9.0 (Caselli and Feyrer 2007; Hsieh and Klenow 2007).[23] Using constant EDA, we retest the extended and post-1990 samples. Several positive and significant interaction term coefficients are noted; however, the majority of the results do not provide support for BD's conclusion. We do note that all significant interactions are in models excluding outliers, which may indicate that BD's measurement of aid is more sensitive to outliers. Compared to previous tables, by switching to constant EDA we find fewer significant interaction terms. This suggests that there are differences based on the measurement of aid but not enough to alter our main conclusion.

Another possible issue is that our EDA measure is calculated by extrapolating Chang, Fernandez-Arias, and Serven's (1999) EDA from ODA. This estimation may be biased if the linear association between EDA and ODA does not hold. Following Dalgaard and Hansen (2001), we simply swap EDA with ODA to check for any difference in the aid-policy-growth association due to measurement issues. We retest both the extended and post-1990 samples. By employing ODA, we find similar results as those reported in Tables 3 and 5.

Together, these findings demonstrate minor differences in the results due to the construction of aid. The results are not tabulated to save space but are available upon request.

D. Alternative Estimation Techniques

To understand the impact of estimation selection, we utilize several advanced estimation methods for robustness. First, we switch from OLS and 2SLS methods to FE, FD, and system GMM. Lu and Ram (2001) find that

21. PWT does not directly report a growth rate measure. We use expenditure-side real GDP at chained PPPs (in million 2005 US$) divided by population (in millions); both variables are collected from PWT 8.1. From this, we calculate the annual percentage change to create a measure of economic growth.

22. Given that in Tables 2–5, BD/ ELR countries and full country samples have very similar results, this section only reports full sample results.

23. The unit value of imports price index is not available before 2000; thus, we use the investment deflator from PWT 9.0. We rescale EDA to constant 2005 dollars with price level of capital formation in PWT 9.0.

country dummies cancel the conditional effect of aid*policy on growth. Hansen and Tarp (2001), however, conclude that FEs increases the significance of aid's overall impact on growth. First differencing addresses issues with omitted variables in cross-country panel data, and is commonly utilized in the aid effectiveness literature (Clemens et al. 2012; Minasyan 2016; Yontcheva and Masud 2005). GMM has become more common in the aid literature as a means of addressing endogeneity (Clemens et al. 2012; Djankov, Montalvo, and Reynal-Querol 2008; Hansen and Tarp 2001; Rajan and Subramanian 2008).

In Table 8 we report BD and ELR's specifications for the full-country, post-Cold War sample using FE, FD, and GMM.[24] We do not observe any significant interaction coefficients. These results, which are arguably superior at addressing endogeneity than prior techniques employed, do not support the claim that policy conditions aid's effect on economic growth. As reported in Panel B, we find no positive and significant marginal effects, but there is some evidence that aid can harm growth in poor policy environments. Compared to earlier tests with BD's original estimation methods, we find fewer significant coefficients, suggesting that BD's results are contingent on estimation selection.

Building from Tan (2009), we next utilize long-panel data techniques, including the PMG estimator and DFE (Pesaran, Shin, and Smith 1999). Both methods have the advantage of allowing for long-run homogeneity and short-run heterogeneity, providing estimates of long-run averages and country-specific short-run estimators.[25] To retain BD's framework, we create an annual sample from 1984 to 2012.[26]

The results of PMG estimation are reported in Table 9. No aid*policy coefficient is significant at the 5% level or lower p value in the long run or averaged short run. More informative, marginal effects are negative at all levels of policy. In countries with average to lower quality policy, aid significantly decreases growth in the long run. These results support the prior post-Cold War findings where aid decreases growth and policy does not condition its effect.

Table S8 presents the country-specific short-run coefficients for all 69 countries in the sample. The evidence is quite consistent: aid*policy does not positively contribute to a country's growth. There are a few exceptions, however. The interaction term coefficient is significant at the 5% level for Indonesia and Tunisia and significant at the 10% level for Togo and Zambia. We also find that for eight countries, including Bangladesh, China, and Syria, aid*policy is negative and significant. Collectively, the long-run and short-run analyses provide evidence suggesting that aid can decrease economic growth.

For robustness, we also use DFE method and find consistent results—policy does not positively condition aid's effect on growth. We do find, however, that aid can hurt growth regardless of policy. These results are not tabulated to save space but are available upon request.

Collectively, the evidence from more advanced techniques supports the findings from the conventional methods (OLS and IV) as employed in BD, ELR, and other earlier works.

In an attempt to exhaust other possible empirical issues, we run several additional sensitivity tests. Arndt et al. (2010, 6) argue "the aid-growth relationship is only likely to emerge over a long time-horizon." Arndt, Jones, and Tarp (2015) find a long-term (30 years) positive effect of aid on growth, and others agree that aid's effect requires a longer time horizon (Rajan and Subramanian 2008). Thus, we re-aggregate the samples with multiple longer run periods, including 12-year averages, the average from 1970 to 1993/1997, a 1990–2013 average, and a full sample average from 1962 to 2013. With the exception of a few specifications, we do not find significant interaction term coefficients, and the marginal effects are insignificant. In BD's framework, the period length appears irrelevant. These results are not tabulated to save space.

