# Aid Effectivenness. New Instrument, New Results

Maria Perrotta Berlin* Emmanuel Frot†

September 17, 2018

**Abstract**

Despite a voluminous literature on the topic, the question of whether aid leads to growth is still controversial. To observe the pure effect of aid, researchers use instrumental variables that must be exogenous to growth and explain well aid flows. This paper argues that instruments used in the past do not satisfy these conditions. We propose a new instrument based on *predicted* aid quantity and argue that it is a significant improvement relative to past approaches. We find a significant and positive effect of aid: the elasticity of GDP per capita growth with respect to aid is found to be 0.3. While this effect is small, it is three times larger than previously estimated in the literature. We compute that, had aid not been disbursed since 1963, GDP per capita in the average developing country would be today about 30% lower. The average developing country citizen would be about 15% poorer, while the median citizen would be 7% poorer.

## 1 Introduction

Foreign aid has been disbursed for decades and is still currently seen as a major tool of development policy. Aid is also an important component of recipient countries budget. In many cases, it can exceed half of GDP. While promises of increasing aid flows are not all fulfilled, and in some cases even under threat of retraction, the trend is still towards expanding aid budgets[1].

---

*SITE, SSE, P.O. Box 6501, SE-113 83 Stockholm, Sweden. Email: maria.perrotta@hhs.se.

†Microeconomix, 5 rue du Quatre Septembre, 75 002 Paris, France; SITE, SSE, P.O. Box 6501, SE-113 83 Stockholm, Sweden. Email: emmanuel.frot@microeconomix.com.

[1]The OECD Development Co-operation Directorate notes that aid flows from DAC donor countries (developed countries that represent a very large share of global aid) totaled USD 146.6 billion in 2017, the highest level ever.

The academic community, however, has not found any robust evidence that aid contributes to development. Some take issue with the faulty empirical strategies employed, some question the very relevance of the question. In particular, the formulation of the problem in terms of aggregate measures of aid and aggregate outcomes such as growth is argued to make a serious empirical investigation unecessarily hard. Disaggregated components of aid flows, or specific aid programs and interventions, as well as particular well defined outcomes would constitute a better object of analysis, for a number of reasons.[2] The more recent impact evaluation literature is moving in this direction.

The contribution of this paper is within the literature on the impact of *aggregate* foreign aid, and it consists mainly of a new instrumental variable. Before arguing on its merits and faults, we jump ahead to the results. We find a significant and positive effect of aid on growth: 1% increase in foreign aid is associated with a significant within-country increase in GDP per capita of around 0.04 percent. In terms of aid to GDP share, a one percentage point increase in the aid share is associated with 0.32 percent increase in GDP per capita. But how large is this effect? How much poorer would developing countries be today, had they never received any aid? Taking our estimates literally, it is possible to compute the growth rate that would have prevailed, had no aid been disbursed at all.[3]

The average GDP per capita growth rate in our data is 2.3%. Without aid, it would have been -1.4%. Figure 1 illustrates the estimated density of growth rates under the actual and the counterfactual scenarios.

The shift in the distribution is non negligible. Many countries would have had negative growth rates over the period 1963-2013 but for aid.

In order to estimate the impact of aid from 50 years of disbursement, we compute the ratio of the counterfactual 2013 GDP to its actual counterpart. The average aid recipient in the regression sample would have a GDP per capita equal to just 57% of its actual level if aid had not been disbursed since 1963. We also looked at population-weighted mean ratios. These represent the effect of aid on the average developing country *citizen*. This potentially makes a significant difference as the average citizen is 26% Chinese and 21% Indian and these countries received little aid in GDP terms. It also reduces the influence of small countries with large aid shares as these usually have small populations. The population-weighted means reveal that without aid the average developing country citizen would be about 5% poorer in 2013. The average citizen in our sample countries would be 11% poorer.

Of course, this exercise is quite demanding for our reduced-form model. Claiming

[2]See for example Qian (2014). We defend our choice of variables later in section 4.

[3]The calculation is detailed in the Appendix.

