NBER WORKING PAPER SERIES

THE EFFECT OF AID ON GROWTH:
EVIDENCE FROM A QUASI-EXPERIMENT

Sebastian Galiani
Stephen Knack
Lixin Colin Xu
Ben Zou

Working Paper 22164
http://www.nber.org/papers/w22164

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2016

We are grateful for valuable comments by several anonymous referees, George Clarke, Michael Clemens, Yingyao Hu, Aart Kraay, Martin Ravallion, and seminar participants at the 2013 Southern Economics Association Annual Meeting, the University of Toronto, Xiamen University, and the Center for Global Development. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research. Ben Zou's time was partially covered by a World Bank short-term consultant contract.

The Effect of Aid on Growth: Evidence from a Quasi-Experiment
Sebastian Galiani, Stephen Knack, Lixin Colin Xu, and Ben Zou
NBER Working Paper No. 22164
April 2016
JEL No. O1,O4

ABSTRACT

The literature on aid and growth has not found a convincing instrumental variable to identify the causal effects of aid. This paper exploits an instrumental variable based on the fact that since 1987, eligibility for aid from the International Development Association (IDA) has been based partly on whether or not a country is below a certain threshold of per capita income. The paper finds evidence that other donors tend to reinforce rather than compensate for reductions in IDA aid following threshold crossings. Overall, aid as a share of gross national income (GNI) drops about 59 percent on average after countries cross the threshold. Focusing on the 35 countries that have crossed the income threshold from below between 1987 and 2010, a positive, statistically significant, and economically sizable effect of aid on growth is found. A one percentage point increase in the aid to GNI ratio from the sample mean raises annual real per capita growth in gross domestic product by approximately 0.35 percentage points.

Sebastian Galiani
Department of Economics
University of Maryland
3105 Tydings Hall
College Park, MD 20742
and NBER
galiani@econ.umd.edu

Stephen Knack
World Bank
1818 H St. NW
Washington DC 20433
sknack@worldbank.org

Lixin Colin Xu
MC 3-420,
World Bank
1818 H Street, N.W.
Washington, DC 20433
lxu1@worldbank.org

Ben Zou
Michigan State University
486 W. Circle Drive 110
Marshall-Adams Hall East
Lansing, MI 48824
zouben@gmail.com

1. Introduction

Whether foreign aid causes economic growth in recipient countries is a highly debated research question. Identification of the causal effect of aid on growth has been elusive so far due to the endogeneity of aid in growth models. An instrumental variable is needed to address these problems. However, as Clemens et al. (2012) conclude in their recent assessment: "the aid-growth literature does not currently possess a strong and patently valid instrumental variable with which to reliably test the hypothesis that aid strictly causes growth."

In this paper we contribute to this literature by instrumenting, in an economic growth equation, the endogenous aid variable exploiting a plausible quasi-experiment created by the income threshold set by IDA (International Development Association), the World Bank's program of grants and concessionary loans to low-income countries. We exploit this new instrument to investigate the causal effect of aid on growth.

This income threshold has been used as a key criterion in allocating scarce IDA resources since 1987, and is adjusted annually only to take into account inflation. Other major donors also appear to use the IDA threshold as an informative signal about where development aid is most needed, and we show that total aid declines significantly once a recipient country crosses the IDA income threshold from below. The IDA threshold is nevertheless an arbitrary income level that does not necessarily represent any structural change in economic growth. Threshold crossing is thus a plausibly valid instrumental variable for aid over years for a recipient country in a panel data model that controls for initial income levels and also includes country and period effects (we group years into 8 three-year periods).

The main concern with successfully identifying an internally consistent estimate is that some countries might cross the threshold by experiencing a series of large positive shocks that are eventually reversed, making the exclusion restriction invalid. We argue that this concern is largely misplaced. On average, our sample of countries grew faster after crossing the IDA threshold than before, consistent with the fact that developing countries in general exhibited better performance in the latter part of our 1987-2010 sample period. Moreover, our empirical growth model accounts for the differential timing at which countries cross the IDA threshold. In our analysis we exploit only the data for the countries that cross the threshold from below during the period studied while our growth model allows countries to grow at different rates over time (by allowing for country-specific effects on growth and by allowing conditional convergence). Additionally, countries start at the beginning of the sample period from different levels below the threshold. Hence, the differential timing at which countries cross the IDA threshold from below exploited for identification in our study does not have to be driven by unobservable shocks.[1] We further investigate this threat to our identification strategy by implementing a battery of tests and robustness checks. In particular,

[1] This assertion applies to countries crossing from below, but perhaps not to the smaller set of countries crossing the threshold from above; at least they display systematic negative growth rates. Moreover, IDA policies are premised on the expectation of growth and eventual graduation, and instances of crossing the threshold from above are dealt with in a more ad hoc fashion. In such cases, IDA has often been "reluctant to accept renewed claims on its scarce concessional resources, especially if this would reward poor performance" (Kapur et al., 1997). Crossing from above and from below may therefore have highly asymmetric effects on aid, and in turn on subsequent growth. In particular, crossing from above would likely be a weaker instrument for aid, and its effects would be less precisely estimated due to the smaller sample of relevant countries. We therefore focus only on the countries that cross the IDA threshold from below.

we fail to reject the null hypothesis of no serial autocorrelation in the error term of the model implied by the threat to the validity of our instrument posed above. We also show that results are robust to controlling for one and two period lags of growth, which would not be the case if the country receives a transitory positive shock around the time of crossing which is eventually reversed. Importantly, we use an alternative instrumental variable based on a smoothed income trajectory which is unaffected by country specific idiosyncratic shocks and obtain similar results. All in all, these and other robustness checks do not point toward rejecting our identification strategy.

Using a sample of 35 countries that crossed the IDA threshold from below between 1987 and 2010, we find that a one percent increase in the aid to GNI ratio raises the annual real per capita short term GDP growth rate by 0.031 percentage points. The mean aid to GNI ratio at the crossing is 0.09, so a one percentage point increase in the aid to GNI ratio raises annual real per capita GDP growth by approximately 0.35 percentage points. This effect is about 1.75 times as large as those reported by Clemens et al. (2012). Using OLS with fixed effects and lagging aid by a period, they find that a one percentage point increase in aid/GDP (at aid levels similar to our sample mean) is followed by at most a 0.2 percentage-point increase in growth of real GDP per capita. We find similar-sized effects in our sample, without instrumenting for aid but merely lagging it by a period and including fixed effects. We also present evidence consistent with the possibility that OLS estimates suffer from attenuation bias due to measurement error in aid, which is exacerbated when the variability in the aid to GNI ratios is exploited to identify the effect of aid on growth in a fixed effects or first-differenced growth equation commonly used in the literature. Thus, one should expect that a valid 2SLS produces larger estimates than OLS. Naturally, this could also be the case if growth affects aid levels.

The sizable effect of aid on growth we find may be attributable in part to the fact that our sample consists of low-income countries that successfully crossed the IDA threshold at some point between 1987 and 2010. Aid may have been more effective in these countries--e.g. due to better economic policies and lower corruption -- than in countries remaining below the threshold. Relative to middle-income countries, countries in our sample likely face more stringent financial constraints, which again should make aid effects stronger. Furthermore, it is possible that reductions in aid after crossing the IDA threshold may have a larger negative effect on growth than any positive effects from aid increases, due to adjustment costs. We provide suggestive evidence, however, that our results might have meaningful external validity to the remaining poor countries as they grow closer to the IDA threshold.

A simple growth accounting exercise and the coefficient on aid in investment regressions suggest that investment could be an important channel through which aid affects growth. We show that the investment rate drops following the reduction in aid. Increasing the aid to GNI ratio by one percentage point increases the investment to GDP ratio by 0.54 percentage points, although this coefficient is generally not significant. The magnitude of the effects on growth and investment is consistent with the average capital stock to GDP ratio for the sample countries, which we estimate to be approximately two.

As in most of the literature relying on panel data covering a short period of time, we estimate the short-run effect of aid on growth, an effect that mainly operates through physical investment. In the long run, aid could affect growth through several other channels, but its identification requires exogenous changes in aid over a very long period of time. Our instrument does not provide such exogenous variability to estimate that parameter.[2]

The rest of the paper is organized as follows. Section 2 briefly reviews past studies testing the causal effect of aid on growth. Section 3 describes the data and the sample. Section 4 presents the effect of IDA threshold-crossing on the volume of aid received. Section 5 presents the empirical model and the baseline results. Section 6 explores some mechanisms and Section 7 discusses the external validity of our results. Section 8 concludes.

2. Previous Aid-Growth Studies

Following the influential studies by Boone (1996) and Burnside and Dollar (2000), many others have emerged, but a basic consensus is still absent. Easterly et al. (2004) show that the key finding of Burnside and Dollar (2000) – namely that aid contributes to growth but only where economic policies are favorable – is not robust to the use of an updated and enlarged dataset. Rajan and Subramanian (2008) and Arndt, Jones and Tarp (2010) are only two of many recent papers that review the bulk of the existing literature yet arrive at differing conclusions.[3]

Identifying the causal effect of aid on growth is fraught with difficulties. First, aid relative to GNI is likely measured with error. The problem of measurement error is exacerbated as the estimated model is often demeaned or first differenced to eliminate the country fixed effects. Second, identification

---

[2] Regressing the average growth rate over a long period of time on the average aid in that period does not identify the long-term effect of aid on economic growth, even if aid were exogenous in that equation.

[3] Also see Temple (2010) and Qian (2015) for two comprehensive reviews of theory, evidence, and practice of foreign aid.

might be confounded by unobserved factors that determine both economic growth and aid. Third, growth itself could also affect aid. In response to these potential problems, previous studies have introduced different instrumental variables to identify the causal effect of aid on growth. In this section we briefly review the major identification strategies exploited in the literature.

Studies using cross-sectional data often rely on population size, economic policies and donor-recipient political connections as instruments for aid (e.g., Boone, 1996; Burnside and Dollar, 2000; Rajan and Subramanian, 2008). These instruments are likely to violate the exclusion restriction, as they are correlated with observable and plausibly also unobserved country-level characteristics that also contribute to growth (Hansen and Tarp, 2001). For example, population size can affect economic growth through channels other than aid (Bazzi and Clemens, 2013). Donor-recipient ties (e.g. colonization, trade or migration) that are correlated with aid flows can also affect growth indirectly through the institutional environment (Acemoglu, Johnson, and Robinson, 2001) or other channels.

Other studies relying on panel data control for time-invariant determinants of growth by first differencing the data or by conditioning on country fixed effects. These studies focus mainly on the short term effect of aid on growth. Many such studies adopt a dynamic panel model and employ difference or system GMM estimators, instrumenting for current aid with lagged values of income and aid, and with other standard cross-country regressors (e.g., Hansen and Tarp, 2001 and Rajan and Subramanian, 2008). However, recent studies show that GMM estimators of dynamic panel models using all mechanical instruments are unstable and potentially severely biased in finite samples, due to the problem of many and weak instruments (Roodman, 2007, 2009a, 2009b;

Bazzi and Clemens, 2013; Bun and Windmeijer, 2010). System GMM estimators, in addition, use additional instruments and require additional assumptions on exclusion restrictions.

Sharing the spirit of our paper, several recent studies exploit donor-recipient connections interacted with over-time variations in total donor contributions. Werker et al. (2009) instrument for aid with the interaction between the price of oil and a dummy for whether recipient countries are Muslim, and find a small and marginally significant effect of lagged aid on growth using annual data. However, using four-year period averages, they find no significant effect of (contemporaneous) aid on growth. Nunn and Qian (2013) find that humanitarian aid extends civil war, using variation over time in U.S. wheat harvests to instrument for the amount of humanitarian aid that a country receives. Temple and Van de Sijpe (2014) find that aid increases net imports and total consumption (both as shares of GDP) by constructing a synthetic measure of aid by taking each recipient country's share of aid in a donor's aid budget in an initial period and multiplying it by the donor's total aid budget in the current period. Dreher and Langlotz (2015) use the interaction between donor government fractionalization (by party) and a recipient government's probability of receiving aid from the donor to instrument for aid, and find no significant effect of it on growth. Arndt et al. (forthcoming) review a range of aid-growth studies published since 2008 that attempt to address the endogeneity of aid in various ways (e.g.. Clemens et al., 2012 and Bruckner, 2013), and conclude that "the large majority…have found positive impacts," particularly when effects are assessed over longer time periods. While these papers rely mostly on historical relationships between bilateral donors and recipients or natural shocks in grain outputs, we adopt a different identification strategy, exploiting a natural experiment based on declines in aid after developing countries surpass a pre-determined level of GNI per capita. Because instrumental variable-based estimates reflect local average treatment effects for the complier group (Angrist, Imbens and Rubin, 1996), our identification strategy only identifies the effect for low-income countries that experience changes in aid after crossing the IDA graduation threshold.

3. Data and Sample

The data for this study are primarily from two sources. Income, investment, economic growth, and other country characteristics are from the World Bank's World Development Indicators (WDI).[4] Aid data are obtained from the OECD's Development Assistance Committee (DAC).[5] Following much of the previous literature, aid is measured by total net Official Development Assistance (ODA) disbursements as a share of GNI, in current US dollars.[6]

We identify 35 countries that crossed the IDA income threshold from below between 1987 and 2010.[7] Table 1 lists these countries and their years of crossing. For countries that crossed the threshold more than once, in our baseline specification we consider only the first crossing in defining the instrumental variable. We show, however, that our results are robust to changes in this criterion. Following the literature, we smooth out fluctuations

[4] http://databank.worldbank.org/ddp/home.do?Step=12&id=4&CNO=2, accessed and extracted in August, 2012. The WDI dataset is usually updated 4 times a year, and sometimes revises historical data on national income and other variables. Most of the revisions are minor.

[5] Data are from DAC Table 2a, available at http://stats.oecd.org/Index.aspx?DatasetCode=TABLE2A#, accessed in August, 2012.

[6] Both aid and GNI are measured in nominal terms, as is the IDA income threshold.