Recent aid literature raises concerns about weak and invalid instrumental variables (Clemens et al. 2012; Dreher, Eichenauer, and Gehring 2016; Dreher and Langlotz 2015). In particular, Bazzi and Clemens (2009) and Deaton (2009) question if BD's instruments, including population, political relations, or historical/colonial, are valid. We attempt to find more suitable instruments but are unable to find

24. We do not tabulate the results for the extended sample as the dynamic ICRG measure for institutions is only available after 1984.

25. We greatly thank an anonymous reviewer for this suggestion.

26. Instead of using the extended sample, we create a sample using annual data between 1984 and 2012 due to limited data availability within BD's framework, for example, the ICRG measure of institutions. We assume I (1) cointegration following the current literature, including Tan (2009). Due to the structure of the BD framework, testing the optimal number of lags does not work. We test PMG method with different number of lags (up to four lags, as recommended by the PMG method), and the conclusion holds. These reports are available upon request.

TABLE 8
Alternative Estimation Techniques, FE, FD, and System GMM, Full Sample, 1990–2013, BD and ELR Regressions 4, 7, 5, and 8

		Outliers Included						Hadi Method, Outliers Excluded					
		All Countries			Lower Income			All Countries			Lower Income		
		4/FE	4/FD	4/GMM	7/FE	7/FD	7/GMM	5/FE	5/FD	5/GMM	8/FE	8/FD	8/GMM
Panel A: BD specification, full sample coefficients for aid, policy, aid*policy and aid^2*policy													
Aid	BD specification	−0.35	0.67	−3.73	0.23	0.93	−4.06	−1.63	0.31	−2.99	0.34	1.24	−4.06
		(0.82)	(0.76)	(2.93)	(1.16)	(1.01)	(4.54)	(1.31)	(0.76)	(4.02)	(1.22)	(0.81)	(4.54)
	ELR specification	−0.36	0.54	−4.47*	0.04	−0.03	−3.42	−1.47	0.34	−4.43	0.08	0.30	−3.81
		(0.76)	(0.72)	(2.47)	(0.97)	(0.86)	(2.37)	(1.09)	(0.69)	(4.46)	(1.23)	(0.82)	(2.66)
Policy	BD specification	−0.03	1.15	−5.29*	0.52	1.78	−3.71	−0.19	1.19	−4.77	3.00	2.74*	−3.71
		(0.47)	(0.75)	(2.98)	(2.18)	(1.61)	(3.13)	(1.00)	(0.82)	(3.14)	(2.05)	(1.47)	(3.13)
	ELR specification	−0.02	1.07	(2.68)	0.48	−0.24	−2.91	−0.09	1.16	−5.11	2.55	0.61	−2.42
		(0.43)	(0.70)	0.26	(1.94)	(1.57)	(2.58)	(0.87)	(0.81)	(3.52)	(1.77)	(1.67)	(3.06)
Aid*policy	BD specification	0.39	−0.11	2.12	0.22	−0.29	1.97	0.74	−0.17	1.71	−0.11	−0.61	1.97
		(0.37)	(0.36)	(1.35)	(0.53)	(0.45)	(1.59)	(0.61)	(0.36)	(1.47)	(0.55)	(0.39)	(1.59)
	ELR specification	0.40	−0.05	(1.24)	0.25	0.06	2.04	0.66	−0.19	2.19	−0.01	−0.17	2.02
		(0.34)	(0.34)	−0.02	(0.44)	(0.38)	(1.31)	(0.52)	(0.34)	(1.68)	(0.57)	(0.40)	(1.51)
Aid^2*policy	BD specification	−0.02	−0.01	−0.03	−0.02	−0.01	0.02			−0.03			0.02
		(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.09)			(0.07)			(0.09)
	ELR specification	−0.02	−0.01	(0.03)	−0.02	−0.00	−0.04			−0.01			−0.03
		(0.02)	(0.01)		(0.02)	(0.01)	(0.04)			(0.07)			(0.04)
Number of observations	BD specification	324	299	261	214	198	172	319	297	256	213	198	172
	ELR specification	324	299	261	199	181	157	315	297	256	195	180	155
Number of countries	BD specification	71		67	35		45	54		51	67		45
	ELR specification	71		67	42		39	70		67	41		39
AR(2) test p value[a]	BD specification			.41			.88			.36			.88
	ELR specification			.76			.73			.66			.54
Sargan/Hansen p value[b]	BD specification			.45			.43			.41			.43
	ELR specification			.51			.40			.44			.32
Panel B: BD specification, full sample coefficients for marginal effects of aid													
Policy at 10th percentile	BD specification	0.26	−0.27	0.39	0.43	−0.29	−0.01	−0.16	−0.34**	−0.09	0.09	−0.37*	0.12
		(0.34)	(0.22)	(0.58)	(0.34)	(0.22)	(1.13)	(0.21)	(0.17)	(0.58)	(0.23)	(0.19)	(0.56)
	ELR specification	0.26	−0.27	0.20	0.43	−0.29	0.13	−0.16	−0.34**	−0.26	0.09	−0.37*	−0.57
		(0.34)	(0.22)	(0.55)	(0.34)	(0.22)	(0.59)	(0.21)	(0.17)	(0.59)	(0.23)	(0.19)	(0.39)
Policy at mean	BD specification	0.34	0.11	0.75	0.45	0.08	0.42	−0.11	−0.03	0.29	0.12	−0.05	0.46
		(0.34)	(0.21)	(0.59)	(0.32)	(0.23)	(0.88)	(0.21)	(0.16)	(0.68)	(0.20)	(0.15)	(0.44)
	ELR specification	0.34	0.11	0.72	0.45	0.08	0.63	−0.11	−0.03	0.34	0.12	−0.05	−0.12
		(0.34)	(0.21)	(0.62)	(0.32)	(0.23)	(0.71)	(0.21)	(0.16)	(0.72)	(0.20)	(0.15)	(0.39)
Policy at 90th percentile	BD specification	0.41	0.34	1.12	0.47	0.35	0.81	−0.06	0.15	0.69	0.15	0.19	0.77
		(0.36)	(0.26)	(0.69)	(0.34)	(0.29)	(0.75)	(0.23)	(0.21)	(0.93)	(0.23)	(0.20)	(0.61)
	ELR specification	0.41	0.34	1.20	0.47	0.35	0.99	−0.06	0.15	0.90	0.15	0.19	0.27
		(0.36)	(0.26)	(0.77)	(0.34)	(0.29)	(0.87)	(0.23)	(0.21)	(1.07)	(0.23)	(0.20)	(0.57)