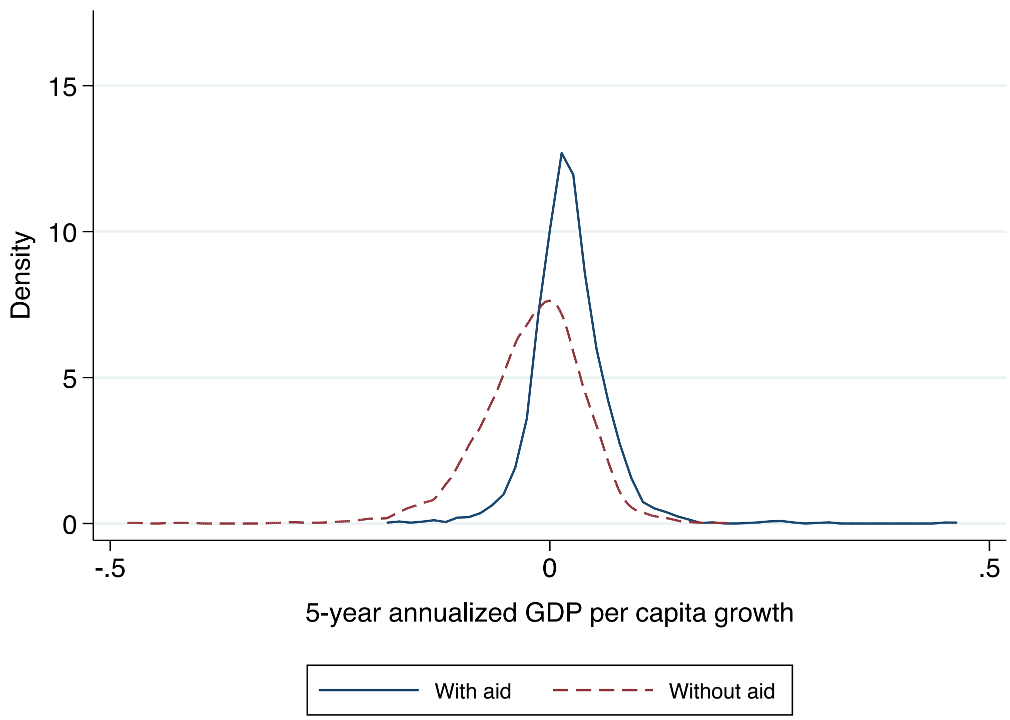


Figure 1: GDP per capita growth density

to being able to approximate the consequences of a shift from all past aid disbursements to none may seem far-fetched. With this we merely mean to represent visually what our estimates imply and to ground the relevance of the research question in a tangible application.

The rest of the paper is organized as follows: in the next section, we draw comparisons with earlier results in the literature. Following, in Section 3 we describe in detail how our instrument is built; we then briefly discuss our methodological choices in terms of data and estimators in Section 4 and present the results in Section 5. In Section 6, the robustness of the results is assessed. . Finally, Section 7 concludes the paper. All variable definitions and data sources can be found in a data appendix at the end of the paper.

## 2 A Brief and Selective Overview of the Literature

The aid effectiveness literature is large and mostly inconclusive. The results vary widely in size and sign, and have often been proven not robust or reversed by new estimations. Even limiting ourselves to the most recent reviews of the field we can observe a broad variety of conclusions. In the popular "The Great Escape", Princeton's Angus Deaton dismisses the idea that aid might contribute to growth as an illusion, that in practice hinders improvements in the lives of the poor (Deaton, 2013). But in his review of the book Georgetown professor (and former director of the World Bank's esearch department) Martin Ravallion retorts that instead "an objective assessment of the evidence (...) suggests a more positive role for aid in the developing world's ongoing efforts to escape poverty" (Ravallion, 2014). Recent overviews also disagree, with conclusions ranging from "aid has an insignificant or minute negative significant impact on per capita income" (Nowak-Lehmann et al., 2012), to "[*our review*] sugests a positive and statistically significant long-run effect of aid on income" (Lof et al., 2014), and from "the macro evidence in the recent (post-2008) literature is more supportive of positive impacts" (Arndt et al., 2014) back to "the improvement is an artifact (of publication bias). [...] the average effect of aid on growth is trivially small" (Doucouliagos and Paldam, 2014).

The reason for this lack of consensus is that, in the words of Clemens et al. (2012), "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."[4] A few recent alternative approaches are very promising. Brückner

[4]Clemens et al. (2012) also offer a critical review of the latest contributions in terms of IV strategies, which we do not go through here due to space constraints.

(2013) estimates separately the effect of growth on aid, getting at reverse causality, and then removes it in the growth equation. Clemens et al. (2012) treat (lagged) aid as exogenous, predetermined with respect to future shocks to growth, after controlling for country fixed effects. Finally **?** rely on a quasi-experiment, the discontinuity around an income threshold used by the World Bank to allocate its concessional lending. Although all these strategies are quite compelling, they provide narrower inference, limited to specific samples or demading in terms of time length of the data.

Among the IV studies in this literature, the dominant approach makes use of variables linked to the aid allocation process (the "supply side" of foreign assistance, mostly historical and political variables), and predict aid flows based on them.[5] We propose a new instrument and argue that it is a significant improvement relative to past approaches. It takes the "supply side" approach one step further. Our identification strategy is similarly based on predicted aid flows; however, unlike existing studies, we exploit a source of variation that we argue not to be subject to the same criticisms. This source of variation is related to the temporal order in which donor-recipient partnerships are established: Frot (2009) shows that *when* a partnership is established and *how long* it lasts are of importance for aid quantities. Our identification strategy relies on the observation that partnerships formed earlier entail higher aid levels. Recipient countries engaged in these earlier partnerships therefore received relatively larger aid quantities. To be valid, this strategy requires partnership formation to be exogenous to subsequent growth, beyond the effect of the larger aid receipts. We provide evidence that makes us confident it is. In addition, we show that our instrument is highly correlated with actual aid levels as weak instrumentaton was common in previous studies.[6]

Before moving on to next section's detailed description of the instrument and our methodology, another way to relate our findings to the literature is to compare the magnitude of our results to previous estimates. The effect identified here is comparable to what found by the three recent contributions listed above: it is roughly the same magnitude as found in **?**, while Clemens et al. (2012)'s effect is about half as large, and Brückner (2013)'s smaller yet (1% increase in foreign aid associated with a significant within-country increase in GDP per capita growth of around 0.1 percentage points). As for older, but still widely cited and influential studies, the elasticities estimated in Dalgaard et al. (2004), Hansen and Tarp (2001) and Burnside and Dollar (2000) are all quite close in size, ranging from 0.09 respectively to 0.1 and

[5]In the working paper version of this paper, Frot and Perrotta (2010), we go through the most influential and widely cited contributions.