[7] Sao Tome and Principe crossed the threshold in 2009. It has only 2 periods of data in the sample and is thus automatically dropped from the analysis and hence also from the sample.

in the annual data by using period averages. Due to the length of our panel dataset, and because IDA has a three-year replenishment cycle, we group years into 8 three-year periods that roughly coincide with the IDA replenishment periods.[8] The first period, with data from calendar years 1987-1989, corresponds roughly to IDA8, covering fiscal years 1988-1990 (July 1, 1987 to June 30, 1990). The final period, with data from 2008-2010, roughly corresponds to IDA15 (July 1, 2008 to June 30, 2011). Donors pledge contributions for each replenishment period, rather than annually. Moreover, policies for allocating IDA funds (e.g. the relative weights assigned to poverty, economic policies, and quality of governance) among eligible recipients are often modified between IDA periods but never within an IDA period. For this reason country allocations should be more correlated from one year to the next within an IDA period than across two replenishment periods. The 3-year IDA periods are therefore a natural way of grouping the data. The timing of actual graduations from IDA also tends to coincide with the end of replenishment periods. The baseline sample contains 247 country-period observations.

Online Appendix Table A lists the definitions and data sources for all variables, and Online Appendix Table B presents summary statistics for the baseline sample. Real per capita GDP of the sample countries grew at an average annual rate of 2.9%. ODA equaled about 8% of GNI for a typical country in a typical year in the sample. Of total ODA, about 9% is from IDA,

[8] Many recent panel studies often group years in 4- or 5-year periods. Temple and Van de Sijpe (2014) use 3-year periods. As Clemens et al. (2012, p.594) observe, “The question of when to test for growth impacts plagues the entire growth literature, not just aid-growth research. Empirical research on the determinants of growth cannot escape the selection of a fixed observation period,” but there is no consensus regarding the time intervals over which to study growth.

67% is bilateral aid from DAC countries, 2% is bilateral aid from non-DAC countries, and 23% is from non-IDA multilateral agencies. The mean value of ODA/GNI is 8.5% in periods when countries are under the income threshold (including the periods in which they cross), and 7.4% in post-crossing periods. The mean of IDA/ODA declines from 11.8% to 5.5% post-crossing, the largest relative drop among the four donor groupings.

4. The IDA Threshold and Aid

Beginning in 1987, a major criterion for IDA eligibility has been whether or not a country is below a certain threshold of per capita income, measured in current US dollars. This "operational threshold" was established for the purpose of rationing scarce IDA funds. Prior to 1987, a higher threshold (now called the "historical cutoff") had been in effect, but economic crises in developing countries increased the demand for IDA funds in the early- and mid-1980s, necessitating a new lower cutoff (World Bank, 1989). Figure 1 shows the evolution of the IDA operational threshold converted in current US dollars between 1987 and 2010. It was originally set at $580, and has been adjusted annually only for inflation, as measured by the SDR deflator.[9] By 2010, the threshold had increased to $1175.

The other criterion for IDA eligibility is lack of creditworthiness, defined as the inability "…to service new external debt at market interest rates over the long term" (World Bank, 1989). China and several other countries graduated from IDA (i.e. were declared ineligible for new loans) while they

[9] The SDR ("Special Drawing Rights," the unit of account for the International Monetary Fund) deflator is a weighted average of the GDP deflators for the U.S., Japan, the U.K. and the euro area. As shown in Figure 1, the threshold declined slightly for several years between 1998 and 2002, because the SDR deflator was negative.

were still under the income threshold, because they were deemed creditworthy. Conversely, Bolivia's graduation from IDA was delayed for many years after it crossed the income threshold due to lack of creditworthiness. The World Bank's assessments of country creditworthiness are highly confidential (Moss and Majerowicz, 2012) and are not even available to most staff members. However, small island economies – those with populations below 1.5 million – are presumptively judged as not creditworthy, due to their vulnerability to shocks.

In contrast to threshold crossing, actual graduation from IDA is likely to be endogenous to economic performance, policies and vulnerability to shocks, even when controlling for a continuous measure of per capita income. Graduation itself, as opposed to threshold crossing, would thus not be a valid instrument for aid.

Once a country has exceeded the IDA income threshold and is judged to be creditworthy, it is considered on track for "graduation" from IDA. Allowance is made for the possibility of income fluctuations, so lending volumes typically are reduced (and repayments accelerated) only after a country has remained over the threshold for three consecutive years. Thus, in most cases, threshold crossing will result in reductions of IDA flows beginning in the next replenishment period, not in the current one (World Bank, 2010). The decline in aid from IDA is amplified by similar behavior from other donors. Some agencies, such as the African Development Bank (AfDB) and Asian Development Bank (AsDB), explicitly use the IDA income threshold in their own aid eligibility criteria. Other donors often view crossing the IDA threshold as a signal that countries are in less need of aid and cut their own aid, reinforcing the decline in aid from IDA (Moss and Majerowicz, 2012;

World Bank, 1989). As a result, although IDA contributes less than one-tenth of the total aid to a typical recipient, crossing the IDA threshold may have a sizable effect on total aid.

The relevance of IDA threshold crossing as an instrument for aid can be tested by looking at its effects on total aid and aid from different donors. We distinguish among four groups of donors: IDA, DAC (OECD Development Assistance Committee) bilateral donors, non-DAC bilateral donors, and other multilateral donors. Following the consensus in the literature (Clemens et al., 2012), lagged aid is the main explanatory variable in our growth regressions, so our instrumental variable is a dummy indicating whether the country has crossed the IDA threshold at least two periods earlier. Throughout the paper we use t to represent a specific year and s to represent a specific period, where s includes years $t-2$, $t-1$, and t. We define $Crossing_{i,s-2}$ equal to 1 if a country's first threshold crossing during the sample period took place at least two periods before period s. Otherwise, $Crossing_{i,s-2}$ equals 0. We estimate the following equation:

$$Aid_{j_{is-1}} = \beta_1\, y_{is-1} + \beta_2\, Crossing_{is-2} + \beta_3\, Pop_{is-1} + \lambda_i + \tau_{s-1} + v_{j_{is-1}} \qquad (1)$$

The dependent variable $Aid_{j_{is-1}}$ is the log of average ratio of aid from donor type j to GNI or the log ratio of average total aid (i.e., the sum of aid from all donor sources) to GNI for country i in period $s-1$, that is, $Aid_{j_{is-1}} =$

$\ln[(\sum_{k=3}^{5} \frac{ODA_{j_{it-k}}}{GNI_{it-k}})/3]$.[10] In this analysis we use $Aid_{j_{is-1}}$ as the dependent variable because it is the key explanatory variable when we estimate the effect of aid on economic growth in the current period, as will become clear in the next section. y denotes log real per capita GDP measured in constant 2000 US dollars. y_{is-1} is measured as log real per capita GDP in the second year of the last period $s-1$, and hence it is equal to y_{it-4}. Pop_{is-1} is the log of average population of period $s-1$. $Crossing_{i,s-2}$ is defined as earlier. This second lag is introduced because the IDA graduation process typically begins only three years after a country crosses the threshold, i.e. in the next replenishment period. The crossing status lagged one period relative to aid also allows time for other donors to respond to threshold crossings.

Table 2 reports the results of estimating Equation 1. For Columns 1 to 5, respectively, the dependent variables are the one-period lag of the logarithm of aid share of GNI from (1) IDA, (2) DAC countries, (3) non-DAC countries, (4) multilateral agencies except for IDA, and (5) all donors. To be conservative, we use two alternative methods to conduct statistical inference throughout the paper. We report both robust standard errors clustered at the country level and the standard errors from the clustered wild bootstrap procedure following Cameron et al. (2008); see Appendix A for a more detailed explanation. Either approach yields very similar statistical inferences; for brevity, we will focus our discussion on the clustered standard errors.

We find that following IDA threshold-crossing, IDA flows as a share of GNI ratio dropped, on average, by about 92% (i.e., $1-e^{-2.5}$). Other donors

[10] We follow the convention of the majority of the literature and measure both GNI and ODA in current US dollars, the same units IDA uses to define its income threshold. A minority of studies, such as Boone (1996), use GNI in purchasing power parity terms, however.

also cut their aid substantially. Estimates of the coefficients associated with threshold crossing are negative and large in magnitude. Except for aid from non-DAC donors, the estimated coefficients are statistically significant at conventional levels. The total aid to GNI ratio dropped, on average, by 59% ($i.e., 1 - e^{-0.88}$). Higher income levels are a strong predictor of aid: a one-percent increase in real per capita GDP is associated with reductions in aid of about 8.6 percent from IDA, 1.4 percent from DAC countries, 4.7 percent from non-DAC countries, 2.5 percent from other multilateral agencies, and 1.5 percent for the overall ODA to GNI ratio.

Potentially, any crossing dummy variable $Crossing_{i,s-p}$ with $p \geq 1$ is a valid instrument. Here we rely on the one that best predicts aid/GNI, namely, $Crossing_{i,s-2}$.[11] In reduced-form tests (not reported in tables), we find that $Crossing_{i,s-2}$ has the strongest and most significant effect on growth, of about -2.4 percentage points.

We conduct two further checks. First, results are robust to controlling for a quadratic relationship between aid and log initial income level (results are shown in Online Appendix Table D). The coefficient of the crossing dummy for ODA aid over GNI, for instance, is now -0.94 and statistically significant at the one percent level, similar to that in Table 2 (i.e., -0.876). Second, we conduct a placebo test to ensure that these effects are not a statistical artifact. Specifically, we replace the true IDA threshold value with a false threshold

[11] In Appendix Table C we report the results of including $Crossing_{i,s-1}$, $Crossing_{i,s-2}$, and $Crossing_{i,s-3}$ in the model while otherwise retaining the specification of Equation (1). Column 1 presents the results. Coefficients associated with all three variables are negative and statistically significant. However, the one associated with $Crossing_{i,s-2}$ has the largest t statistic. Reduced-form results are in column (2) of Online Appendix Table C.

equal to 50% of the true value, and re-estimate equation 1 using a threshold-crossing dummy variable based on this false threshold. In the analysis we retain only country-period observations prior to the period in which countries cross the actual threshold, so the regression sample is unaffected by the effect of actually crossing the true threshold.[12] Crossing the false threshold has no significant effect on aid results are reported in Online Appendix Table E).

Another concern is that IDA threshold crossing may not be a good instrument to identify the direct or structural effect of aid on growth, because country leaders may endogenously alter policies to take advantage of potential complementarity (or substitutability) between aid revenues and policies. Ex ante it is unclear whether and how aid might affect the quality of policymaking— e.g., aid could facilitate policy reform if it is used to compensate losers, or worsen policy by stimulating rent seeking (Rodrik, 1996)—and whether country leaders can engineer quick policy changes to have an immediate effect on growth along desired directions. Thus, how aid affects policies is an empirical question. We investigate this potential threat to our identification strategy by estimating equation (1) after replacing aid as the dependent variable by a set of variables measuring policymaking and institutional quality. These variables include measures of civil liberty and political rights from Freedom House, the World Bank CPIA (Country Policy and Institutional Assessment), broad money (M2) as a percentage of GDP, inflation as measured by changes in the GDP deflator, and dummy variables indicating respectively bank, currency and debt crises (see Appendix Table A

[12] For sample countries with per capita GNI always above the fake threshold, the crossing dummy is replaced with 0.

for definitions of these variables). Crossing the IDA income threshold turns out to have no statistically significant effect on any of the policy variables considered (results reported in Online Appendix Table F). We show below that our growth results are also robust to controlling for these variables.

5. Foreign Aid and Economic Growth

5.1 An illustration

The effects of aid should be most pronounced two periods after crossing, since aid volumes drop most precipitously in the period after crossing (see earlier discussion of results in Appendix Table C). To see this, for the group of countries that cross the threshold at least two periods before the end of our sample period, Figure 2 shows the relationship between per capita real GDP growth and the once-lagged log aid to GNI ratio during the second period after each country crosses the threshold. We first-difference the two variables to get rid of the time-invariant effects specific to each country. Almost all countries experienced a significant drop in aid compared to the last period, and larger reductions in aid are associated with larger declines in growth.[13]

5.2 Econometric models

We postulate the following model in order to test the null hypothesis that aid does not affect growth:

$$g_{is} = \beta_1 y_{is-1} + \beta_2 Aid_{is-1} + \boldsymbol{X}_{is} \cdot \beta_3 + \lambda_i + \tau_s + \varepsilon_{is}, \qquad (2)$$

[13] The fitted line has a slope equal to 0.8 and a standard error equal to 0.4. Our aid and growth measures both have some measure of national income in their denominators. In controlling for both income (per capita) and population, however, and measuring them (and aid/GNI) in logs, we minimize the possibility of spurious correlation, produced by errors in national accounts or population estimates (Kronmal, 1993).

where s denotes non-overlapping 3-year periods. Period s includes years $t, t-1, t-2$. y denotes log real per capita GDP. g_{is}, constructed as $(y_{it} - y_{it-3})/3$, is the average log difference of real GDP per capita of country i in period s. y_{is-1} is measured as the log real per capita GDP in the second year of the previous period (i.e., y_{it-4}).[14] We expect β_1, which captures conditional convergence, to be negative. Aid_{is-1} is the log of average aid received by country i as a share of GNI in the previous period.[15] We use the one-period lag of aid instead of contemporaneous aid to allow time for aid to take effect (Clemens et al., 2012). $\boldsymbol{X}_{is}$ is a vector of time-varying variables, including log population, assumed to be strictly exogenous. In the literature, population is almost always on the right hand side of aid allocation regressions (smaller countries receive more aid per capita), and it is commonly on the right hand side in growth regressions (usually with scale effects in mind) as well. While here we show only a parsimonious model of time-varying variables, we show later that our results remain robust to controlling for other time-varying growth determinants. λ_i is the country i fixed effect. τ_{s-1} is the period s fixed effect.

We use three alternative empirical methods in this subsection to estimate

[14] Notice that, by construction, y_{is-1} is not mechanically correlated with the dependent variable. Some studies in the literature use per capita real GDP in purchasing power parity terms to measure income level and to calculate growth (e.g., Boone, 1996). Real per capita GDP based on current exchange rates (in constant dollar terms, and using the Atlas method) and real per capita GDP in PPP terms are highly correlated (at over .95) across countries in our sample. Growth rates constructed from the two versions are essentially the same. We use per capita GDP based on current exchange rates (in constant dollars) because there are fewer missing observations in the WDI database than for the PPP measure. Using instead the PPP measure we obtain almost identical results for our basic specifications in Table 3.

[15] For further justification of why we use the log form, see Appendix B. In addition, the results are essentially unchanged when Aid_{is-1} is measured as the ODA share of GDP.

equation (2). Each method requires particular assumptions. We thus report baseline results and a series of robustness tests using each of the three methods.