Notes: Bootstrap standard errors are reported in parentheses for FE and FD; robust standard errors are reported in parentheses for GMM due to insufficient observations for bootstrapping. Time fixed effects are included in all regressions. See Table S3 for detailed variable description.

[a]The null hypothesis is that the error term exhibits no second-order serial correlation.

[b]The null hypothesis is that the instruments are valid and not correlated with the residuals.

*$p < .1$.
**$p < .05$.
***$p < .01$.

TABLE 9
Alternative Method and Sample: Pooled Mean Group with Annual Data, Outliers Included, 1984–2012

	All Countries					Lower Income				
	EC	Aid	Policy	Aid* Policy	Aid²* Policy	EC	Aid	Policy	Aid* Policy	Aid²* policy
Panel A: Long-run coefficients and averaged short-run coefficients										
Long-run		−0.54***	0.61***	0.10*	0.01***		−0.44***	1.04***	0.03	0.01***
		(0.13)	(0.22)	(0.06)	(0.00)		(0.15)	(0.27)	(0.06)	(0.00)
Averaged short-run	−0.77***	25.94	−0.46	−12.12	19.50	−0.87***	28.33	0.22	−10.96	31.14
	(0.05)	(22.01)	(1.74)	(8.00)	(25.59)	(0.07)	(32.04)	(1.63)	(11.51)	(30.44)
Panel B: Marginal effects for long-run aid coefficients										
Policy at 10th percentile			−0.33***					−0.28***		
			(0.08)					(0.08)		
Policy at mean			−0.21**					−0.19***		
			(0.08)					(0.07)		
Policy at 90th percentile			−0.11					−0.12		
			(0.10)					(0.09)		

Notes: "EC" refers to as the convergence coefficient (Error-Correction term). Due to data availability, the annual sample is between 1984 and 2012. Region dummies are also dropped with PMG method; hence, there is no difference between BD and ELR specifications in this table. Robustness standard errors are reported in parentheses. Time fixed effects are included in all regressions.

*$p < .1$.
**$p < .05$.
***$p < .01$.

satisfactory alternatives. For example, we use two United Nations measures, voting alignment and membership on the United Nations Security Council (Dreher and Sturm 2012; Dreher, Sturm, and Vreeland 2011). Both variables are associated with political alignment of donors and more aid dollars to recipient countries. Using both instruments together, we find no significant interaction term coefficients. However, our specifications do not pass the Cragg-Donald test (Cragg and Donald 1993). These additional results are available upon request.[27]

We also attempt to use a new instrument for aid, an indicator variable equal to one if a country exceeds the International Development Association income threshold (Galiani et al. 2014). However, this method does not provide sufficient observations in our sample.

E. Alternative Measures of Well-Being

Early foreign aid studies, including BD and ELR, focus strictly on a macro perspective examining if foreign aid could increase economic growth. Following the millennium development goals and SDGs, a more micro perspective is adopted to include social dimensions such as poverty alleviation (Briggs 2017; Chong, Gradstein, and Calderon 2009; Cogneau and Naudet 2007; Guillaumont and Wagner 2014; Mosley, Hudson, and Verschoor 2004), income inequality (Asongu and Nwachukwu 2017; Chong, Gradstein, and Calderon 2009; Holtham and Hazelwood 2010), unemployment rate (Page and Shimeles 2015), and human development (Breitwieser and Wick 2016; Furukawa 2016; Gillanders 2016; Watkins et al. 2005; Williamson 2008). To add to this current discussion, we shift our focus from growth to different measures of well-being.