[6]Rajan and Subramanian (2008) and Bazzi and Clemens (2009), among others, review the literature and question the validity of the instruments used in past studies.

0.18.[7] The already mentioned meta-analysis by Doucouliagos and Paldam (2009) claims that "the best estimate we can make of the elasticity of the real product to aid is about 0.13". Our effect is, in other words, close to the upper bound of the range of estimates in the literature.

# 3 The instrument

This section focuses on describing in more detail our new instrument. Earlier research based on the supply-side approach looks at the determinants of different countries' aid receipts. Total aid $A_{it}$ to recipient $i$ in year $t$ can be decomposed as

$$A_{it} = \sum_j s_{ijt} D_{jt} \tag{1}$$

where donors are indexed by $j$, $D_{jt}$ is donor $j$'s total aid budget in year $t$ and $s_{ijt}$ is the share of this budget allocated to recipient $i$. Previous work by Frot (2009) analyzes the behavior of the shares $s_{ijt}$ over time, and we reproduce some of his results here. In Figure 2 recipients are grouped into six cohorts based on their entry dates into donors' portfolios: recipients with an entry date of one receive aid from a donor in the first year of the donor's activitiy, and so on. The plot presents the average *normalized*[8] share received by recipients in each cohort in each year.[9] In other words, it shows how much recipients in each cohort get in deviation from equal sharing. It is important to notice that the same recipient belongs to different cohorts as relative to different donors. For example, Afghanistan is in cohort 2 with respect to Finland's portfolio and in cohort 11 for the European Community; so the aid share of Finland to Afghanistan and the aid share of EC to Afghanistan enter two different averages in Figure 2.

As shown by the figure, early entrants into donors' portfolios receive on average persistently larger aid shares. There is some convergence across cohorts but even many years after portfolios were formed, it is still the case that entry dates and aid shares are correlated. Stratification by cohorts is visible in any year. This fact is the main basis for the construction of our instrument.

[7] This last result only refers to the subsample of countries with good policies, according to the definition in Burnside and Dollar (2000).

[8] See definition in the Appendix.

[9] In Figure 2 the sample is restricted to donors that have been present from 1960 to 2007 in order to make recipient coohorts comparable. In general donors can enter, and sometimes exit, the market in different years.