The standard way to estimate equation (2) is to eliminate the unobservable country-specific effects, λ_i, by including a set of country dummy variables in the model, which is equivalent to demeaning equation (2) and estimating the transformed equation by OLS. This estimator, however, is potentially inconsistent due to aid being also affected by growth, measurement error, and time-varying unobservable variables. We therefore instrument aid in period $s-1$ $(i.e., Aid_{is-1})$ with a dummy variable indicating whether the country has crossed the IDA threshold by the end of period $s-2$, that is, $Crossing_{i,s-2}$. To address the endogeneity of the initial income level y_{is-1}, we instrument the initial income level with further lags of the income level.

Our instrumental variable is based on per capita (nominal) GNI crossing the IDA threshold two periods earlier. Per capita (nominal) GNI level in period $s-2$ is correlated with the idiosyncratic shock to (real) economic growth of that period, ε_{is-2}. Thus, estimating equation 2 by means of the fixed effects estimator, which first de-means the equation, mechanically introduces a correlation between the instrumental variable and the demeaned error term, $\ddot{\varepsilon}_{is} = (\varepsilon_{is} - \bar{\varepsilon}_i)$. However, if ε_{is} is not serially correlated, the correlation between $\ddot{\varepsilon}_{is}$ and ε_{is-2} will be small if the time dimension of the panel is large. Our sample has 8 periods, which is not considered short in the literature. We also show below that we do not reject the null hypothesis of no serial correlation of the error terms in equation (2).[16] We use two other methods to

[16] When the panel is long, $\ddot{\varepsilon}_{is}$ is less correlated with the error term from a particular period, and the bias will be small. To get a sense of the potential bias in the 2SLS fixed effect model

circumvent this potential problem. The first approach is to first-difference equation (2) and estimating the following equation:

$$\Delta g_{is} = \beta_1 \Delta y_{is-1} + \beta_2 \Delta Aid_{is-1} + \Delta \boldsymbol{X}_{is} \cdot \beta_3 + \Delta \tau_s + \Delta \varepsilon_{is,} \quad (4)$$

In using $\Delta Crossing_{is-2}$ to instrument ΔAid_{is-1}, our identification strategy exploits only the sharp variability in aid at the period after crossing the IDA threshold. Under treatment heterogeneity, both in terms of the effect of threshold-crossing on aid and of the latter on economic growth, this strategy will identify a particular local average causal effect. Alternatively, we can also use just $Crossing_{is-2}$ to instrument for ΔAid_{is-1}.

The validity of the exclusion restriction in this case requires the error terms ε_{is} to be serially uncorrelated, i.e., the instrumental variable will be invalid if the unobservable idiosyncratic error term in the growth equation is serially correlated up to two periods. Below we investigate the validity of this assumption.

In the first differenced model, even if the error terms were *i.i.d.*, the transformed error terms will not be, and will exhibit first-order serial correlation. Standard GMM inference when using optimal weights takes into account this feature while 2SLS does not. Thus, through the rest of the paper,

---

due to the mechanical correlation between the instrument and the demeaned error term, we conduct a Monte Carlo simulation. We assume the error term is *i.i.d.* We use the predicted values from the OLS FE model and add an *i.i.d.* error to simulate the outcome variable (g_{is}^{sim}). We reconstruct our instrument as $Crossing_{is-2}^{sim} = \mathbf{1}\{y_{is-3} + g_{is}^{sim} \geq \bar{y}_{s-2}\}$, where $\bar{y}_s$ is the IDA threshold in the second year of period s. In the FE model, the instrument is thus mechanically correlated with the demeaned error term. We then estimate the 2SLS FE model using $Crossing_{is-2}^{sim}$ as the instrumental variable for aid and gauge the magnitude of the bias. We repeat this procedure 1,000 times, take the mean of 2SLS FE estimate and compare it with the OLS FE estimate (the true parameter). We estimate a negligible bias of less than 2% of the true parameter value.

as in the previous section, we rely on two alternative methods to conduct statistical inference. We report in parentheses robust standard errors clustered at the country level, which allow for within-country correlation. We also report in brackets the standard errors from the wild clustered bootstrap procedure. Both methods render similar statistical conclusions for our main parameters.

The second alternative to the 2SLS-FE model relies on a smoothing method of the latent process that determines our instrumental variable. Intuitively, we form a "synthetic control" for each crossing country using countries that are not in our sample. The synthetic control is constructed in such a way that the distance between income trajectories of the crossing country and its synthetic control is minimized. We then use the income trajectories of the synthetic control to predict the year of crossing, which is a function of shocks to other countries. Under the assumption that shocks across countries are not correlated (after controlling for all the covariates in equation (2)), the predicted crossing is not correlated with ε_{is-2}.

Specifically, we proceed to construct our synthetic control as follows. We include a panel of 130 other developing countries that were official DAC aid recipient countries between 1987 and 2010, together with the 35 countries in our original sample.[17] We demean all the series in our extended panel (of 165 countries) by projecting the annual log of nominal per capita GNI onto a set of country fixed effects, denoted $\hat{y}_{\mathrm{i}}$. We then take the residuals, $\hat{e}_{\mathrm{it}}$. For each of the 35 countries in our working sample, we construct a set of weights $\mathrm{w}_j \in \{\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_\mathrm{J}\}$ bounded between 0 and 1 for the 130 recipient

[17] Since both our sample and the extended dataset are unbalanced panels, for each of the 35 countries in our sample we use only a balanced panel of available donors.

countries such that the following distance function is minimized:

$$D_i = ||\boldsymbol{e}_i - \sum_j w_j \cdot \boldsymbol{e}_j||, \qquad (3)$$

where $||\cdot||$ is the Euclidean distance operator. The vector $\boldsymbol{e}$ includes the residuals $\hat{e}_{\mathrm{ht}}$. For each country i, we use all the sample years to minimize the influence of a given observation around the period of crossing the IDA threshold. Denote the optimal weight assigned to country j as w_{ij}^{*}. We then define $\hat{e}_{it} = \sum_{\mathrm{j}} \mathrm{w}_{ij}^{*} \cdot \hat{e}_{jt}$, and construct the predicted log per capita nominal GNI as $\hat{y}_{it} = \hat{y}_{\mathrm{i}} + \hat{e}_{it}$.[18] Since $\hat{y}_{it}$ is, by the nature of construction, plausibly uncorrelated with the error in equations (2), the predicted crossing of the IDA threshold based on $\exp(\hat{y}_{it})$ is also likely to be uncorrelated with the error term. Accordingly, the predicted crossing two periods earlier, $Crossing_{is-2}^{pred}$, satisfies the exclusion restriction needed for identification, and is a valid instrumental variable for Aid_{is-1}. On the other hand, due to the limited number of comparison countries, predicted crossing may contain measurement error from an imperfect synthetic matching, which in turn results in a weaker first stage. For 10 out of 35 countries, the predicted period of crossing differs from the actual period.[19]

[18] This algorithm is in the spirit of the synthetic control approach in Abadie et al. (2010).

[19] Using the predicted crossing as the instrumental variable introduces sampling error not accounted for in the synthetic control step, which leads to bias in the estimation of the standard error. Usually this is adjusted by bootstrapping the whole procedure. We did not do this because bootstrap fails for the K-nearest neighbor matching (Abadie and Imbens, 2008). The intuition is that in the K-nearest neighbor matching the number of matches does not increase smoothly with the sample size. Synthetic matching does allow the number of matches to rise as the sample size increases, but the increase is not smooth everywhere because we restrict the weights to be bounded between 0 and 1. More sophisticated bootstrap procedures

5.3 Baseline results

Table 3 reports the baseline results from estimating the models discussed in the previous subsection. Column 1 reports the estimate of equation (2) without instrumenting aid, while column 2 reports the estimate of equation (4) also without instrumenting aid. Columns 3 through 7 are all estimated by the 2SLS estimator where we instrument only for aid. Column 3 estimates equation (2) using $Crossing_{is-2}$ as the instrument for Aid_{is-1}. Column 4 uses the crossings based on predicted per capita GNI from the smoothing exercise, $Crossing_{is-2}^{pred}$, as the instrumental variable. Finally, columns 5 to 7 report estimates of equation (4) using respectively $\Delta Crossing_{is-2}$, $Crossing_{is-2}$ and $Crossing_{is-2}^{pred}$ as the instrumental variable.

The fixed effect model in Column 1 and first difference estimate in Column 2 show that aid is positively correlated with real economic growth. The estimated coefficient is statistically significant at the 5% level, but is small in magnitude. A one percent increase in the aid to GNI ratio increases annual real per capita GDP growth by 0.011 and 0.013 percentage points.

Studies on aid and growth typically calculate the increase in annual real per capita GDP growth rate in percentage *points* implied by a one percentage *point* increase in ODA's share of GNI. Because we take logs of ODA/GNI, the implied effect of a percentage point increase in ODA/GNI depends on its level. We use the average aid to GNI ratio at the period of crossing (i.e., 0.09)

---

may recover the correct inference, but we are not aware of a widely accepted method (also see Imbens and Wooldridge (2009) for a discussion of this issue).

because this is the most relevant value given our identification strategy. The OLS estimate in Column 1 suggests that a one percentage point increase in the aid to GNI ratio from the average level at the period of crossing is associated with a 0.12 percentage point increase in real per capita GDP growth. The result in Column 2 implies that a 1 percentage point increase in the aid to GNI ratio from the same level is associated with a 0.14 percentage point increase in growth, consistent with the magnitudes obtained by Clemens et al. (2012). They address endogeneity of aid simply by lagging it one period in a fixed-effects regression, as in our Column 1, and find that a one percentage point increase in aid/GDP from the sample mean increases annual real per capita GDP growth by 0.1 to 0.2 percentage points in the next (4-year) period. Werker et al. (2009) estimate a slightly larger effect of 0.22 percentage points instrumenting aid with the interaction between oil and a dummy indicating a Muslim country.

Columns 3 through 7 are estimated using the 2SLS method. The point estimates of the aid coefficient are more than twice as large as those estimated by OLS. In column 3 we use $Crossing_{is-2}$ as the instrument for Aid_{is-1}. Now a one percent increase in the aid to GNI ratio raises growth by 0.028 percentage points. The first stage is strong, with an F-statistic of about 16. Growth is negatively correlated with lagged income, supporting conditional convergence.[20] In column 4 we use the predicted crossings based on the smoothed per capita GNI trajectory, $Crossing_{is-2}^{pred}$, as the instrument. The

[20] Column 3 of Appendix Table C presents the estimates of the same specification as in Column 3 of Table 3, but using $Crossing_{is-1}$, $Crossing_{is-2}$, and $Crossing_{is-3}$ as instrumental variables. The point estimate of the effect of aid on growth remains very similar, but the first stage is weaker than in our baseline specification.

estimated coefficient associated with Aid_{is-1} increases slightly to 0.035, and is statistically significant at the 5% level. The first stage is weaker with an F-statistic of 7.4. We also report the Anderson-Rubin (AR) 95% confidence intervals for the coefficient associated with aid. These confidence intervals are robust to potential weak instruments (Finlay and Magnusson, 2009). Almost all of these confidence intervals are greater than 0. Note the similarity of the point estimates of the effect of aid on growth in columns 3 and 4.

Columns 5 through 7 estimate the first differenced model in equation 4. The coefficients are all larger than those in Column 3 and Column 4. Column 5 uses $\Delta Crossing_{is-2}$ as the instrumental variable for ΔAid_{is-1}. In this specification, we estimate the coefficient of aid using only the variability from the one period after crossing. The aid coefficient is 0.047. Column 6 uses $Crossing_{is-2}$ as the instrumental variable for ΔAid_{is-1}. The first stage and the estimated coefficients are essentially unchanged from those in column 5. Column 7 uses the predicted crossing, $Crossing_{is-2}^{pred}$, as the instrument. The first stage is strong, with an F-statistic of 24. The aid coefficient is 0.055, statistically significant at the 1% level. Overall, our instrumental variable estimates are robust and consistently larger than the OLS estimates.[21]

As discussed earlier, for the first differenced model, our instrumental variable will be invalid if the unobservable idiosyncratic error term in the growth equation (equation 2) is serially correlated. We test for the presence of serial auto-correlation in the error terms in equation 2 following Arellano and Bond (1991). The Arrellano-Bond test for serial correlation tests the n^{th}

[21] If we re-estimate the model in column 3 of Table 3 while including quadratic and cubic terms of y_{is-1}, the results remain similar (as shown in Online Appendix Table H).

order of serial correlation of the first differenced error to infer the $(n-1)^{th}$ order of serial correlation of the error terms in the original equation (see also Roodman, 2009a). We report the $p-$values of the Arellano-Bond test for AR(2) after estimations in Columns 5, 6, and 7. None of the tests rejects the null hypothesis of no serial correlation in the errors in equation 2.

So far we have treated Aid_{is-1} as the only endogenous variable. However, the initial income level, y_{is-1}, could also be endogenous. In particular, when we first difference the model, Δy_{is-1} is mechanically correlated with the first-differenced error term. Table 4 reports estimates where we also instrument y_{is-1} following the standard procedures in the literature. We therefore test for whether the model is under-identified. The p-values of the Kleibergen-Paap rank Lagrange Multiplier (LM) test uniformly reject the null hypothesis that the model is under-identified. Columns 1 and 2 re-estimate the models in columns 3 and 4 of Table 3, respectively, but also instrumenting y_{is-1} (i.e., y_{it-4}) with y_{it-5}. Column 3 re-estimates the model in column 6 of Table 3 and uses y_{it-8} to instrument for Δy_{is-1} (i.e., $y_{it-4} - y_{it-7}$). The aid coefficient remains similar. Columns 4 and 5 are estimated by the difference GMM estimator, which is widely used in this literature. Given the potential problems with many instruments in finite samples, we use a parsimonious set of instruments (Roodman, 2007, 2009a, 2009b; Bazzi and Clemens, 2013; Bun and Windmeijer, 2010), namely y_{it-8}, y_{it-9}, and y_{it-10}. We use $Crossing_{i,s-2}$ as an instrument in Column 4 and $Crossing_{i,s-2}^{pred}$ in Column 5. With more instruments than endogenous variables, we can test the validity of the over-identifying restrictions, and both the Sargan and Hansen tests do not reject the null hypothesis of their validity. The GMM estimation does not automatically provide F-statistics. In order to compare with other columns in terms of instrument strength, we report the first-stage F-statistics based on estimating the corresponding 2SLS models with the same specification as in the GMM model.