Specifically, we use four different social measures of well-being as our dependent variables: the human development index (HDI), which measures life expectancy, education, and income per capita (United Nations Development Programme 2018); income inequality measured by the GINI coefficient (World Bank Group 2018); the poverty rate measured as share of population at national poverty line (World Bank Group 2018); and the unemployment rate (WDI 2018/International Labour Organization 2017). If foreign aid positively influences social well-being, it should positively correlate with HDI but negatively relate to income inequality, poverty rate, and unemployment rate.

To test this associate, we employ system GMM as the estimation technique in order to minimize endogeneity. Due to data availability, we are restricted to the 1990–2013 sample. Table 10 presents the results for all countries and low-income countries (BD and

27. This is due to three endogenous regressors in the first stage regression, aid, aid*policy, and aid2*policy (Stock, Wright, and Yogo 2002). Specifications using BD's original instruments from both BD/ELR's original samples and our updated sample, also fail the Cragg-Donald test.

TABLE 10

Alternative Dependent Variables: Measures of Social Well-Being, Outliers Included, 1990–2013

		All Countries				Lower Income			
		4/HDI	**4/Gini**	**4/Poverty**	**4/Unemp**	**7/HDI**	**7/Gini**	**7/Poverty**	**7/Unemp**
Panel A: 1990–2013 full sample coefficients for aid, policy, aid*policy and aid^2*policy									
Aid	BD specification	−0.03	3.29	105.69	3.29	0.33	2.63	20.47	2.63
		(0.80)	(6.39)	(121.62)	(6.39)	(0.36)	(10.86)	(84.18)	(10.86)
	ELR specification	−0.47	4.17	118.64	4.17	0.18	−8.96	14.78	−8.96
		(0.93)	(7.41)	(137.35)	(7.41)	(0.29)	(19.16)	(51.72)	(19.16)
Policy	BD specification	−0.89	0.78	2.62	0.78	0.14	0.12	1.77	0.12
		(1.85)	(0.54)	(3.13)	(0.54)	(1.76)	(0.83)	(2.80)	(0.83)
	ELR specification	−1.98	0.83	2.73	0.83	0.37	−0.75	0.35	−0.75
		(2.15)	(0.62)	(3.62)	(0.62)	(0.60)	(3.38)	(2.52)	(3.38)
Aid*policy	BD specification	0.07	−0.13	−2.83	−0.13	−0.51	−0.10	−0.71	−0.10
		(1.26)	(0.14)	(3.13)	(0.14)	(0.56)	(0.21)	(2.12)	(0.21)
	ELR specification	0.76	−0.16	−3.17	−0.16	−0.29	0.21	−0.49	0.21
		(1.46)	(0.16)	(3.53)	(0.16)	(0.44)	(0.44)	(1.32)	(0.44)
Aid^2*policy	BD specification	−0.00	0.00	0.03*	0.00	−0.00	0.00	0.02**	0.00
		(0.00)	(0.01)	(0.02)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
	ELR specification	−0.00**	0.00	0.02*	0.00	0.00	−0.00	0.01	−0.00
		(0.00)	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
Number of observations	BD specification	251	226	117	226	162	141	83	141
	ELR specification	251	226	117	226	147	126	72	126
Number of countries	BD specification	65	59	45.00	59	43	39	32.00	39
	ELR specification	65	59	45	59	37	33	28	33
AR(2) test *p* value	BD specification	.06	.93	.16	.93	.47	.67	.20	.67
	ELR specification	.13	.91	.16	.91	.97	.91	.30	.91
Hansen-J *p* value	BD specification	.39	.60	.81	.60	.03	.25	.42	.25
	ELR specification	.42	.63	.88	.63	.49	.12	.08	.12
Panel B: 1990–2013 full sample, coefficients for marginal effects of aid									
Policy at 10th percentile	BD specification	0.01	−1.46	1.49	−1.46	0.01	−0.69	−3.22	−0.69
		(0.01)	(2.17)	(5.31)	(2.17)	(0.01)	(2.68)	(7.31)	(2.68)
	ELR specification	0.01	−1.45	1.65	−1.45	0.00	−0.80	−1.93	−0.80
		(0.01)	(2.17)	(5.84)	(2.17)	(0.00)	(1.82)	(3.50)	(1.82)
Policy at mean	BD specification	0.01	−1.79	−3.55*	−1.79	0.00	−1.07	−4.74	−1.07
		(0.01)	(2.07)	(2.08)	(2.07)	(0.00)	(2.00)	(3.64)	(2.00)
	ELR specification	0.01	−1.83	−3.51*	−1.83	0.00	−0.30	−2.64	−0.30
		(0.01)	(2.02)	(1.93)	(2.02)	(0.00)	(1.20)	(2.87)	(1.20)
Policy at 90th percentile	BD specification	0.01	−1.95	−7.52	−1.95	−0.00	−1.20	−5.55	−1.20
		(0.02)	(2.04)	(5.50)	(2.04)	(0.01)	(1.83)	(3.45)	(1.83)
	ELR specification	0.02	−2.01	−7.58	−2.01	0.00	−0.08	−3.14	−0.08
		(0.01)	(1.98)	(5.24)	(1.98)	(0.01)	(1.17)	(3.28)	(1.17)