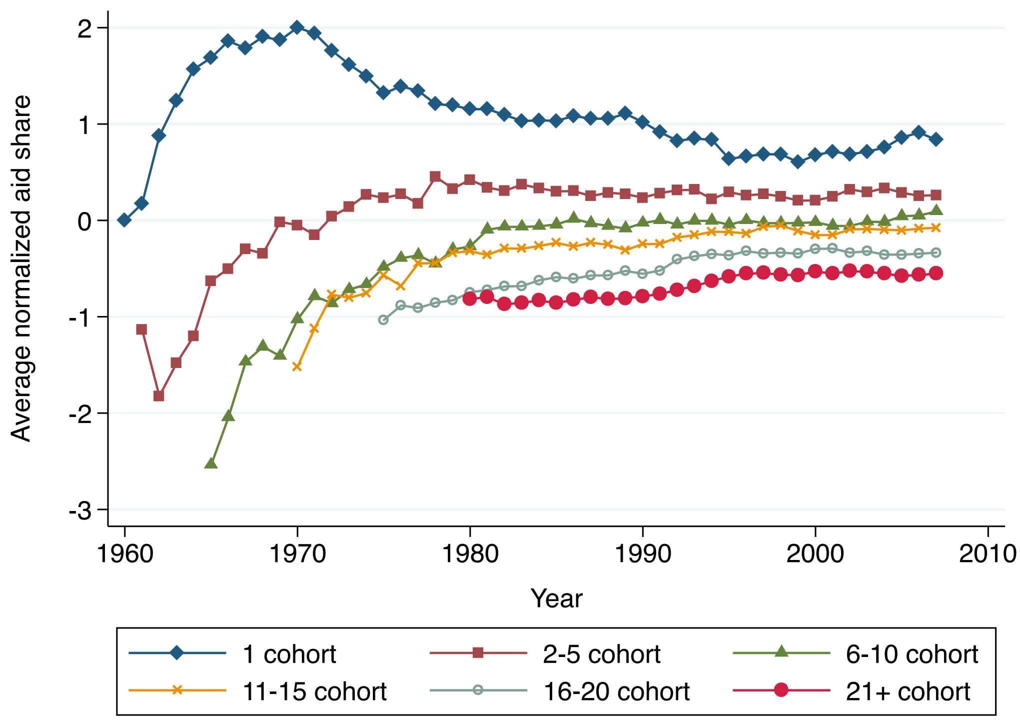


Figure 2: Average aid share in deviation from equal sharing, by recipient cohort

## 3.1 Design

Each donor-recipient pair $(i, j)$ in a given year $t$ is characterized by two features, related to the date when the partnership was established: the entry date order in a sequence of all partnerships established by the donor, and how long ago this partnership was created. We call $\kappa_{ij}$ country $i$'s entry date position in an ordered sequence of all partnerships established by donor $j$. For instance, $\kappa_{ij} = 1$ for recipients that received aid from $j$ in the first year $j$ started to give aid, and so on.[10] More formally, let $\eta_{ij}$ be the entry date for the pair $(i, j)$, or the first year in which donor $j$ gives aid to recipient $i$, and let $\pi_j$ be the first year in which donor $j$ disburses aid to any country. The entry date order $\kappa_{ij}$ is then defined as

$$\kappa_{ij} = \eta_{ij} - \pi_j + 1 \tag{2}$$

Let $\tau_{ijt}$ be the difference between $t$ and the entry date.[11] Predicted aid shares are then the fitted values of the regression

$$\sigma_{ijt} = a + b\kappa_{ij} + c\tau_{ijt} + u_{ijt} \tag{3}$$

Predicted aid shares for any observation (i.e. a given $(i, j)$ pair in a given year) are fully defined by their entry date order and partnership length. In other words, to any partnership characterized by entry date of order $\kappa$ and length $\tau$ we associate a predicted aid share $\hat{\sigma}_{\kappa\tau} \equiv \hat{a} + \hat{b}\kappa + \hat{c}\tau$ that is not related to $i$, $j$, or $t$. $\hat{\sigma}_{\kappa\tau}$ is the *typical* share (in fact, the average share) that any recipient gets from a donor if their partnership was established in the $\kappa^{th}$ year of activity of this donor, $\tau$ years ago.

The instrument for aid is finally the predicted aid quantity

$$\hat{A}_{it} = \sum_j \hat{\sigma}_{\kappa_{ij}\tau_{ijt}} D_{jt} \tag{4}$$

In words, we start from the aid share that our model predicts each donor allocates to each recipient, based on the partnership characteristics. We then multiply these predicted aid shares by the donors' aid budgets to obtain a predicted aid quantity for each recipient and year. The intuition is as follows. The instrument artificially recreates a situation where a country receives more aid in a given period, independently

[10] To be precise, our data only starts in 1960, so the ordered sequence of recipients' cohorts is approximate. This is a data limitation which is akin to censoring, but on an independent variable; the econometric literature has surprisingly little to say about how to deal with this issue; see Manski and Tamer (2002) and Rigobon and Stoker (2009) for contributions.

[11] We call this *length*, but it is rather a predicted length, given that we do not take into account whether the partnership was ended at some point in time since this would be an endogenous decision.

of the "fundamentals" of its economy, but rather for one or more of the following reasons: because it on average had an earlier order of entry with respect to other recipients in its average donor's portfolio; because it was in its average partnership for a longer period of time[12]; finally, because its average donor's budget for aid happened to be larger that year.

In section 5.3, we will argue that predicted aid is a strong instrument, relevant for predicting actual aid flows, and that its only effect on growth occurs through the actual aid flows it proxies. But first we describe our empirical specification.

# 4 Data and specification

Table A.5 in the Appendix reports our estimation samples. The largest sample includes 108 countries.[13] In some specifications, we control for a standard set of variables established in the literature.[14]This allows us to draw comparisons with past studies.

In order to avoid business cycle fluctuations and also for consistency with the literature, we focus on 5-year periods. Time period 1 represents years 1960-1965, and so on. The dependent variable, GDP growth, is the annualized growth over the five-year period. For aid, we sum the flows over the period.[15] All other variables are five-year arithmetic averages.

[12]Notice that even length is in some sense predicted, in that we make no use of exit dates or partnership terminations. This would reintroduce endogeneity in our instrument, because if the donor decides to terminate a partnership this obviously reflects the recipient conditions. Here instead length represents rather how long time has passed since the entry date, at the moment we observe the country, to allow for a linear time trend in the behavior of the aid share.

[13]The results are robust to excluding very small countries, with a population lower than 100,000: Antigua and Barbuda, Dominica, Kiribati, Seychelles, and St. Kitts and Nevis.

[14]Population size; the Barro and Lee (2010) average years of primary schooling (whether we use primary or secondary schooling does not make much difference) as a measure of investments in human capital; inflation as a measure of macroeconomic policies; liquid assets (M2/GDP), commonly used as a measure of financial depth; institutional quality, measured by the International Country Risk Guide (ICRGE) index; a commonly used measure of trade openness, the ratio of total trade (exports plus imports) to GDP.