The failures to reject the null hypothesis of the over-identification restriction and to detect an AR(1) structure in the error term in equation 2 suggest that the error terms in the growth equation are serially uncorrelated. Besides these two pieces of evidence, we also provide a third piece of evidence. If the error terms were serially correlated, including lagged values of the dependent variable will alter the estimates of the aid coefficient. Including lagged values of the dependent variable on the right-hand-side also controls for any positive shocks to growth before crossing which are eventually reversed. We thus re-estimate the model in Column 3 of Table 3 but include the once-lagged value of the dependent variable as a control variable (Column 6 of Table 4), or its twice-lagged value (Column 7), and then both the once- and twice-lagged values (Column 8). Columns 9, 10, and 11 of Table 4 repeat Columns 6, 7, and 8 but using $Crossing_{i,s-2}^{pred}$ to instrument aid. Estimates of the effect of aid remain sizable and similar to those in Table 3.[22]

The estimated effects of aid on growth in Columns 1 to 5 in Table 4 are all very similar to each other. Taking the point estimate in Column 2 we observe that a one percent increase in the aid to GNI ratio increases real per capita GDP growth by 0.031 percentage points. A one percentage point increase in the aid to GNI ratio from its average value at the period of crossing

[22] In column 3 of Table 4, we find that the clustered standard errors and the robust standard errors are very similar (results not shown). The similarity between the two sets of standard errors is consistent with the evidence of lack of serial correlation of the error term in equation (2).

(0.09) thus raises the growth rate by 0.35 percentage points (i.e.,.031*[.01/.09]*100).[23]

5.4 Measurement error in aid

Aid/GNI is likely to be measured with error, and this possibility is consistent with the observation that the 2SLS estimates are larger than the OLS estimates. There are various reasons why aid is measured with error. Not all donors report their aid to the DAC in all years. For example, aid from the former Soviet Union and from China in the Mao era to other communist countries was not reported to the DAC. Aid from China and other emerging donors has significantly increased in recent years but is mostly not reported. The official definition of aid counts $1 in grants the same as $1 in concessionary loans (for any loan with a grant element of at least 25%). Additionally, but better understood, the denominator, GNI, is also measured with sizeable error for many less developed countries (Jerven, 2013). With classic measurement error, the OLS estimate of the effect of aid is biased towards zero. Demeaning or first differencing the model likely exacerbates the bias.

The natural experiment we exploit in this paper provides a unique

[23] We report the robust first-stage F statistic for the overall strength of the first stage (Kleibergen-Paap rank Wald F statistic) for specifications in Table 4. With multiple endogenous variables, the first stage F statistic is less informative than the case of one instrument. We thus construct 95% Anderson-Rubin confidence intervals for the coefficients associated with endogenous variables that are robust to weak instruments (Finlay and Magnusson, 2009). With two endogenous variables, the confidence interval of a particular coefficient depends on the value of the coefficient associated with the other endogenous variable. Thus the confidence interval for both endogenous variables will be a two-dimensional figure. We report these graphs for specifications in Table 4 in Appendix Figure A. Overall, the 95% confidence intervals for the aid coefficient are greater than 0 (except for Column 4 in which we obtain significance only at the 10% level, and we show the plot for the 90% confidence interval).

opportunity to gauge the magnitude of the attenuation bias in the OLS estimation due to measurement error. We have shown that the amount of aid a country receives declines substantially following its crossing of the IDA threshold. Assuming that the measurement error is *i.i.d.*, it would contribute much less to the total variation in ΔAid_{is} in periods closer to threshold crossings. Thus the OLS estimates of equation 4 using only periods in the neighborhood of the crossings are likely to provide more accurate estimates of aid's effect on growth than the one exploiting all of the variability in aid from all periods. To test this, we re-estimate Equation 4 by OLS and 2SLS (using $Crossing_{i,s-2}$ to instrument aid), successively narrowing the window of periods used in the analysis around the crossing point of each country.[24] We expect the OLS estimate of the aid effect on growth to increase as we narrow the window of estimation. Naturally, the coefficient in the 2SLS estimate of the first-differenced model should remain stable as the window narrows, because it uses only a single period for identification. We find exactly that pattern in Table 5, Panel B.

Panel A of Table 5 reports the OLS estimates. As we narrow the window, the estimated coefficient associated with foreign aid monotonically and gradually increases. The coefficient in Column 6 (with maximal two periods around the crossing) is 0.0201, more than 50% larger than that in Column 1 (with maximal 7 periods around the crossing). A generalized Hausman test of the null hypothesis that the coefficient associated with Aid_{is-1} is the same in Column 1 and in Column 6 is rejected with a p-value of 0.07. These findings

[24] We rely on first differenced models in this exercise because changing the number of periods also affects the estimation of the country fixed effects, and we want to hold everything constant except for the signal to noise ratio in aid.

are consistent with the existence of significant measurement errors in aid.[25]

5.5 Bunching

Our identification strategy hinges on the large decline in the amount of aid received following the crossing of a pre-determined threshold. If countries manipulate their income data to remain below the threshold, then threshold crossing may not be a valid instrument for aid.

Endogenous manipulation of the income level is unlikely to be prevalent for three reasons. First, the GNI estimates used by the World Bank are by no means entirely within a government's control. The national accounts data produced by national statistical agencies are merely one of several inputs into the World Bank's income estimates (Jerven, 2013). Governments cannot perfectly predict (1) the adjustments to those national accounts data often made by World Bank staff, (2) the exchange rates used, or (3) the population estimates used in constructing GNI per capita. Second, crossing also depends on the current IDA threshold, and its annual adjustments for global inflation rates cannot be predicted perfectly either. Finally, income level with respect to the threshold is not the only criterion for IDA eligibility; e.g. countries that cross the threshold from below can remain eligible if they are judged to be non-creditworthy for non-concessionary lending.

Even if governments manipulated GNI to delay their graduation from IDA, the resulting bias would work against our main finding. Note that GNI per capita and its growth would be understated prior to crossing the threshold,

[25] Needless to say, as discussed extensively through the paper, this is not the only source of potential endogeneity in aid. Furthermore, we note that the discussion above is based on the assumption of a homogeneous treatment effect. Relaxing this assumption, the results found in Table 5 would also be consistent with the presence of heterogeneous treatment effects where aid has the largest effect around the IDA threshold (instead of when countries were poorer).

when aid is relatively high. After crossing, there would be little reason to continue understating GNI per capita, and correcting it would overstate growth for at least one period after crossing, when aid is relatively low.

Nevertheless, we tested for evidence of data manipulation. Figure 3A is a histogram that shows the distance between a country's current GNI per capita and the contemporaneous IDA threshold. All countries that were ever eligible for IDA between 1987 and 2010 are included, and each GNI per capita value in each country-year is treated as a separate observation. We group country-year observations in 100-dollar bins according to the distance between the income level and the contemporaneous IDA threshold. If many governments understate GNI to stay below the IDA threshold, we should observe significant "bunching" of observations just below the threshold, relative to the number of observations just above it. Specifically, we should observe the bin just to the left of the threshold to be abnormally high relative to the neighboring bins. If there is no bunching, the numbers of observations in each bin should cross the threshold of zero smoothly. As shown in Figure 3A, there is no visual evidence that countries bunch right below the IDA threshold. A formal test confirms this result. Using a density test proposed by McCrary (2008), we find no significant evidence of bunching. Figure 3B is a density graph that shows the fitted kernel density functions at both sides of the threshold. The density crosses the IDA threshold smoothly, and the minor difference is by no means statistically significant.

5.6 Further robustness checks

We next present various additional robustness tests, for which we report two sets of results. These tests all use either $Crossing_{i,s-2}$ or $Crossing_{i,s-2}^{pred}$ as the instrument for aid. They are all based on our preferred specification in

Column 3 of Table 3 (or Column 4 of Table 3 when $Crossing_{i,s-2}^{pred}$ is used as the instrument). As we shall see, these robustness checks yield results that are largely consistent with our main findings.[26]

Throughout the study, we control for period fixed effects, log of initial income, and log of population. Period fixed effects take account of any time-specific shocks that affect all countries. Initial income and population are among the key time-varying factors that affect economic growth. We also control for country fixed effects which account for any time-invariant cross-country variability in economic growth, and most slow-moving factors (over our relatively short 1987-2010 period) such as the quality of governance. Here we further test whether other time-varying factors could be confounding the effect of aid on economic growth, by adding to the baseline regression a host of economic and political variables, including the primary school enrollment rate, the Freedom House index of civil liberty and political rights, the World Bank's Country Policy and Institutional Assessment (CPIA) ratings, total trade as a percentage of GDP, broad money as a percentage of GDP, inflation as measured by changes in the GDP deflator, dummies for whether the country is experiencing a banking crisis, currency crisis, or debt crisis (central government debt/GNI is available for fewer than half of the countries in our sample), and a survey-based measure of country creditworthiness from Institutional Investor. Due to missing values, we add these variables in separate groups to maintain a reasonable sample size for each regression. Table 6 shows the results of these exercises. Few of the additional regressors have a statistically significant effect on growth in either set of estimations for

[26] Results are also robust to other specifications in Table 3.

our relatively small sample of countries. The aid coefficients remain similar in magnitude to their counterparts in the baseline regressions; the coefficient of our key aid variable range from 0.028 to 0.046, all statistically significant at conventional levels.

Recall that our instrument is a dummy variable that switches from 0 to 1 when a country has crossed the income threshold from below two periods earlier. The countries in our sample all crossed the threshold at some point between 1987 and 2010, but for countries that crossed the threshold in the last and the next-to-the-last periods (i.e. in periods 7 or 8) the instrumental dummy variable is always zero. We keep these countries in the sample because they satisfy our simple rule for sample selection and they provide relevant information for estimating the effects of the control variables. We now check whether our results still hold when we drop these countries. In Column 1 of Panel A in Table 7, we drop the seven countries that crossed the threshold in the final period of our sample (2008-2010). The point estimate increases (to 0.041) and remains statistically significant at conventional levels. In Column 2 we further drop the seven countries that crossed the threshold in the next-to-last period (2005-2007). The coefficient is slightly larger (i.e., at 0.045) than that in Column 1 and remains statistically significant at the 5% level. The first stage remains strong in each case. The first stage is weaker when we use $Crossing_{i,s-2}^{pred}$ as the instrument in panel B, but the results are qualitatively similar.

As shown in Table 1, a few countries crossed the IDA threshold from below more than once during the sample period. These countries must have crossed the threshold from above after its first crossing from below, then crossed from below again. These cases might be a threat to our identification strategy. If income drops below the threshold again immediately after the first crossing for any reason other than a decline in aid, the estimated effect of aid might be confounded. In Column 3 of Panel A of Table 7, we drop countries with multiple crossings. The estimated effect of aid (0.026) remains similar. In Column 4, we use the last threshold crossing (from below) instead of the first one to construct the instrumental variable. The estimated effect of aid on growth (0.024) again remains similar relative to the baseline specification. It remains statistically significant in panel A, but in panel B the standard errors increase and we cannot reject the null hypothesis of no effect at conventional levels.

There are two exceptional groups of countries in the sample. First, several countries in the sample benefit from the "small island country exception," which permits island nations with populations below 1.5 million to remain IDA eligible, even after surpassing the income threshold. Second, a few countries were never classified as IDA eligible in the 1987-2010 period due to various reasons, despite having income levels below the threshold for one or more years. We include them in the sample because the IDA income threshold potentially serves as a useful benchmark for donors other than IDA. In Column 5 of Table 7 we drop the three small island countries. In Column 6, we drop the four countries that were never eligible for IDA throughout the sample period. The estimated effects of aid are robust to these sample changes, with the coefficient ranging from 0.027 to 0.03.

6. Potential Mechanism

Given the positive effect found in this paper, an important question is the mechanism by which aid boosts growth. In the short run, an important channel through which aid could cause growth is through fostering physical investment.

Irrespective of the form aid takes, it constitutes a flow of funds to recipient countries, which, if they are financially and perhaps fiscally constrained (see Kraay, 2012), would release resources in the economy that could be invested. Because our sample countries were all financially constrained by definition, this is a relevant scenario. Admittedly, our findings here do not rule out other mechanisms nor can it necessarily be extrapolated to countries that are more developed and have better access to credit markets.

We consider the expected effect of an increase of one percentage point in aid if it were fully invested in physical capital. Approximating the technology at the aggregate level of the economy linearly, the rate of economic growth would increase in the inverse of the capital-output ratio.[27] Using the standard perpetual inventory method, we estimate this ratio to be approximately equal to 2 in our sample.[28] Thus, one would expect that growth could increase by as much as 0.5 percentage points if all of the aid were invested.

How much aid fosters physical investment is therefore an empirical question. To this end, in Appendix Table I we re-estimate the baseline specifications in Table 3, replacing economic growth with the period average investment to GDP ratio as the dependent variable. The OLS estimates in Column 1 and Column 2 of the effect of aid on investment are statistically insignificant and have the "wrong" sign. Instrumenting aid in Columns 3 to 7,

---

[27] Even under a linear technology such as the AK model, the effect of aid on growth could be strictly concave since, for instance, the effect of aid on physical investment might be decreasing in aid. Moreover, the aggregate technology could be non-linear as suggested by the robust finding of conditional convergence in the empirical growth literature.

[28] Note that this figure is also consistent with standard growth accounting assumptions. Assuming capital per capita depreciation rate of 10 percent per year, and an investment rate of 25 percent of GDP per year, a country with a capital-output ratio of 2 would grow, in per capita terms, at 2.5 percent per year, which is consistent with the figures in our sample.

we find that the coefficient associated with aid becomes positive and ranges between 0.5 and 0.8. In one case it is significant at the 10% level, but in general the p-values are around 0.15. In column 4 we see that a one percent increase in the aid to GNI ratio increases the investment to GDP ratio by 0.049 percentage points. Evaluated at the average level of aid to GNI ratio at the period of crossing, an increase of 1 percentage point in the aid to GNI ratio increases the investment to GDP ratio by 0.54 percentage points. Using the estimated capital-output ratio of 2, a 0.54 percentage point increase in physical investment would raise growth by 0.27 percentage points, not far away from our back-of-the-envelope calculation of the effect of aid on growth.[29]

7. Suggestive Evidence on External Validity

The results found in this paper are Local Average Treatment Effects (LATE) from a sample of aid-recipient countries that all successfully crossed the IDA income threshold from below between 1987 and 2010. How would our results apply to other aid-recipient countries, particularly those that are still below the IDA cutoff? On the one hand, if the sample countries crossed the IDA threshold level because they have fundamentally different attributes, our results may have little relevance for those countries remaining very poor. On the other hand, if the difference between crossing and non-crossing countries is mainly due to being in different stages of development, and the two samples have similar growth patterns conditional on the initial income, then our results may apply to these countries as well when their income level approaches the

[29] Admittedly, the back-of-the-envelope calculation changes somewhat when alternative growth models are used. Using the A-K model with per capital output function $y = Ak^{\alpha}$, when $\alpha = 0.35$ and $K/Y = 2$, investment accounts for a smaller, but still substantial 29% of the total effect on growth.