Notes: GMM estimation technique; full country sample, 1990–2013, BD and ELR regressions 4 and 7. Aid measure is EDA. HDI is Human Development Index; GINI is GINI coefficient; Poverty is national standard poverty rate; Unemp is unemployment rate. Robustness standard errors are reported in parentheses. Time fixed effects are included in all regressions. See Table S3 for detailed variable description.

*$p < .1$.

**$p < .05$.

***$p < .01$.

ELR specifications, including outliers). Across all specifications, the coefficients for aid*policy are statistically insignificant.[28] The marginal effects are also insignificant with one exception: foreign aid may reduce the poverty rate in countries with an average level of policy; however, the coefficient is significant at the 10% level. When outliers are excluded, the significance disappears. We do not tabulate the specifications excluding outliers because they are qualitatively similar to those reported (available upon request).

28. We retest these correlations with sectoral aid to show that the results are not sensitive to measurement of aid. We match specific sectoral aid to each dependent variable. For HDI and GINI aid is ODA to social infrastructure. For Poverty and Unemp, aid is the sum of ODA to economic infrastructure and production sectors. All aid*policy coefficients are insignificant. We find three significant marginal effects indicating that aid may increase income inequality at each level of policy. For example, aid's conditional impact for low income, good policy countries is positive and highly significant. We also find for low income, good policy countries aid may reduce unemployment (10% significance). Otherwise, the results are qualitatively unchanged, and not tabulated to save space.

VI. CONCLUSION

Collectively, our results support ELR more so than BD's claim that aid can increase growth in a good policy environment. Our tests indicate that BD's results are driven by county-year sample, choice of GDP measurement, and

estimation selection. In fact, we are unable to replicate BD's original findings unless we reintroduce their unique observations. This finding reflects the data sensitive feature of the aid-growth literature (Roodman 2007a).

Donors, researchers, and the aid community emphasize that in order to make aid effective at achieving its targeted goals, donors need to be more selective in allocating aid to countries with better institutions (Dollar and Levin 2006; Easterly and Pfutze 2008; Hagen 2015; OECD 2005, 2008; World Bank Group 2005, 171). Even if donors become more selective and give aid to better governed countries, our work suggests that aid is unlikely to achieve economic growth or other goals outlined by the SDGs. Worse, we find evidence suggesting that aid may hinder economic and social development in recipient countries. Combined, our findings challenge the calling for more aid, as aid may not be "fundamental to equitable 'progress' for 'all'" (United Nations 2016, 44).

Economic reasoning tells us that countries most in need of aid are in need because they lack growth-enhancing institutions (Bauer 1972 Leeson 2008). Thus, where aid is needed it might not have its intended effects due to corruption and lack of accountability. On the contrary, in countries with sound economic and political institutions, aid may not be needed. This interpretation of the aid selectivity literature is directly applicable to our findings.

Hansen and Tarp (2000, 394) warn against using single-cause explanations and simple aid allocation rules to guide policymakers. Our empirical exercises support this warning. In addition, we encourage academics and policymakers to not solely rely on empirical results to guide policy. Instead, economic reasoning should remain central to any policy recommendation. Reexamining the aid-policy-growth debate remains an important undertaking as policymakers continue to operate as if aid can be made effective if given under the 'right' conditions.

REFERENCES

Achta, M., M. O. Toman, and T. Rainer. "Corrupt Governments Do Not Receive More State-to-State Aid: Governance and the Delivery of Foreign Aid through Non-State Actors." *Journal of Development Economics*, 114, 2015, 20–33.

Alesina, A., A. Devleeschauwer, W. Easterly, S. Kurlat, and R. Wacziarg. "Fractionalization." *Journal of Economic Growth*, 8, 2003, 155–94.

Ali, A. M., and H. S. Isse. "An Empirical Analysis of the Effect of Aid on Growth." *International Advances in Economic Research*, 11(1), 2005, 1–11.

Alvi, E., D. Mukherjee, and E. K. Shukralla. "Aid, Policies, and Growth in Developing Countries: A New Look at the Empirics." *Southern Economic Journal*, 2008, 693–706.

Arndt, C., S. Jones, and F. Tarp. "Aid, Growth, and Development: Have We Come Full Circle?" *Journal of Globalization and Development*, 1(2), 2010.

———. "What Is the Aggregate Economic Rate of Return to Foreign Aid?" *The World Bank Economic Review*, 30(3), 2015, 446–74.

Asongu, S. A., and J. C. Nwachukwu. "Foreign Aid and Inclusive Development: Updated Evidence from Africa, 2005–2012." *Social Science Quarterly*, 98(1), 2017, 282–98.