[15]See discussion in Section 6.2

The empirical specification is hence as follows:

$$\begin{aligned}\Delta lny_{it} =& \beta_1 lny_{it-5} + \beta_2 ln(\sum_{t-10}^{t-5} a_{is}) + \beta_3 lnpop_t + \beta_4 M_{t-5}^{t-1} ln(1+infl_{is}) + \beta_5 M_{t-5}^{t-1} m2_{is} \\ &+ \beta_6 school_{it-5} + \beta_7 M_{t-5}^{t-1} ICRG_{is} + \beta_8 M_{t-5}^{t-1} open_{is} + \beta_9 reg + \alpha_t + \mu_i + \xi_{it}\end{aligned} \quad (5)$$

where $M_{t-5}^{t-1}$ indicates the arithmetic average over the time period, $i$ indexes recipients, $j$ donors, and $t$ years. A detailed description of the variables with sources is in the appendix. All regressions in the paper include year effects $\alpha_t$ and recipient fixed effects $\mu_i$.

# 5 Results

## 5.1 Preliminary stage

As mentioned above, our strategy consists in first predicting aid shares by regressing actual shares on entry date order and partnership length (and their squares).[16] We then compute predicted aid quantities $\hat{A}_{it}$ by summing the predicted aid shares multiplied by donors' aid budgets. The predicted aid quantity is then used in the "second stage" growth regression (equation (5)).

## 5.2 Second stage

The log-log specification adopted in equation (5) implies that the coefficient on the aid variable can be interpreted as the elasticity of GDP per capita with respect to aid. We start by not instrumenting the aid variable, and present naive estimates, with and without country fixed effects in Table 1. Column (1) confirms the traditional finding that, when not instrumented, aid has no effect on GDP growth. The inclusion of country fixed effects only reinforces this conclusion. However, we know there is little to learn from regressions where aid is not instrumented. We move on to column (2) where aid is instrumented using our predicted aid quantities. [17] The consequence

[16] Section 5.1 goes through the details of the preliminary stage. The specification we use to predict aid shares corresponds to Table 3 column (4).

[17] Because a major concern in the literature is the weakness of instrumentation for aid, we provide two statistics. The first is the $p$-value of the Angrist and Pischke (2009) test of excluded instruments. With a single endogenous regressor, this is simply the $F$-statistic of the first stage. The second is the Kleibergen and Paap (2006) Wald statistic. Both are tests of instrument weakness, therefore a failure to reject the null hypothesis would cast doubts on the choice of instrument.

Table 1: OLS and 2SLS regressions

| | (1) FE OLS | (2) FE IV |
|---|---|---|
| Aid | 0.0052 | $0.040^{*}$ |
| | (0.0028) | (0.019) |
| GDP p/c | $-0.038^{***}$ | -0.019 |
| | (0.0063) | (0.010) |
| Population | $0.00016^{***}$ | $0.00024^{***}$ |
| | (0.000021) | (0.000062) |
| Inflation | $-0.037^{***}$ | $-0.020^{*}$ |
| | (0.0057) | (0.0087) |
| M2/GDP | 0.00000045 | 0.000016 |
| | (0.0000016) | (0.000011) |
| School | 0.0030 | 0.0093 |
| | (0.0043) | (0.0063) |
| Openness | 0.025 | 0.014 |
| | (0.015) | (0.0088) |
| Observations | 708 | 701 |
| Countries | 109 | 109 |
| KP F stat | | 9.60 |
| KP p-val | | 0.0019 |
| KPW F stat | | 9.49 |

Note: KP: Kleibergen-Paap LM underidentification test. KPW: Kleibergen-Paap Wald weak identification test.The dependent variable is the annualized growth rate in per capita income over the five- or three-year period. All the regressions include year effects. Robust standard errors clustered at the recipient level in parentheses. * (p<0.10), ** (p<0.05), *** (p<0.01).

for the aid coefficient is quite dramatic. It is much larger than in column (1) and comfortably passes the 5% significance threshold. The Kleibergen and Paap Wald statistic is quite high, suggesting that predicted aid is not a weak instrument for aid.[18] The null hypothesis of the Angrist and Pischke test is strongly rejected, too. The first-stage is shown in column (1) of Table 2. It confirms that our instrument is a strong predictor of actual aid. What the 2SLS procedure does is regressing actual aid (which is actual shares multiplied with donors' budgets) on predicted aid (which is predicted shares multiplied with the same donors' budgets) in the first stage. Because the instrument is obtained in a preliminary stage where aid shares are already regressed on the instrumenting variables observed at the bilateral level, this classical first stage seems redundant. Due to the dimension of the preliminary stage and second stage variables, though, we cannot use a normal two-stage procedure with the instrumenting variables directly entering the second stage. As an alternative method, we introduce the predicted aid variable directly in equation (5), replacing actual aid. The standard errors are bootstrapped to take into account the fact that predicted aid is estimated in the preliminary stage. The resulting point estimates and standard errors are similar, which supports the use of standard 2SLS procedures.[19]

The estimated effect implies an elasticity of GDP per capita with respect to aid around 0.04. This elasticity is relatively small: a 1% increase in disbursed aid increases per capita GDP by, on average, 0.04%. So doubling aid disbursements, arguably a large shock, would increase GDP growth by 4%. More realistically, raising the ratio of aid to GDP from the 11.5% mean to say 13.5%, which would require a 25% increase in aid disbursements, would lead to an associated growth rate in per capita GDP of around 0.75%.