IDA threshold.

We thus investigate whether the crossing countries have systematically higher growth rates (conditional on the initial income level) than the non-crossing countries using a simple regression. We include all country-year observations with per capita GNI level below the threshold, and linearly project the annual real per capita GDP growth onto a dummy variable indicating whether the observation belongs to a country in the crossing sample, controlling for a one-year lag of log per capita real GDP and its square, as well as for year dummies. When their income levels were still below the IDA threshold, real per capita GDP in countries that eventually crossed were growing on average 1.98% per year. During the same period, real per capita GDP in countries that had not crossed the threshold by 2010 were growing on average 1.11% per year. In Appendix Table J we show that, after we control for the initial real GDP per capita level (Column 1) as well as its quadratic form (Column 2), the difference in average annual GDP per capita growth between the crossing countries and the non-crossing countries is approximately 0.7 percentage points and not statistically significant. In Column 3, we compare the average annual GDP per capita growth of the two groups of countries within each quartile of the distribution of the lagged income levels. When the income level is in the lowest quartile, countries that eventually crossed the threshold were growing at a much higher rate than those that have not crossed. In other higher quartiles of income level, the differences in annual growth rates between the two groups become smaller in magnitude and not statistically significant. Thus, for most of the income distribution below the IDA threshold, we do not find significant differences in pre-crossing growth rates between the sample of countries studied in this paper and those still below the IDA threshold. This finding suggests that that our estimates of aid's growth effects, based on countries that recently crossed the income threshold, may be relevant to countries that have not yet crossed, as the latter group is not systematically different from the former group.

8. Conclusion

We present new evidence on the effect of foreign aid on recipient countries' economic growth, estimating a model similar to those in the recent aid-growth literature, but addressing the identification challenge using a novel instrument for aid. This instrument is derived from an exogenously-determined aid allocation policy, exploiting the substantial drop in aid after a country crosses an exogenous income threshold set by the World Bank for IDA eligibility. Using a sample of countries that have crossed the IDA threshold since 1987, we find that rapid reductions in aid subsequent to crossing the IDA threshold have a sizable negative effect on growth. Reducing the aid to GNI ratio by one percentage point from its average value at the period of crossing decreases real per capita GDP growth by approximately 0.35 percentage points. Our finding of a positive impact of aid on growth is consistent with a majority of recent studies (Arndt et al., forthcoming), although our estimated effect is somewhat larger than those reported in Clemens et al. (2012) and most other studies.

We address various identification concerns, and our results remain robust throughout. We provide a new way of constructing predicted income trajectories and crossing periods using a smoothing technique.

The estimates are based on a relatively small group of countries, but this group is particularly interesting because it is comprised of poor and financially constrained countries that receive large amounts of aid (averaging 8% of GNI).

We provide suggestive evidence that our results may generalize to countries that are still under the IDA threshold as they grow closer to the threshold.

We present evidence that aid also likely increases the investment rate, although this effect is less precisely estimated. A back-of-the-envelope calculation is consistent with physical investment being a main channel through which aid operates in the short-run.

Identification of causal effects is a daunting task—especially at the macroeconomic level—so all causal estimates of country level parameters should be interpreted cautiously. Still, at the micro level, researchers need to evaluate on a case by case basis which aid projects work better, if at all. Our evidence shows only that overall foreign aid increases economic growth among poor countries where aid is a large source of funding. Moreover, even at the macro level, aid may have heterogeneous effects depending on recipient characteristics, aid modalities, and donor motives (Mekasha and Tarp, 2013). For example, aid provided by some bilateral institutions for political or commercial reasons may be less effective (Dreher et al., 2014), and may be less sensitive to crossing the IDA threshold. Our relatively large effect may apply to less politicized aid. Following the end of the Cold War, however, the share of aid that is highly politicized has arguably fallen significantly, with geopolitical motives declining in importance relative to developmental concerns (Headey, 2008).

Our relatively small and homogeneous sample is not ideal for testing heterogeneous effects of aid. Moreover, because we identify only the effect of aid on growth in the short term, our evidence does not contradict any view of aid's effects on long-term development. Despite these caveats, we believe our evidence contributes to understanding the effect of aid on economic growth in the short-term for poor countries that are financially constrained.

Our results also contribute to the empirical literature on donors' aid allocation decisions across recipient countries (e.g. Alesina and Dollar, 2000; Chong and Gradstein, 2008). Specifically, they support the conjecture by Moss and Majerowicz (2012) that bilateral donors use IDA policies – and specifically its income eligibility threshold – as an informative signal of recipient need. Patterns of donor "herding" measured by Frot and Santosi (2011) may be partially due to donors' common responses to recipient countries' crossing the IDA income threshold.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010). “Synthetic Control Methods for Comparative Case Study: Estimating the Effect of California’s Tobacco Control Program.” *Journal of the American Statistical Association*, 105(490), 493-505.

Abadie, Alberto, and Guido W. Imbens. "On the Failure of the Bootstrap for Matching Estimators." *Econometrica* 76.6 (2008): 1537-1557.

Acemoglu, Daron, Simon Johnson, and James A. Robinson (2001). “The Colonial Origins of Comparative Development: An Empirical Investigation.” *American Economic Review* 91(5), 1369-1401.

Alesina, Alberto and David Dollar (2000). “Who Gives Foreign Aid to Whom and Why?” *Journal of Economic Growth* 5, 33-64.

Angrist, Joshua, Guido W. Imbens, Donald B. Rubin. 1996. “Identification of Causal Effects Using Instrumental Variables,” *Journal of the American Statistical Association* 91 (434), 444-455.

Arellano, Manuel, and Stephen Bond (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58(2), 277-297.

Arndt, Channing, Sam Jones, and Finn Tarp (forthcoming). “What is the Aggregate Economic Rate of Return to Foreign Aid?” *World Bank Economic Review.*

Arndt, Channing, Sam Jones and Finn Tarp (2010). “Aid, Growth and Development: Have We Come Full Circle?” *Journal of Globalization and Development* 1(2), 5.

Bazzi, Samuel and Michael Clemens (2013). “Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth.” *American Economic Journal: Macroeconomics*, 5(2), 152-186.

Boone, Peter (1996). “Politics and the Effectiveness of Foreign Aid.” *European Economic Review* 40(2), 289-329.

Bruckner, M. (2013). “On the Simultaneity Problem in the Aid and Growth Debate.” *Journal of Applied Econometrics* 28(1), 126-50.

Burnside, Craig and David Dollar (2000). “Aid, Policies, and Growth.” *American Economic Review* 90(4), 847-68.

Bun, Maurice J.G. and Frank Windmeijer (2010). “The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models.”

Econometrics Journal 13(1), 95-126

Cameron, A. Colin, Jonah Gelbach and Douglas Miller (2008). “Bootstrap-based Improvement for Inference with Clustered Errors.” *Review of Economics and Statistics* 90(3), 414-427.

Chong, Alberto and Mark Gradstein (2008). “What Determines Foreign Aid: The Donors’ Perspective.” *Journal of Development Economics* 87, 1-13.

Clemens, Michael A., Steven Radelet, Rikhil R. Bhavnani and Samuel Bazzi (2012). “Counting Chickens When They Hatch: The Short-Term Effect of Aid on Growth.” *Economic Journal* 122(561), 590-617.

Delgaard, C.J., H. Hansen, and F. Tarp (2004). “On the Empirics of Foreign Aid and Growth.” *Economic Journal* 114(496), F191-F216.

Dreher, Axel, Vera Eichenauer and Kai Gehring (2014). “Geopolitics, Aid and Growth.” University of Heidelberg, Department of Economics, Discussion Paper Series, No. 575.

Dreher, Axel and Sarah Langlotz (2015). “Aid and Growth: New Evidence Using an Excludable Instrument.” *CESifo Working Paper* No. *5515.*

Easterly, William, Ross Levine and David Roodman (2004). “Aid, Policies, and Growth: Comment.” *American Economic Review* 94(3), 774-80.

Finlay, Keith and Leandro M. Magnusson (2009). “Implementing Weak-Instrument Robust Tests for a General Class of Instrumental-Variables Models.” *Stata Journal* 9(3), 398-421.

Frot, Emmanuel and Javier Santiso (2011). “Herding in Aid Allocation.” *Kyklos* 64(1), 54-74.

Hansen, Henrik and Finn Tarp (2001). “Aid and Growth Regressions.” *Journal of Development Economics* 64(2), 547-570.

Headey, Derek (2008). “Geopolitics and the Effect of Foreign Aid on Economic Growth: 1970-2001.” *Journal of International Development* 20, 161-180.

Imbens, Guido W., and Jeffrey M. Wooldridge. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47.1 (2009): 5-86.

Jerven, Morten (2013). “Comparability of GDP Estimates in Sub-Saharan Africa: The Effect of Revisions in Sources and Methods since Structural Adjustment.” *Review of Income and Wealth* 59, S16-S36.

Kapur, Devesh; John Prior Lewis and Richard C. Webb (1997). "IDA: The Bank as a Dispenser of Concessional Aid." Chapter 17 (pp. 1119-1159) in *The World Bank: Its First Half Century*, Vol. 1, edited by D. Kapur, J. P. Lewis and R.C. Webb.

Kraay, Aart (2012). "How Large is the Government Spending Multiplier? Evidence from World Bank Lending." *Quarterly Journal of Economics* 127, 829-887

Kronmal, Richard A. 1993. "Spurious Correlation and the Fallacy of the Ratio Standard Revisited." *Journal of the Royal Statistical Society.* 156(3), 379-392

McCrary, Justin (2008). "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142(2), 698-714.

Mekasha, Tseday Jemanah and Finn Tarp (2013). "Aid and Growth: What Meta-Analysis Reveals." *Journal of Development Studies* 49(4), 564-83.

Moss, Todd and Stephanie Majerowicz (2012). "No Longer Poor: Ghana's New Income Status and Implications of Graduation from IDA." Center for Global Development Working Paper 300, Washington, D.C.

Nunn, Nathan and Nancy Qian (2013). "Aiding Conflict: The Effect of U.S. Food Aid on Civil War". Manuscript.

OECD (2012). "Statistics on Resource Flows to Developing Countries."

http://www.oecd.org/dac/aidstatistics/statisticsonresourceflowstodevelopingcountries.htm.

Qian, Nancy (2015). "Making Progress on Foreign Aid." *Annual Review of Economics* 7(1).

Rajan, Raghuram and Arvind Subramanian (2008). "Aid and Growth: What Does the Cross-Country Evidence Really Show?" *Review of Economics and Statistics* 90(4), 643-665.

Rodrik, Dani (1996). "Understanding Economic Policy Reform." *Journal of Economic Literature* 34, 9-41.

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

Roodman, David (2009a). "How to Do xtabond2: An Introduction to Difference and System GMM in Stata." *Stata Journal* 9(1), 86-136.

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

Temple, Jonathan (2010). "Aid and Conditionality." *Handbook of development economics* 5: 4415-4523.

Temple, Jonathan, Nicolas Van de Sijpe (2014). “Foreign Aid and Domestic Absorption.” CESifo Working Paper No. 5029.

Werker, Eric, Faisal Z. Ahmed and Charles Cohen (2009). “How is Foreign Aid Spent? Evidence from a Natural Experiment.” *American Economic Journal: Microeconomics* 1(2), 225-244.

World Bank (1989). “IDA Eligibility.” IDA 9 Discussion Paper No. 3. March. Washington, DC: The World Bank.

World Bank (2010). “A Review of IDA’s Long Term Financial Capacity and Financial Instruments” (IDA16 Replenishment Report). Washington, DC: The World Bank.

Appendix A: Standard Errors

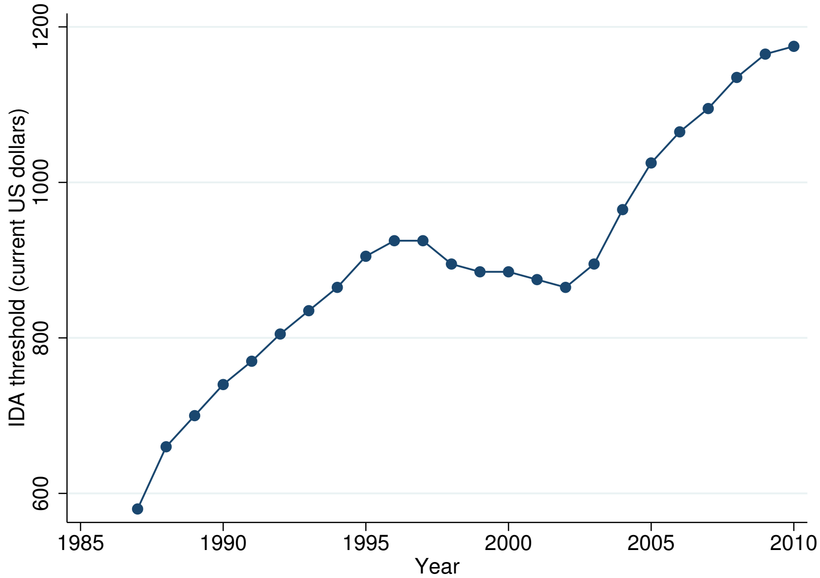
We first report robust standard errors clustered at the country level, which allow for arbitrary within-country correlation. There are 35 countries in our sample. Standard asymptotic tests might over-reject the null hypothesis under the presence of few clusters. Although 35 clusters is not a small number, for robustness we also report the standard errors from the clustered wild bootstrap procedure following Cameron et al. (2008). Cameron et al. (2008) recommends the clustered wild bootstrap-t procedure for better asymptotic properties, as the t statistic is “pivotal”. However, in order to compare with the clustered standard errors, we report the standard errors from clustered wild bootstrap procedure. Inference based on the bootstrap-t procedure is quantitatively similar to that based on bootstrapped standard errors. Either approach yields very similar statistical inferences. Throughout the paper, we report both sets of standard errors. For brevity, we focus our discussion on the clustered standard errors, and in tables, we mark asterisks after these standard errors to indicate conventional levels of statistical significance.