Banks, A. S., and K. A. Wilson. *Cross-National Time-Series Data Archive*. Binghamton, NY: Databanks International, 2016.

Bauer, P. T. "Dissent on Development: Studies and Debate in Development Economics." *International Affairs*, 48(4), 1972, 650–2.

Bazzi, S., and M. Clemens. "Blunt Instruments: On Establishing the Causes of Economic Growth." Working Paper No. 171, Center for Global Development, 2009.

Boone, P. "Politics and the Effectiveness of Foreign Aid." *European Economic Review*, 40, 1996, 289–329.

Breitwieser, A., and K. Wick. "What We Miss by Missing Data: Aid Effectiveness Revisited." *World Development*, 78, 2016, 554–71.

Briggs, R. C. "Does Foreign Aid Target the Poorest?" *International Organization*, 71(1), 2017, 187–206.

Brumm, H. J. "Aid, Policies, and Growth: Bauer Was Right." *Cato Journal*, 23(2), 2003, 1167–74.

Burman, L. E., W. R. Reed, and J. Alm. "A Call for Replication Studies." *Public Finance Review*, 38(6), 2010, 787–93.

Burnside, C., and D. Dollar. "Aid, Policies, and Growth." *American Economic Review*, 90(4), 2000, 847–68.

———. "Aid, Policies, and Growth: Reply." *American Economic Review*, 94(3), 2004, 781–4.

Caselli, F., and J. Feyrer. "The Marginal Product of Capital." *Quarterly Journal of Economics*, 122(2), 2007, 535–68.

Chang, C. C., E. Fernandez-Arias, and L. Serven. "Measuring Aid Flows: A New Approach." World Bank Working Paper No. 2050, 1999.

Chatelain, J., and K. Ralf. "Spurious Regressions and Near-Multicollinearity, with an Application to Aid, Policies and Growth." *Journal of Macroeconomics*, 39, 2014, 85–96.

Chong, A., M. Gradstein, and C. Calderon. "Can Foreign Aid Reduce Income Inequality and Poverty?" *Public Choice*, 140(1–2), 2009, 59–84.

Clemens, M. A., S. Radelet, R. R. Bhavnani, and S. Bazzi. "Counting Chickens When They Hatch: Timing and the Effects of Aid on Growth." *The Economic Journal*, 122(561), 2012, 590–617.

Cogneau, D., and J. D. Naudet. "Who Deserves Aid? Equality of Opportunity, International Aid, and Poverty Reduction." *World Development*, 35(1), 2007, 104–20.

Collier, P., and J. Dehn. "Aid, Shocks, and Growth." World Bank Working Paper No. 2688, 2001.

Cragg, J. G., and S. G. Donald. "Testing Identifiability and Specification in Instrumental Variable Models." *Econometric Theory*, 9(2), 1993, 222–40.

Dalgaard, C., and H. Hansen. "On Aid, Growth and Good Policies." *Journal of Development Studies*, 37(6), 2001, 17–41.

Dalgaard, C., H. Hansen, and F. Tarp. "On the Empirics of Foreign Aid and Growth." *The Economic Journal*, 114(496), 2004, 191–216.