## 5.3 Discussion

### 5.3.1 Strength of the instrument

For predicted aid $\hat{A}_{it}$ to be a good instrument, it must be the case that entry date order and length are strong determinants of aid shares. Figure 2 alone does not offer enough evidence that entry date order plays a decisive role in determining aid shares, neither does it exclude the case that other factors are behind the correlation between entry date order and aid receipts. It is likely that donors created partnerships prioritizing poor and heavily populated countries, and that such countries have

[18] Although critical values only exist for the Cragg-Donald Wald statistic, which is not robust to heteroskedasticity, the 25% maximal IV size value is 5.53.

[19] Standard procedures have the advantage of offering established diagnostic tools. The IV estimates are not reported but can be obtained upon request.

Table 2: First stages

| | (1) Aid | (2) Aid | (3) Trade |
|---|---|---|---|
| Predicted aid, lagged | 1.20*** | 1.37*** | .12 |
| | (.22) | (.37) | (.12) |
| Predicted trade, lagged | | -.32 | 1.73*** |
| | | (.66) | (.30) |
| GDP per capita, lagged | -.14 | -.18 | 1.11*** |
| | (.28) | (.28) | (.14) |
| Population | .67 | .61 | .82*** |
| | (.74) | (.74) | (.25) |
| Inflation | -.16 | -.16 | .013 |
| | (.13) | (.13) | (.075) |
| Money | .0048 | .0047 | .011*** |
| | (.0037) | (.0036) | (.0023) |
| Schooling | -.077 | -.094 | .19* |
| | (.18) | (.18) | (.10) |
| Institutional quality | .056*** | .054*** | -.030** |
| | (.020) | (.019) | (.012) |
| Openness | .068 | .046 | .083 |
| | (.11) | (.11) | (.076) |
| Countries | 58 | 58 | 58 |
| $R^2$ | .36 | .37 | .78 |
| Observations | 340 | 339 | 339 |

Note: Column (1) is the first stage of the regression in Table 1 column (3). Columns (2) and (3) are the first stages of the regression in Table 5 column (3). The instruments for aid and trade are built from fitted values of the preliminary stage estimated at the bilateral level, and then aggregated at the country level. All the regressions include country and year effects. Robust standard errors clustered at the country level in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

received larger aid shares because of these characteristics, and not because of their entry dates. However, Frot (2009) also shows that the explanatory power of entry dates is robust to controlling for these characteristics. In order to disentangle these different possible effects, the normalized aid share of each recipient is regressed on a set of controls. The following equation is estimated:

$$\sigma_{ijt} = \alpha + \beta\tau_{ijt} + \gamma\tau_{ijt}^2 + \delta\kappa_{ij} + \boldsymbol{x_{ijt}\varphi} + \varepsilon_{ijt} \tag{6}$$

where $\kappa_{ij}$ is entry date order, $\tau_{ijt}$ is the number of years the partnership has existed ($\tau_{ijt} = t - \eta_{ij} + 1$), $\boldsymbol{x_{ijt}}$ is a vector of controls including recipient GDP per capita, recipient population size, a dummy variable for whether donor and recipient shared a colonial relationship, and the distance in kilometers between $i$ and $j$, and $\varepsilon_{ijt}$ is an error term uncorrelated with the independent variables. The variable $\tau_{ijt}^2$ enters the equation to allow for convergence among countries with different entry dates.[20] Table 3 presents the results. Column (1) shows that entry dates are indeed affected by recipient and recipient-donor characteristics, as expected: donors did prioritize countries with a larger population, lower GDP per capita, geographically closer to them and those with which a colonial relationship had been in place.

The remaining columns indicate that, as suggested by Figure 2, earlier entrants indeed receive larger aid quantities, even after controlling for such recipient and recipient-donor characteristics. Columns (4) and (5) acknowledge the censored nature of aid shares that are bound to lie between 0 an 1, and thus present censored regression estimates. Results are very similar to the OLS estimates.

The effects of entry dates are sizable. Consider two hypothetical aid recipients A and B from the same portfolio. A and B's characteristics are identical, except that A's entry date order is $\kappa = 1$ and B's is $\kappa = 10$ (corresponding roughly to a one-standard deviation difference). The difference in A and B's aid shares in year $\tau = 20$ (20 years after they started receiving aid) is 0.99% using estimates from column (3), and 1.45% from column (5). This is as large as between 25% and 40% of the standard deviation of the shares distribution. To put this number into perspective, we compare it with the GDP differential that would result in such a difference. We ask the question, for B to have the same aid share as A, how much smaller should its per capita GDP be? From the estimates of Table 3, B's income per capita would have to be USD 7071 to 5814 lower than that of A, using columns (3) and (5), respectively. The mean income per capita in the sample is USD 1712, with a standard deviation

[20] The exact functional form of the dependence of the normalized share $\sigma_{ijt}$ on $\kappa_{ij}$ is debatable. Equation (6) assumes that it is linear. Figure 2 suggests something more complex, with a falling effect of entry dates on aid shares (curves get closer when one moves downward). To capture such non-linearities we also estimate equation (6) by adding $\kappa_{ij}^2$ as a regressor.