Appendix B: The Functional Form of Aid

Our measure of aid is slightly different from most of the literature, which often uses the aid to GDP or GNI ratio as the main explanatory variable. We take the log of aid since the previous body of evidence suggests that the marginal effect of aid on growth is decreasing. The logarithmic form is a parsimonious way to introduce concavity while preserving our ability to identify aid's causal impact with only one exogenous binary instrumental variable. However, note that instrumented aid still takes on a large number of values on its domain, since each country's aid is shifted by the instrument starting from different values (over time). The logarithmic specification is admittedly less flexible than a quadratic specification; in particular, it does not allow the marginal effect of aid to change its sign. Clemens et al. (2012), however, find that the effect of aid on growth does not turn negative until aid exceeds roughly 15% of GDP. In our sample, over 90% of the observations are below 11% of GDP. Thus, a logarithmic specification provides a good approximation over the range of observed values on aid. Additionally, we report as a robustness test in Table G results from using $Aid^{*}_{is-1} = (\sum_{k=3}^{5} \frac{ODA_{it-k}}{GNI_{it-k}})/3$ (i.e. not logged) as the measure for aid. Aid^{*}_{is-1} has a positive, quantitatively large, and marginally significant coefficient, despite a lower first stage F statistic. We find that a 1 percentage point increase in the aid to GNI ratio raises annual per capita GDP growth by 0.57 percentage point at the sample mean of the aid to GNI ratio. This estimate is even larger than our baseline result, reported later in Section 5.

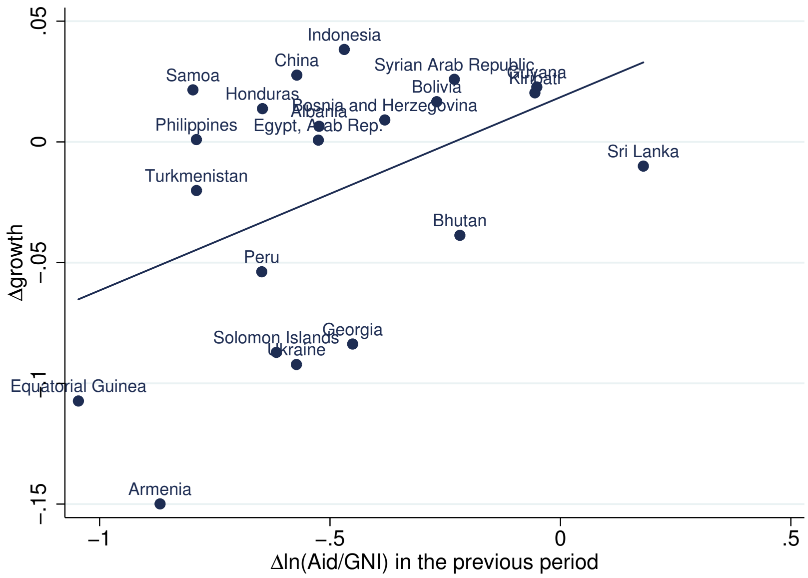
Figures and Tables

Figure 1: IDA Threshold 1987-2010



Note: IDA threshold is originally nominated in current international dollars. We convert it in current US dollars.

Figure 2: Changes in Growth and Once-Lagged Changes in Aid Two Periods after Crossing



Note: Each dot is a country. We show the relationship between changes in real per capita GDP growth two periods after crossing (Δy_{is}, where period s is two periods after crossing, y-axis), and changes in aid to GNI ratio in the previous period ($\Delta \ln(Aid/GNI)_{is-1}$, x-axis). The slope of the fitted line is 0.08, with a stadard error of 0.04.

Figure 3A: Histogram of Income

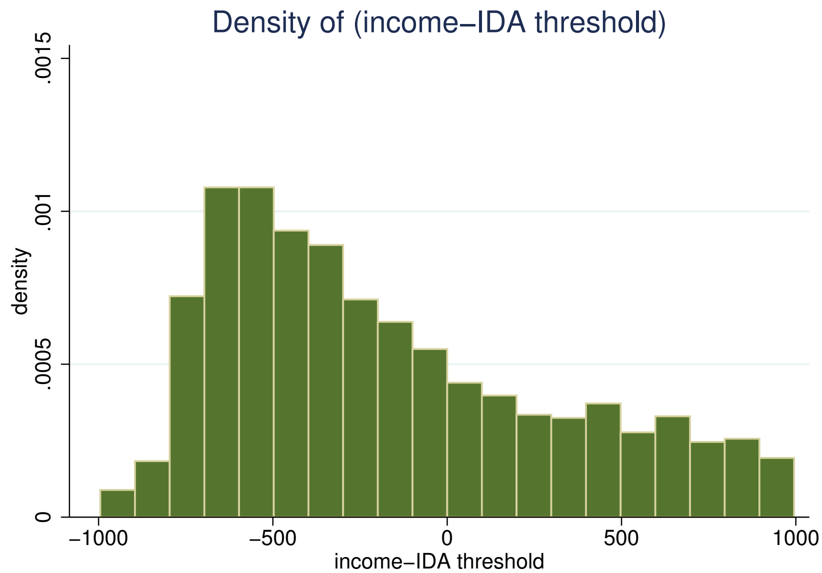
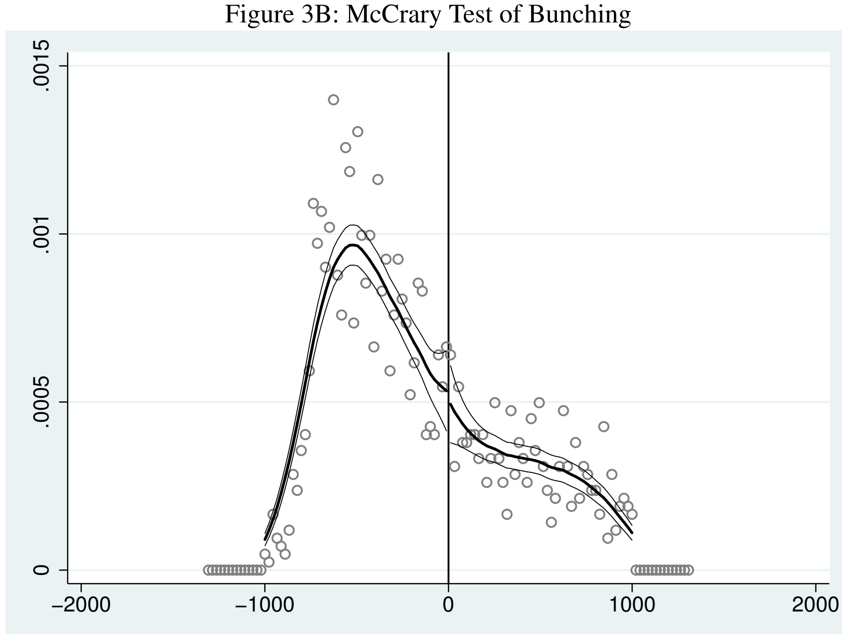


Figure 3B: McCrary Test of Bunching



Note: There are 1,920 country-year observations from 112 countries that were ever on the DAC list between 1987 and 2010. For each country-year observation, we calculate the distance of the current per capita GNI (y_{it}) from the current IDA threshold ($\bar{y}_t$). We restrict the distance ($y_{it} - \bar{y}_t$) between -1000 and 1000. Graph A is a histogram of country-year observations against ($y_{it} - \bar{y}_t$), grouped in 100-dollar bins. Graph B shows the McCrary density test. The discontinuity estimate (log difference in height from left to right) is -0.0476, with a standard error of 0.1776.

Table 1: Sample Countries and Years of Crossing the IDA Threshold

Country Name	Year of Crossing (graduation)	Country Name	Year of Crossing (graduation)
Albania	1999 (2008)	India	2010 (2014)[5]
Angola	2005 (2014)	Indonesia	1994
Armenia	2003 (2014)		2004 (2008)
Azerbaijan	2005 (2014)	Kiribati	1988
Bhutan	2004[2]		1992[3]
Bolivia	1997	Moldova	2007[2]
	2005[4]	Mongolia	2006[2]
Bosnia and Herzegovina	1997 (2014)	Nigeria	2008[2]
Cameroon	2008[2]	Papua New Guinea	2009[2]
China	2000 (1999)	Peru	1990[6]
Congo, Rep.	2006[2]	Philippines	1994 (1993)
Djibouti	2007[4]	Samoa	1995[3]
Egypt	1995 (1999)	Solomon Islands	1997
Equatorial Guinea	1998 (1999)	Sri Lanka	2003[2]
	2000	Sudan	2008[1]
Georgia	2003 (2014)	Syrian Arab Republic	1998[6]
Ghana	2009[4]	Timor-Leste	2006[2]
Guyana	1999	Turkmenistan	2002[6]
	2005[4]	Ukraine	2003[6]
Honduras	2000[4]	Uzbekistan	2010[2]

Note: Countries that crossed the IDA threshold from below between 1987 and 2010.

1. Inactive countries: no active IDA financing due to protracted non-accrual status.

2. Blend countries: IDA-eligible but also creditworthy for some IBRD borrowing.

3. Small island economy exception: small islands (with less than 1.5 million people, significant vulnerability due to size and geography, and very limited credit-worthiness and financing options) have been granted exceptions in maintaining their eligibility.

4. Borrowing on blend terms: countries that access IDA financing only on blend credit terms.

5. India graduated from IDA at the end of FY14 but will receive transitional support on an exceptional basis through the IDA17 period (FY15-17)

6. Never IDA-eligible.

Categorization of current borrowing countries from http://www.worldbank.org/ida/borrowing-countries.html (accessed in November 2015).

Table 2: IDA Threshold and Aid

	(1)	(2)	(3)	(4)	(5)
	IDA	DAC	NDAC	MLA	ODA
$Crossing_{is-2}$	-2.485	-0.961	-2.222	-0.750	-0.876
	(1.371)*	(0.238)***	(1.776)	(0.302)**	(0.216)***
	[1.214]	[0.209]	[1.540]	[0.263]	[0.188]
y_{is-1}	-8.587	-1.443	-4.739	-2.508	-1.535
	(1.691)***	(0.420)***	(1.964)**	(0.865)***	(0.324)***
	[1.515]	[0.369]	[1.744]	[0.794]	[0.286]
Country FE	X	X	X	X	X
Period FE	X	X	X	X	X
N	247	247	247	247	247
N countries	35	35	35	35	35

Note: Each observation is a country-period. Dependent variables are the log average share of aid in GNI by donor in the last period and share of total aid in GNI in the last period. There are 35 countries in the sample. Country fixed effects, period fixed effects, and log population in the last period are controlled in all columns. $Crossing_{is-2}$ is a dummy variable indicating whether the country crossed the IDA cutoff at least two periods earlier. y_{is-1} is the log real GDP per capita in the second year of the last period, y_{it-4}. Cluster-robust standard errors are in parentheses, $*p < 0.10$, $**\,p < 0.05$, $***p < 0.01$. Wild cluster bootstrapped standard errors are reported in brackets.

Table 3: Baseline Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Main Specification	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Aid_{is-1}	0.0105	0.0133	0.0281	0.0352	0.0475	0.0485	0.0552
	(0.00455)**	(0.00615)**	(0.0100)***	(0.0147)**	(0.0239)**	(0.0177)***	(0.0190)***
	[0.004]	[0.006]	[0.010]	[0.014]	[0.023]	[0.018]	[0.019]
y_{is-1}	-0.0675	-0.161	-0.0371	-0.0249	-0.0976	-0.0957	-0.0835
	(0.0246)***	(0.0231)***	(0.0256)	(0.0322)	(0.0516)*	(0.0387)**	(0.0415)**
	[0.022]	[0.022]	[0.026]	[0.033]	[0.051]	[0.039]	[0.043]
Period FE	X	X	X	X	X	X	X
Country FE	X		X	X			
First differenced		X			X	X	X
IV			X	X	X	X	X
IV from predicted income				X			X
IV first differenced					X		
N	247	212	247	247	212	212	212
Number of countries	35	35	35	35	35	35	35
First stage F statistic (K-P Wald)			16.50	7.385	19.52	16.16	24.06
95% A-R CI			[.0118, .0627]	[.0136,0.1247]	[-0.0083,0.1034]	[0.0212,0.1166]	[0.0258,0.1207]
AR(2) p-value					0.729	0.830	0.824
Olea - Pflueger robust weak IV test							
Effective F-Stat			16.91	7.57	20.00	16.56	24.65
Critical value for % worst case bias							
value for 5% worst case bias			37.42	37.42	37.42	37.42	37.42
value for 10% worst case bias			23.11	23.11	23.11	23.11	23.11
value for 20% worst case bias			15.06	15.06	15.06	15.06	15.06
value for 30% worst case bias			12.04	12.04	12.04	12.04	12.04

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild cluster bootstrapped standard errors are reported in brackets. See text for more details. Standard errors in curly brackets are from a bootstrapping procedure in which only crossing countries are resampled with replacement.

Table 4: Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Main Specification	2SLS	2SLS	2SLS	GMM	GMM	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Aid_{is-1}	0.0258	0.0312	0.0427	0.0298	0.0308	0.0198	0.0229	0.0205	0.0331	0.0359	0.0377
	(0.00966)***	(0.0138)**	(0.0182)**	(0.0122)**	(0.0133)**	(0.00984)**	(0.0101)**	(0.00909)**	(0.0149)**	(0.0152)**	(0.0155)**
	[0.010]	[0.013]	[0.018]			[0.010]	[0.010]	[0.010]	[0.014]	[0.015]	[0.015]
y_{is-1}	-0.0525	-0.0431	-0.137	-0.128	-0.136	-0.0540	-0.0543	-0.0518	-0.0326	-0.0334	-0.0245
	(0.0226)**	(0.0290)	(0.0679)**	(0.0554)**	(0.0529)***	(0.0184)***	(0.0217)**	(0.0201)**	(0.0282)	(0.0281)	(0.0261)
	[0.023]	[0.029]	[0.069]			[0.019]	[0.022]	[0.022]	[0.022]	[0.028]	[0.026]
Period FE	X	X	X	X	X	X	X	X	X	X	X
Country FE	X	X				X	X	X	X	X	X
First differenced			X	X	X						
IV for y_{is-1}	X	X	X	X	X	X	X	X	X	X	X
Predicted crossing		X			X				X	X	X
lagged dependent variables						1	2	1,2	1	2	1,2
N	247	247	212	212	212	245	229	229	245	229	229
Number of countries	35	35	35	35	35	35	35	35	35	35	35
First stage F statistic (K-P Wald)	8.098	3.601	11.46	4.453	6.164	5.951	6.520	6.448	3.263	4.083	4.103
Under-id (K-P rank LM) (p-value)	0.001	0.012	0.000	0.004	0.001	0.002	0.002	0.002	0.015	0.010	0.009
Number of IVs				12	12						
Hansen test for over-id (p-value)				0.330	0.174						
Sargan test for over-id (p-value)				0.275	0.106						
AR(2) p-value				0.950	0.849						

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. Instrumental variable for y_{is-1} is y_{it-5} all columns except for Columns 3, 4, 5. y_{is-1} is instrumented by y_{it-8} in Column 3, and is instrumented by y_{it-8}, y_{it-9}, and y_{it-10} in Columns 4 and 5. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild cluster bootstrapped standard errors are in brackets. See text for more details.