Dayton-Johnson, J., and J. Hoddinott. "Aid, Policies, and Growth, Redux." Mimeo, Department of Economics, Dalhousie University, 2003.
Deaton, A. S. "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development." National Bureau of Economic Research, No. w14690, 2009.
Djankov, S., J. G. Montalvo, and M. Reynal-Querol. "The Curse of Aid." *Journal of Economic Growth*, 13(3), 2008, 169–94.
Dollar, D., and V. Levin. "The Increasing Selectivity of Foreign Aid, 1984-2003." *World Development*, 34, 2006, 2034–46.
Doucouliagos, H., and M. Paldam. "Conditional Aid Effectiveness: A Meta-Study." *Journal of International Development*, 22(4), 2010, 391–410.
Dreher, A., and S. Langlotz. "Aid and Growth. New Evidence Using an Excludable Instrument." CESifo Working Paper Series No. 5515, 2015.
Dreher, A., and J. Sturm. "Do the IMF and the World Bank Influence Voting in the UN General Assembly?" *Public Choice*, 151(1–2), 2012, 363–97.
Dreher, A., J. Sturm, and J. R. Vreeland. "Buying Votes and International Organizations." Working Paper, 2011.
Dreher, A., V. Eichenauer, and K. Gehring. "Geopolitics, Aid and Growth: The Impact of UN Security Council Membership on the Effectiveness of Aid." *The World Bank Economic Review*, 32(2), 2016, 268–86.
Dunning, T. "Conditioning the Effects of Aid: Cold War Politics, Donor Credibility, and Democracy in Africa." *International Organization*, 58(2), 2004, 409–23.
Easley, R. W., C. S. Madden, and V. Gray. "A Tale of Two Cultures: Revisiting Journal Editors' Views of Replication Research." *Journal of Business Research*, 66(9), 2013, 1457–9.
Easterly, W. "Can Foreign Aid Buy Growth?" *Journal of Economic Perspectives*, 17(3), 2003, 23–48.
Easterly, W., and T. Pfutze. "Where Does the Money Go? Best and Worst Practices in Foreign Aid." *Journal of Economic Perspectives*, 22(2), 2008, 29–52.
Easterly, W., and C. R. Williamson. "Rhetoric versus Reality: The Best and Worst of Aid Agency Practices." *World Development*, 39(11), 2011, 1930–49.
Easterly, W., R. Levine, and D. Roodman. "New Data, New Doubts: A Comment on Burnside and Dollar's 'Aid, Policies, and Growth'." *American Economic Review*, 94(3), 2004, 774–80.
Evanschitzky, H., and J. S. Armstrong. "Research with In-Built Replications: Comment and Further Suggestions for Replication Research." *Journal of Business Research*, 66, 2013, 1406–8.
Feenstra, R. C., R. Inklaar, and M. P. Timmer. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10), 2015, 3150–82.
Frot, E., A. Olofsgard, and M. Perrota. "Aid Effectiveness in Times of Political Change: Lessons from Post-Communist Transition." *World Development*, 56, 2014, 127–38.
Furukawa, M. "Aid Fragmentation and Effectiveness for Infant and Child Mortality and Primary School Completion," in *The Fragmentation of Aid*, edited by S. Klingebiel, T. Mahn, and M. Negre. London: Palgrave Macmillan, 2016, 165–84.
Galiani, S., S. Knack, L. C. Xu, and B. Zou. "The Effect of Aid on Growth: Evidence from a Quasi-Experiment." *Journal of Economic Growth*, 2014, 1–33.
Gillanders, R. "The Effects of Foreign Aid in Sub-Saharan Africa." *The Economic and Social Review*, 47(3), 2016, 339–60.
Greenhill, R., A. Prizzon, and A. Rogerson. "The Age of Choice: Developing Countries in the New Aid Landscape," in *The Fragmentation of Aid*, edited by S. Klingebiel, T. Mahn, and M. Negre. London: Palgrave Macmillan, 2016, 137–51.
Griffin, K. "Foreign Aid after the Cold War." *Development: Critical Concepts in the Social Sciences*, 3, 2000, 202.
Guillaumont, P., and L. Chauvet. "Aid and Performance: A Reassessment." *Journal of Development Studies*, 37(6), 2001, 66–92.
Guillaumont, P., and L. Wagner. "Aid Effectiveness for Poverty Reduction: Lessons from Cross-Country Analyses, with a Special Focus on Vulnerable Countries." *Revue d'économie du développement*, 22(HS01), 2014, 217–61.
Gwartney, J., R. Lawson, and J. Hall. "Economic Freedom of the World Annual Report." *The Fraser Institute*, 2016.
———. "Economic Freedom of the World Annual Report." *The Fraser Institute*, 2017.
Hagen, R. J. "Dancing to the Donors' Tune? Policy Choice in Aid-Dependent Countries." *Scandinavian Journal of Economics*, 117, 2015, 126–63.
Hall, J. C., and R. A. Lawson. "Economic Freedom of the World: An Accounting of the Literature." *Contemporary Economic Policy*, 32(1), 2014, 1–19.
Hansen, H., and F. Tarp. "Aid Effectiveness Disputed," in *Foreign Aid and Development: Lessons Learnt and Directions for the Future*, edited by F. Tarp. London: Routledge, 2000, 103–28.
———. "Aid and Growth Regressions." *Journal of Development Economics*, 64(2), 2001, 547–70.
Holtham, G., and A. Hazelwood. *Aid and Inequality in Kenya: British Development Assistance to Kenya*. London: Routledge, 2010.
Hsieh, C., and P. J. Klenow. "Relative Prices and Relative Prosperity." *American Economic Review*, 97(3), 2007, 562–85.
Hudson, J., and P. Mosley. "Aid Policies and Growth: In Search of the Holy Grail." *Journal of International Development*, 13(7), 2001, 1023–38.
International Labour Organization. *International Labour Organization Database, ILO Modelled Estimates*. Geneva, Switzerland: ILO, 2017.
Islam, M. N. "Regime Changes, Economic Policies and the Effect of Aid on Growth." *Journal of Development Studies*, 41(8), 2005, 1467–92.
Juselius, K., N. F. Møller, and F. Tarp. "The Long-Run Impact of Foreign Aid in 36 African Countries: Insights from Multivariate Time Series Analysis." *Oxford Bulletin of Economics and Statistics*, 76(2), 2014, 153–84.
Kohama, H., Y. Sawada, and H. Kono. "Aid, Policies and Growth Reconsidered." Working Paper, 2003.
Leeson, P. T. "Escaping Poverty: Foreign Aid, Private Property, and Economic Development." *Journal of Private Enterprise*, 23(2), 2008, 39–64.
Lensink, R., and H. White. "Are There Negative Returns to Aid?" *Journal of Development Studies*, 37(6), 2001, 42–65.
Lu, S., and R. Ram. "Foreign Aid, Government Policies, and Economic Growth: Further Evidence from Cross-Country Panel Data for 1970-1993." *International Economics*, 54, 2001, 15–29.
McCullough, B. D., K. A. McGeary, and T. D. Harrison. "Do Economics Journal Archives Promote Replicable Research?" *Canadian Journal of Economics*, 41(4), 2008, 1406–20.
Mekasha, T. J., and F. Tarp. "A Meta-Analysis of Aid Effectiveness: Revisiting the Evidence." World Institute for Development Economic Research (UNU-WIDER) Working Paper No. 044, 2018.