Table 3: Determinants of aid shares

| | (1) Entry | (2) Aid share | (3) Aid share | (4) Aid share | (5) Aid share |
|---|---|---|---|---|---|
| GDP per capita | .00036*** | | -.00014*** | | -.00024*** |
| | (.000077) | | (.000013) | | (.000018) |
| Population, mil | -.012*** | | .0056*** | | .0059*** |
| | (.0011) | | (.0010) | | (.00090) |
| Colony | -3.91*** | | 2.67*** | | 2.94*** |
| | (.52) | | (.71) | | (.72) |
| Distance | .17*** | | -.060*** | | -.086*** |
| | (.041) | | (.014) | | (.016) |
| Entry | | -.13*** | -.11*** | -.17*** | -.15*** |
| | | (.013) | (.017) | (.014) | (.019) |
| Entry, squared | | .0032*** | .0033*** | .0036*** | .0037*** |
| | | (.00034) | (.00046) | (.00043) | (.00058) |
| Length | | .057*** | .084*** | .056*** | .10*** |
| | | (.0047) | (.0064) | (.0056) | (.0083) |
| Length, squared | | -.00099*** | -.0016*** | -.00079*** | -.0018*** |
| | | (.00011) | (.00015) | (.00012) | (.00018) |
| Constant | 6.98*** | .0018 | .18 | -.30*** | .035 |
| | (.36) | (.11) | (.20) | (.082) | (.18) |
| $\sigma$ | | | | 3.76*** | 3.60*** |
| | | | | (.015) | (.025) |
| Observations | 71918 | 132798 | 71620 | 132798 | 71620 |
| $R^2$ | .057 | .021 | .099 | .008 | .028 |

Note: Standard errors clustered at the partnership level in parentheses. Columns (4) and (5) estimate a censored-normal regression. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

of USD 2043, so this difference is extremely large. This implies that entry dates have a large effect when compared to per capita GDP's. The small percentage difference is also significant in monetary terms, as it represents between USD 14 and 20 million (in 2010 USD). Entry dates, together with partnership length, are therefore good predictors of aid shares, on top of more traditional determinants of aid.

### 5.3.2 Exogeneity of the instrument

Exogeneity of our unique aid instrument cannot be statistically tested in the absence of other valid instruments. In absence of a test, we will provide evidence in support of our claim.

**Reverse causality** Unlike actual aid $A_{it}$, predicted aid $\hat{A}_{it}$ is not influenced by shocks to economic performance in the recipient country, so it is not affected by reverse causality. Notice that, while entry dates themselves are endogenous (consider the example of a crisis year for a specific country, that might cause it to seek assistance and enter new partnerships in that date), we can still claim that the *order* of entry is exogenous to the country's performance in the year of entry, because it depends on other factors such as how long the new partners have already been active as donors and how many other partnerships they have established during the years. A stronger concern is that some unobserved trait of the recipient country that promotes growth also has a direct effect on the starting date and/or the duration of the donor-recipient relationship, *i.e.* the building blocks of our instrument. For example, if donors are reluctant or unable to establish partnerships in countries with despotic rulers or with persistent conflicts, this could delay the entry of those countries into donors' portfolios and at the same time limit growth. This would result in a negative correlation between entry date and growth, biasing upward the coefficient on aid in the main regression. Since entry dates are observed at the partnership level, each recipient country has many entry dates (one for each donor) and only one growth rate for each time period. We therefore average the entry date order at the recipient country level, using aid quantities as weights for donors. The average entry date of a country is earlier if the most important donors, in terms of disbursed aid, started their partnership with this country earlier. In Table 4 we show the correlation between this average entry date order and growth, with different controls and in different time periods (each coefficient in the table comes from a separate regression). In the last row of the first column, we see that, indeed, countries with a later entry date did experience a lower average growth rate over the following years. This simple correlation disappears, though, after controlling for the initial level of GDP and population size, arguably strong determinants of subsequent growth rates. The correlation is almost always zero if we consider the period growth rates separately: with only two exceptions, early entrants did not significantly grow faster than late entrants, with or without consideration for initial conditions. We also control for total aid received. The idea is to check if, although the time of entry in a development cooperation partnership has an effect on growth, this effect goes through aid and only through aid. The direct inclusion of aid quantity in this regression is problematic, given the endogeneity of aid to growth, so we do not put too much weight on this last model. All in all, Table 4 shows that entrants with different entry dates do not on average differ from a GDP growth point of view, and offers thus further suggestive evidence that our instrument is indeed exogenous.

**Exclusion restrictions - Entry order** We also need to ensure that there are no other confounding effects that go from entry date order to growth through other channels than aid, *i.e.* that exclusion restrictions hold. For instance, countries engaged in a long-term aid partnership may also exchange valuable information about innovation or technological progress that have nothing to do with aid, but that reflect the specific nature of the relationship between the two countries. In such a case, we would erroneously attribute to aid the better growth performance observed. A response to this concern is to control for those potential factors correlated with entry dates and affecting growth in 5 and show that aid has an independent effect on top of them. However, it is very difficult to directly control for all of these exchanges, that can take many forms. An important variable very likely to be influenced by partnership characteristics is trade. We would expect that two countries engaged in a very strong aid partnership would also engage in other economic exchanges, and that trade would be a prominent one. If our instrument is just a correlate of trade, then it is likely that the effect we are measuring comes from trade, but not from aid.