Table 5: Narrowing Periods

Panel A: OLS - first differenced	(1)	(2)	(3)	(4)	(5)	(6)
# of maximal periods around the crossings	7	6	5	4	3	2
Aid_{is-1}	0.0133	0.0133	0.0137	0.0142	0.0154	0.0201
	(0.00615)**	(0.00615)**	(0.00640)**	(0.00631)**	(0.00814)*	(0.00946)**
	[0.006]	[0.006]	[0.006]	[0.006]	[0.008]	[0.009]
y_{is-1}	-0.161	-0.161	-0.162	-0.163	-0.161	-0.157
	(0.0231)***	(0.0232)***	(0.0232)***	(0.0235)***	(0.0256)***	(0.0298)***
	[0.022]	[0.022]	[0.022]	[0.022]	[0.024]	[0.028]
Period FE	X	X	X	X	X	X
N	212	211	203	188	165	133
Number of countries	35	35	35	35	35	35
Test for Aid_{is-1} ((6)-(1), $p-$ value)						0.069
Panel B: 2SLS - first differenced	(1)	(2)	(3)	(4)	(5)	(6)
# of maximal periods around the crossings	7	6	5	4	3	2
Aid_{is-1}	0.0485	0.0481	0.0460	0.0442	0.0427	0.0527
	(0.0177)***	(0.0177)***	(0.0181)**	(0.0157)***	(0.0192)**	(0.0222)**
Wild cluster bootstrap-t p-value	[0.018]	[0.018]	[0.018]	[0.016]	[0.019]	[0.022]
y_{is-1}	-0.0957	-0.0969	-0.103	-0.107	-0.113	-0.104
	(0.0387)**	(0.0389)**	(0.0397)***	(0.0369)***	(0.0441)**	(0.0488)**
Wild cluster bootstrap-t p-value	[0.039]	[0.040]	[0.041]	[0.038]	[0.045]	[0.048]
Period FE	X	X	X	X	X	X
N	212	211	203	188	165	133
Number of countries	35	35	35	35	35	35
First stage F statistic (K-P Wald)	16.159	16.353	15.741	19.946	16.577	19.971
Test for Aid_{is-1} ((6)-(1), $p-$ value)						1.000

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. The growth equation is first differenced before estimation. IV in the first differenced equation is $Crossing_{is-2}$. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors from the wild cluster bootstrap procedure are reported in brackets.

Table 6: Adding Covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	IV is $Crossing_{is-2}$						IV is $Crossing^{pred}_{is-2}$					
	baseline	schooling	political	CPIA	econ cond	credit rating	baseline	schooling	political	CPIA	econ cond	credit rating
Aid_{is-1}	0.0281	0.0307	0.0287	0.0336	0.0308	0.0303**	0.0352	0.0393	0.0368	0.0383	0.0456	0.0331***
	(0.010)***	(0.011)***	(0.0102)***	(0.0105)***	(0.011)***	(0.0125)	(0.0147)**	(0.0152)***	(0.0158)**	(0.0148)***	(0.0205)**	(0.0134)
y_{is-1}	-0.0371	-0.0402	-0.0361	-0.0232	-0.0119	-0.0740**	-0.0249	-0.0245	-0.0222	-0.0153	0.0101	-0.0330
	(0.0256)	(0.0273)	(0.0262)	(0.0241)	(0.0269)	(0.0303)	(0.0322)	(0.0332)	(0.0343)	(0.0321)	(0.0377)	(0.0378)
log population	-0.0086	-0.0149	-0.0114	0.0423	0.0133	-0.0824	0.0161	0.0203	0.0149	0.0584	0.0571	0.0067
	(0.0738)	(0.0753)	(0.0754)	(0.0767)	(0.0556)	(0.0857)	(0.0859)	(0.0869)	(0.0879)	(0.0903)	(0.0858)	(0.1030)
primary school enrolment rate		-0.0003						-0.0003				
		(0.0004)						(0.0005)				
Freedom House civil liberty			0.0023	-0.0036	0.0038				0.0021	-0.0045	0.0021	
			(0.0073)	(0.0062)	(0.0063)				(0.0078)	(0.0061)	(0.0071)	
Freedom House political rights			0.0013	0.0004	-0.0018				0.0026	0.0011	-0.0001	
			(0.0042)	(0.0040)	(0.0039)				(0.0044)	(0.0039)	(0.0041)	
World Bank CPIA-Z score				0.0151*						0.0156*		
				(0.0082)						(0.0082)		
total trade as % of GDP					0.0007						0.0006	
					(0.0005)						(0.0005)	
broad money					-0.0003						-0.0002	
					(0.0006)						(0.0006)	
inflation (GDP deflator)					-0.0000**						-0.0000**	
					(0.0000)						(0.0000)	
bank crisis during the period					0.0056						0.0045	
					(0.0117)						(0.0134)	
currency crisis during the period					-0.0023						0.0023	
					(0.0097)						(0.0110)	
debt crisis during the period					-0.0095						-0.0080	
					(0.0162)						(0.0154)	
credit rating						0.0008*						0.0005
						(0.0004)						(0.0004)
Period FE	X	X	X	X	X	X	X	X	X	X	X	X
Country FE	X	X	X	X	X	X	X	X	X	X	X	X
N	247	224	247	238	225	195	247	224	247	238	225	147
Number of countries	35	34	35	35	33	34	35	34	35	35	33	25
First stage F statistic (K-P Wald)	16.495	14.469	16.389	16.459	13.647	8.258	7.385	7.909	6.590	7.732	5.259	6.482

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. All columns are estimated using 2SLS. Columns 1 through 6 use actual crossings as instrument. Columns 7 through 12 use predicted crossings as instrument. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Model Robustness Checks

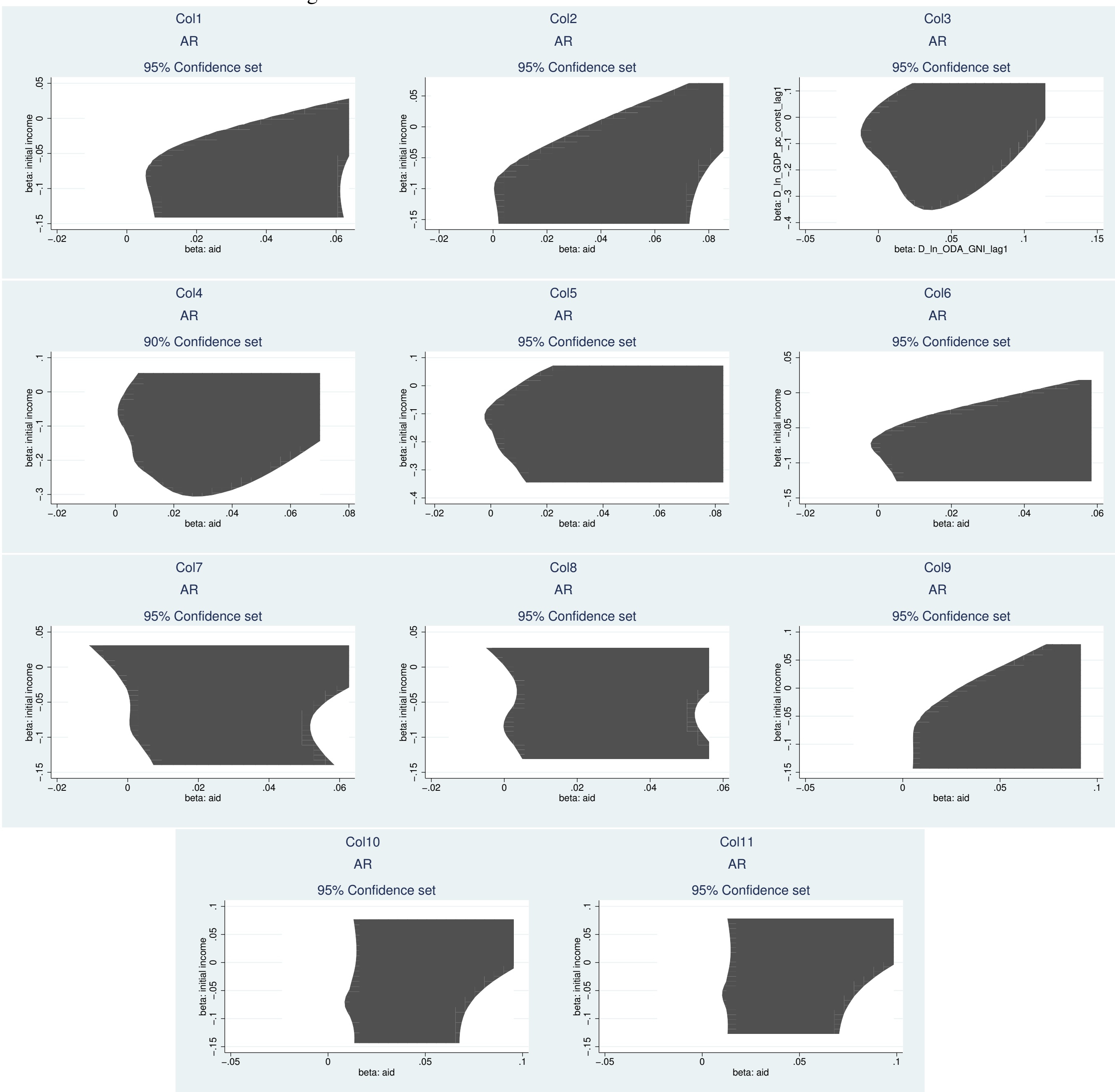
Panel A: IV is $Crossing_{is-2}$	(1) excl. last xing	(2) excl. last 2 xing	(3) multiple crossings	(4) last crossings	(5) small islands	(6) non IDA
Aid_{is-1}	0.0413	0.0451	0.0257	0.0241	0.0273	0.0299
	(0.0172)**	(0.0186)**	(0.00991)***	(0.0112)**	(0.0093)***	(0.0120)**
	[0.018]	[0.026]	[0.010]	[0.011]	[0.010]	[0.012]
y_{is-1}	-0.0194	-0.0200	-0.0717	-0.0440	-0.0396	-0.0194
	(0.0337)	(0.0366)	(0.0263)***	(0.0320)	(0.0251)	(0.0259)
	[0.035]	[0.042]	[0.026]	[0.032]	[0.025]	[0.026]
Period FE	X	X	X	X	X	X
Country FE	X	X	X	X	X	X
N	193	151	208	247	225	220
Number of countries	28	21	30	35	32	31
First stage F stat	8.635	20.06	18.48	19.04	14.39	11.97

Panel B: IV is $Crossing^{pred}_{is-2}$	(1) excl. last xing	(2) excl. last 2 xing	(3) multiple crossings	(4) last crossings	(5) small islands	(6) non IDA
Aid_{is-1}	0.0580	0.0883	0.0248	0.0258	0.0324	0.0382
	(0.0294)**	(0.0570)	(0.0138)*	(0.0159)	(0.0124)***	(0.0172)**
	[0.034]	[0.124]	[0.014]	[0.014]	[0.013]	[0.017]
y_{is-1}	0.00745	0.0387	-0.0736	-0.0411	-0.0308	-0.00346
	(0.0493)	(0.0861)	(0.0332)**	(0.0406)	(0.0289)	(0.0345)
	[0.057]	[0.157]	[0.034]	[0.032]	[0.030]	[0.034]
Period FE	X	X	X	X	X	X
Country FE	X	X	X	X	X	X
N	193	151	208	247	225	220
Number of countries	28	21	30	35	32	31
First stage F statistics (K-P Wald)	3.401	1.905	6.438	8.158	7.958	5.946

Note: Each observation is a country-period. The dependent variable is the period average real per capita GDP growth rate. Sample restrictions are marked in the short handle in each column. All columns are estiamted using 2SLS. Panel A uses actual crossings as the instrumental variable. Panel B uses predicted crossings as the instrumental variable. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild cluster bootstrapped standard errors are reported in brackets.

Appendix Figures and Tables

Figure A: Weak IV Robust Confidence Intervals for Table 4



Note: Anderson-Rubin confidence intervals robust for potential weak instruments for two endogenous variables. Each graph corresponds to a column in Table 4. Except for Column 4 which depicts the 90% confidence region, all other graphs show 95% confidence regions. The number of column increases first from left to right, then from top to down.