Minasyan, A. "Your Development or Mine? Effects of Donor–Recipient Cultural Differences on the Aid-Growth Nexus." *Journal of Comparative Economics*, 44(2), 2016, 309–25.

Mosley, P., J. Hudson, and A. Verschoor. "Aid, Poverty Reduction and the 'New Conditionality'." *The Economic Journal*, 114(496), 2004, F217–43.

OECD. *Paris Declaration on Aid Effectiveness*. Paris, France: OECD, 2005.

OECD. *Third High Level Forum on Aid Effectiveness.* Accra, Ghana, September 2-4, 2008.

Page, J., and A. Shimeles. "Aid, Employment and Poverty Reduction in Africa." *African Development Review*, 27(S1), 2015, 17–30.

Pesaran, M. H., Y. Shin, and R. P. Smith. "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels." *Journal of the American Statistical Association*, 94, 1999, 621–34.

PRS Group. "International Country Risk Guide." New York: PRS Group, 2016.

Qian, N. "Making Progress on Foreign Aid." *Annual Review of Economics*, 7(1), 2015, 277–308.

Rajan, R. G., and A. Subramanian. "Aid and Growth: What Does the Cross-Country Evidence Really Show?" *Review of Economics and Statistics*, 90(4), 2008, 643–65.

Ram, R. "Roles of Bilateral and Multilateral Aid in Economic Growth of Developing Countries." *Kyklos*, 56(1), 2003, 95–110.

Ram, R., and S. Ural. "Comparison of GDP Per Capita Data in Penn World Table and World Development Indicators." *Social Indicators Research*, 116(2), 2014, 639–46.

Roodman, D. "The Anarchy of Numbers: Aid, Development, and Cross-Country Empirics." *The World Bank Economic Review*, 21(2), 2007a, 255–77.

———. "Macro Aid Effectiveness Research: A Guide for the Perplexed." Working Paper, Center for Global Development, 2007b.

Sachs, J., and A. Warner. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity*, 1995(1), 1995, 1–118.

Stock, J. H., J. H. Wright, and M. Yogo. "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments." *Journal of Business & Economic Statistics*, 20(4), 2002, 518–29.

Tan, K. Y. "A Pooled Mean Group Analysis on Aid and Growth." *Applied Economics Letters*, 16(16), 2009, 1597–601.

Tashrifov, Y. "Foreign Financial Aid, Government Policies and Economic Growth: Does the Policy Setting in Developing Countries Matter?" *Zagreb International Review of Economics and Business*, 15(1), 2012, 1–22.

United Nations. *The Sustainable Development Goals Report*. New York: UN, 2016.

United Nations Development Programme. *Human Development Data*. New York: UN, 2018.

Verschoor, A., and A. Kalwij. "Aid, Social Policies and Pro-Poor Growth." *Journal of International Development*, 18(4), 2006, 519–32.

Watkins, K., H. Fu, R. Fuentes, A. Ghosh, C. Giamberardini, C. Johansson, C. Kuonqui, A. Montes, D. Stewart, C. Ugaz, and S. Yaqub. *Human Development Report 2005*: *International Cooperation at a Crossroads: Aid, Trade and Security in an Unequal World*. New York: UNDP, 2005.

Williamson, C. R. "Foreign Aid and Human Development: The Impact of Foreign Aid to the Health Sector." *Southern Economic Journal*, 2008, 188–207.

World Bank Group. "Review of World Bank Conditionality." Working Paper, World Bank, 2005.

———. *World Development Indicators*. Washington, D.C.: World Bank Group, 2018.

Yontcheva, B., and N. Masud. "Does Foreign Aid Reduce Poverty? Empirical Evidence from Nongovernmental and Bilateral Aid." International Monetary Fund Working Paper No. 05/100, 2005.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Differences in Samples Between BD, ELR and the New Data 1962–2013

Table S2. Country Differences in BD/ELR Samples and New Data, Full Sample, 1990–2013/1962–2013

Table S3. Variable Descriptions

Table S4. Summary Statistics, 1962–2013 Full Sample

Table S5. Outliers Excluded from Regressions

Table S6. Country List of BD and ELR Regional Dummy Variables

Table S7. Countries with Trade Openness Status Changed Between BD/ELR Samples and New Sample

Table S8. Alternative Method: Country-Specific Short-Run Coefficients, Pooled Mean Group with Annual Data, Outliers Included, 1984–2012