Table A: Construction and Sources of Key Variables

Variable	Notation	Source*	Description
Period	τ_s		Each period consists of 3 consecutive years. The first period is from 1987 to 1989.
IDA threshold		WB	Denoted in current US dollars, available since 1987.
Crossing 2 periods earlier	$Crossing_{is-2}$		Country i crossed the IDA cutoff for the first time in the sample at least two periods earlier.
Foreign aid	Aid_{is}	WDI/DAC	$Aid_{is} = \sum_s (ODA/GNI)/3$. Total net Official Development Aid (ODA) in current US dollars is from the DAC. GNI in current US dollars is from the WDI.
Initial income level	y_{is-1}	WDI	Real per capita GDP in 2000 constant US dollars in the second year of the last period, y_{it-4}.
Real per capita GDP growth	g_{is}	WDI	Denote real per capita GDP for country i in year t as y_{it}, annual real per capita GDP growth is $\ln(y_{it}) - \ln(y_{it-1})$. Period real pre capita GDP growth is the mathematical average of annual real per capita GDP for years in the period.
Aid by donor	$Aid_{j_{is}}$	DAC	Donor groups (j) include IDA, DAC countries, non- DAC countries, and multilateral agencies (MLA) except for IDA.
Investment		WDI	Gross capital formation as ratio of GDP. Investment in a period arithmetic average of annual gross capital formation as ratio of GDP.
Population+		WDI	Population
Primary school enrollment+		WDI	Gross primary school enrollment ratio. It is the total enrollment in primary education, regardless of age, expressed as a percentage of the population of official primary education age.
Trade+		WDI	Measured as merchandise trade as percentage of GDP.
Money supply+		WDI	Broad money as percentage of GDP
Inflation+		WDI	GDP deflator (percentage annual)
Crisis−		WB	Dummy variables indicating whether there is any bank, currency, or debt crisis during the years within the period.
Bank crisis−		WB	Whether the country experiences a bank crisis.
Currency crisis−		WB	Whether the country experiences a currency crisis.
Debt crisis−		WB	Whether the country experiences a debt crisis.
Political rights+		FH	Freedom House political rights indicator. It ranges from 0 to 7, with a higher number indicating less political rights.
Civil liberties+		FH	Civil liberties indicator. It ranges from 0 to 7, with a higher number indicating less civil liberty.
Bureaucratic quality+		ICRG	0-4, with a higher number indicating less risk in bureaucratic quality.
Corruption+		ICRG	0-6, with a higher number indicating less risk in corruption.
Rule of law+		ICRG	0-6, with a higher number indicating less risk in rule of law.
Ethnic tension+		ICRG	0-6, with a higher number indicating less risk in ethnic tension.
CPIA z-score+		WB	Public sector management and institutions cluster average. 1=low to 6=high.
Credithworthiness+		II	0 (low) -100 (high), based on survey of international bankers.

* WB is short for the World Bank; WDI is short for the the World Development Indicators from the World Bank; DAC represents the OECD Development Assistance Committee. FH is short for the Freedom House. ICRG is short for International Country Risk Guide. II is short for Institutional Investors.

+ averaged within each period.

− summed over each period.

Table B: Summary Statistics

Variable	N	mean	$s.d.$	25^{th}	50^{th}	75^{th}
real GDP per capita growth	247	.029	.054	.007	.028	.048
log real GDP per capita 4 years earlire	247	6.635	.551	6.274	6.666	6.988
lag of log(ODA/GNI)	247	-3.282	1.472	-4.473	-2.910	-2.184
lag of log(IDA/GNI)	247	-3.282	1.472	-4.473	-2.91	-2.184
lag of log(DAC/GNI)	247	-10.618	8.022	-21.043	-6.256	-4.689
lag of log(NDAC/GNI)	247	-3.774	1.574	-4.941	-3.356	-2.697
lag of log(Other MLA/GNI)	247	-11.281	6.073	-13.267	-9.246	-6.641
lag of ODA/GNI	247	.081	.094	.011	.054	.113
lag of IDA/GNI	247	.007	.012	0	.002	.009
lag of DAC/GNI	247	.054	.068	.007	.035	.067
lag of NDAC/GNI	247	.002	.005	0	0	.001
lag of other MLA/GNI	247	.019	.027	.002	.009	.024
crossed IDA threshold 2 periods earlier	247	.2308	.4222	0	0	0
lag of Investment/GDP	231	0.253	0.119	0.187	0.235	0.297
log population	247	15.950	2.259	14.765	15.874	17.108
lag of terms of trade (year 2000=100)	167	100.548	15.894	92.328	99.983	104.495
CPIA z-score	238	-.214	.934	-.876	.0192	.439
civil liberty	247	4.306	1.604	3	4	6
political rights	247	4.273	1.981	2	4	6
primary school enrollment	224	97.508	19.290	91.358	101.563	110.337
merchandized trade as % of GDP	247	64.469	31.215	41.355	60.568	83.855
broad money as % of GDP	225	37.383	25.520	18.216	31.825	50.164
inflation (%)	247	94.608	537.018	5.107	8.726	18.506
bank crisis (dummy)	247	.052	.224	0	0	0
currency crisis (dummy)	247	.109	0.313	0	0	0
debt crisis (dummy)	247	.020	.141	0	0	0
creditworthiness	195	28.116	13.458	18.367	26.2	33.933

Note: Each observation is a country-period. For each variable, the mean, standard deviation, median, 25^{th}percentile, and 75^{th} percentile are reported. Missing values for ODA and ODA by donor are treated as zeros, following the precedent of Arndt, Jones and Tarp (2010) (Page 14). In this sample there are no missing values in total ODA. Zero values in aid by donor are replaced with 1 dollar before taking logarithm. See Appendix Table A for more details in construction and sources of these variables.

Table C: More Crossing Dummies

	(1)	(2)	(3)	(4)	(5)
	First Stage	Reduced Form		TSLS	
dep var	Aid_{is-1}	g_{is}	g_{is}	g_{is}	g_{is}
Aid_{is-1}			0.0219***	0.00687	0.0224**
			(0.00835)	(0.0177)	(0.00873)
$Crossing_{is-1}$	-0.368*	0.00477			
	(0.190)	(0.0115)			
$Crossing_{is-2}$	-0.483***	-0.0226**			
	(0.112)	(0.00905)			
$Crossing_{is-3}$	-0.563***	-0.00736			
	(0.181)	(0.00866)			
y_{is-1}	-1.475***	-0.0817***	-0.0478*	-0.0738**	-0.0469
	(0.355)	(0.0260)	(0.0245)	(0.0288)	(0.0289)
Period FE	X	X	X	X	X
Country FE	X	X	X	X	X
IV for aid, lags of crossing			1,2,3	1	3
N	247	247	247	247	247
Number of countries	35	35	35	35	35
First state F statistic (K-P Wald)			7.674	5.919	14.49

Note: Column 1 reports the first stage with multiple crossing dummies. Column 2 reports the reduced form with multiple crossing dummies. Column 3 reports the 2SLS estimation using multiple crossing dummies as instrumental variables. Log population controlled for in all specifications. Standard errors in parentheses, clustered at the country level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D: IDA Threshold and Aid with Quadratic Log Initial Income Levels

	(1)	(2)	(3)	(4)	(5)
	IDA	DAC	NDAC	MLA	ODA
$Crossing_{is-2}$	-2.960	-1.020	-2.747	-1.027	-0.940
	(1.264)**	(0.272)***	(1.725)	(0.420)**	(0.243)***
	[1.096]	[0.237]	[1.495]	[0.362]	[0.209]
y_{is-1}	-30.28	-4.133	-28.67	-15.12	-4.462
	(21.24)	(3.995)	(17.17)	(8.215)*	(3.120)
y^2_{is-1}	1.562	0.194	1.724	0.909	0.211
	(1.464)	(0.277)	(1.147)	(0.555)	(0.217)
Country FE	X	X	X	X	X
Period FE	X	X	X	X	X
N	247	247	247	247	247

Note: Each observation is a country-period. Dependent variables are the log average share of ODA in GNI by donor in the last period, Aid_{is-1}. 35 countries are in the sample. Country fixed effects, period fixed effects, and log population in the last period are controlled in all columns. Standard errors are in parentheses, clustered at the country level.$*p < 0.10$, $**p < 0.05$, $***p < 0.01$. Standard errors from the wild cluster bootstrap are also reported.

Table E: IDA Threshold and Aid - Placebo Threshold at 50% of the True Level

	(1)	(2)	(3)	(4)	(5)
	IDA	DAC	NDAC	MLA	ODA
$Crossing_{is-2}$	3.068	0.404	2.565	-0.118	0.299
	(1.867)	(0.661)	(2.396)	(0.383)	(0.514)
	[1.621]	[0.587]	[2.175]	[0.395]	[0.472]
y_{is-1}	-12.29	-2.251	-8.808	-4.622	-2.245
	(3.778)***	(0.859)**	(4.367)*	(1.507)***	(0.657)***
	[3.200]	[0.685]	[3.737]	[1.273]	[0.506]
Period FE	X	X	X	X	X
Country FE	X	X	X	X	X
N	162	162	162	162	162
N countries	34	34	34	34	34
F-statistic on $Crossing_{is-2}$					0.57

Note: Each observation is a country-period from countries that crossed the IDA threshold between 1987 and 2010. Country-period observations included in the sample are prior to the period of crossing the real threshold. There are 34 countries and 162 observations in the regression. Dependent variables are one period lag of average shares of ODA in GNI by donor. Country and period fixed effects, and log population in the last period are controlled in each column. F-statistic on $Crossing_{is-2}$ is reported in Column 5. y_{is-1} is log per capita real GDP in the second year of the last period, y_{it-4}. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Wild cluster bootstrapped standard errors are reported in brackets.

Table F: First Stage - Policies as Outcome Variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Freedom House		Word Bank		Inflation		Crisis		credit
	civil liberty	political rights	CPIA z-score	broad money	(GDP deflator)	bank	currency	debt	rating
$Crossing_{is-2}$	0.1255	0.2955	0.1032	0.2007	-21.2578	0.0028	0.0482	0.0427	2.0734
	(0.2527)	(0.4039)	(0.1867)	(2.6788)	(100.4773)	(0.0404)	(0.0489)	(0.0292)	(3.7842)
	[0.226]	[0.352]	[0.163]	[2.364]	[92.780]	[0.037]	[0.044]	[0.027]	[3.292]
y_{is-1}	-0.0573	-0.1061	-0.2203	7.7313	483.0089	0.0668	0.0229	0.0617	1.4058
	(0.1675)	(0.2603)	(0.1624)	(8.6056)	(385.9596)	(0.0642)	(0.0744)	(0.0511)	(3.8376)
	[0.146]	[0.228]	[0.142]	[7.694]	[343.112]	[0.058]	[0.067]	[0.047]	[3.369]
Country FE	X	X	X	X	X	X	X	X	X
Period FE	X	X	X	X	X	X	X	X	X
N	255	255	244	232	255	255	255	255	197
# of countries	35	35	35	33	35	35	35	35	34

Note: Each observation is a country-period. The dependent variable for each column and its source are indicated on top of the column. See Appendix Table A for description of these variables. All columns are estimated using 2SLS. Log population one period lag is included as a covariate. Standard errors in parentheses, clustered at the country level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors from the wild cluster bootstrap procedure are reported in brackets.

Table G: Functional Form of Aid

	(1)	(2)
Aid^{*}_{is-1}	0.573	0.525
	(0.264)**	(0.251)**
	[0.252]	[0.237]
y_{is-1}	-0.0211	-0.0381
	(0.0475)	(0.0384)
	[0.048]	[0.039]
Period FE	X	X
Country FE	X	X
Observations	247	247
# of Countries	35	35
First Stage F statistic (K-P Wald)	8.062	4.111

Note: $Aid^{*}_{is-1} = \sum_{k=3}^{5}(ODA_{it-k}/GNI_{it-k})/3$ is the period average aid to GNI ratio in the last period. There are 212 country-period observations from 35 countries. 2SLS estimator is used in both columns. In both columns, ΔAid^{*}_{is-1} is instrumented with $\Delta Crossing_{is-2}$. Δy_{is-1} is also treated as endogenous in Column 2 and y_{it-8} is used as an additional instrumental variable. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors from the wild cluster bootstrap procedure are reported in brackets.

Table H: Baseline Results with Polynomials of Initial Income Level

	(1)	(2)	(3)
Aid_{is-1}	0.0281	0.0297	0.0342
	(0.0100)***	(0.0105)***	(0.0148)**
	[0.01]	[0.01]	[0.014]
Polynomials of y_{is-1} included	1	1,2	1,2,3
Period FE	X	X	X
Country FE	X	X	X
Observations	247	247	247
# of Countries	35	35	35
First stage F statistic (K-P Wald)	16.495	14.935	10.560

Note: Column 1 replicates Column 3 of Table 3. Column 4 and Column 5 adds quadratic and cubic terms of y_{is-1}. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors from wild cluster bootstrap are reported in brackets.

Table I: Effects of Aid on Investment

Dep var $(Inv/GDP)_{is-1}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Main Specification	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Aid_{is-1}	-0.0166	-0.0103	0.0580	0.0490	0.0542	0.0831	0.0719
	(0.0150)	(0.0158)	(0.0350)*	(0.0333)	(0.0511)	(0.0624)	(0.0465)
	[0.013]	[0.015]	[0.035]	[0.033]	[0.052]	[0.063]	[0.046]
y_{is-2}	-0.0638	-0.0962	0.0308	0.0194	-0.0503	-0.0298	-0.0378
	(0.0392)	(0.0339)***	(0.0514)	(0.0523)	(0.0409)	(0.0461)	(0.0399)
	[0.036]	[0.034]	[0.052]	[0.053]	[0.042]	[0.047]	[0.041]
Period FE	X	X	X	X	X	X	X
Country FE	X		X	X			
First differenced		X			X	X	X
IV			X	X	X	X	X
IV predicted				X			X
IV first differenced					X		
Equation first differenced		X			X	X	X
N	206	171	206	206	171	171	171
Number of countries	34	34	34	34	34	34	34
First stage F statistic (K-P Wald)			9.223	9.714	10.94	4.724	10.51

Note: Each observation is a country-period. The dependent variable is the period average investment to GDP ratio. Standard errors clustered at the country level are reported in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors from the wild cluster bootstrap are in brackets. See text for more details.

Table J: Growth Trajectories of Crossing and Non-Crossing Countries

	(1)	(2)	(3)
Crossing sample = 1	0.00669	0.00672	
	(0.00669)	(0.00698)	
Crossing sample = 1 $\times$			
1^{st} quartile of y_{it-1}			0.03078
			(0.00606)***
2^{nd} quartile of y_{it-1}			0.01327
			(0.00868)
3^{rd} quartile of y_{it-1}			-0.00049
			(0.01047)
4^{th} quartile of y_{it-1}			0.00739
			(0.01018)
y_{it-1}	0.00799	0.00940	0.01848
	(0.00656)	(0.13174)	(0.14883)
y^2_{it-1}		-0.00012	-0.00085
		(0.01133)	(0.01294)
Year FE	X	X	X
N	1303	1303	1303

Note: Each observation is a country-year. The dependent variable is annual log per capita real GDP growth. There are 78 countries that were ever eligible for IDA between 1987 and 2010 as well as the 35 countries in our baseline sample. The key variable of interest is a dummy variable indicating whether the country belongs to the crossing sample. The sample consists of country-year observations between 1987 and 2010 that have per capita GNI level below the IDA threshold. Year fixed effects are controlled. Log real GDP per capita in the last year and its quadratic terms are included in the regressions. Standard errors are reported in parentheses, clustered at the country level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.