To exclude this possibility, we include trade, defined as the sum of exports toward and imports from donor countries, in the previous specifications. We construct a trade instrument using the same strategy as for the aid instrument. Using aid entry dates, we compute a predicted trade quantity for each bilateral trade partnership and obtain a predicted trade quantity by summing these up.[21]

Table 5 shows that controlling for trade only reinforces the results. The effect of aid in the 2SLS fixed-effect regressions has a similar size – slightly larger when directly including predicted aid and trade – and is significant. Our trade instrument is though

[21]This is done because otherwise the simultaneity between trade and growth would once more bias the estimations. A reason for using the aid partnership entry dates to instrument trade flows is that we are especially interested in capturing the part of those flows that correlates with our aid instrument. We attempted the same approach for foreign direct investment flows, but then abandoned this part of the analysis due to serious limitations in the bilateral FDI data.

Table 4: Correlation between entry date order and growth

| | Controls | | |
|---|---|---|---|
| | None | Income and Population | Income, Population, and Aid |
| **Time period** | | | |
| 1963-1967 | 0.002 | 0.000 | 0.008 |
| *p-value* | 0.780 | 0.952 | 0.528 |
| 1968-1972 | -0.004 | -0.005 | 0.010 |
| *p-value* | 0.791 | 0.761 | 0.572 |
| 1973-1977 | -0.015 | -0.014 | -0.026** |
| *p-value* | 0.152 | 0.209 | 0.023 |
| 1978-1982 | -0.011 | -0.012 | -0.008 |
| *p-value* | 0.393 | 0.396 | 0.594 |
| 1983-1987 | -0.015 | -0.017* | -0.017 |
| *p-value* | 0.140 | 0.085 | 0.106 |
| 1988-1992 | -0.003 | -0.002 | -0.010 |
| *p-value* | 0.595 | 0.578 | 0.375 |
| 1993-1997 | -0.004 | -0.004 | 0.000 |
| *p-value* | 0.221 | 0.234 | 0.940 |
| 1998-2002 | -0.002 | -0.003 | -0.005 |
| *p-value* | 0.706 | 0.538 | 0.283 |
| 2003-2007 | -0.005 | -0.001 | -0.002 |
| *p-value* | 0.256 | 0.803 | 0.614 |
| Average | -0.006** | -0.007 | -0.007 |
| *p-value* | 0.033 | 0.131 | 0.139 |

Note: The dependent variable is the growth rate over the indicated time period. The variable of interest is the recipient's average order of entry, where each donor is weighted with the total aid quantity donated to that specific recipient. Each coefficient is from a separate regression of growth on the average entry date order, with the indicated controls. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: OLS and IV regressions, with trade flows

| | (1)<br>OLS | (2)<br>FE OLS | (3)<br>FE 2SLS | (4)<br>FE IV |
|---|---|---|---|---|
| Aid, lagged | .0037* | .0050* | .030*** | |
| | (.0021) | (.0030) | (.012) | |
| Predicted aid, lagged | | | | .040*** |
| | | | | (.011) |
| Trade, lagged | -.0082** | -.0034 | -.0062 | |
| | (.0033) | (.0063) | (.013) | |
| Predicted trade, lagged | | | | -.020 |
| | | | | (.030) |
| GDP per capita, lagged | .0036 | -.049*** | -.044** | -.057*** |
| | (.0037) | (.011) | (.020) | (.0096) |
| Population | .0057** | -.069*** | -.083*** | -.070*** |
| | (.0029) | (.022) | (.026) | (.021) |
| Inflation | -.016*** | -.012* | -.0054 | -.010 |
| | (.0056) | (.0069) | (.0092) | (.0065) |
| Money | 1.1e-05 | .00013 | 3.0e-05 | .00010 |
| | (.00013) | (.00021) | (.00022) | (.00016) |
| Schooling | .0029 | .0015 | .0085 | .0045 |
| | (.0020) | (.0056) | (.0089) | (.0055) |
| Institutional quality | .0017* | .00040 | -.00062 | .0012 |
| | (.00090) | (.00087) | (.0012) | (.00092) |
| Openness | .0095*** | .0051 | .0037 | .0045 |
| | (.0032) | (.0040) | (.0053) | (.0044) |
| Observations | 339 | 339 | 339 | 339 |
| Countries | 58 | 58 | 58 | 58 |
| AP test ($p$-val), aid | | | .0013 | |
| AP test ($p$-val), trade | | | 6.2e-08 | |
| KP $F$ stat | | | 6.38 | |
| $R^2$ | .28 | .34 | .092 | .37 |

Note: AP: Angrist-Pischke. KP: Kleibergen-Paap. The dependent variable is the annualized growth rate in per capita income over five years. All the regressions include year effects. Robust standard errors clustered at the recipient level in parentheses in column (1)-(3). Bootstrapped standard errors clustered at the recipient level in parentheses in column (4), 10,000 repetitions. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

not as strong: the null hypothesis of the Angrist-Pischke test is still not rejected, but the Kleibergen-Paap $F$ statistic is lower than when aid is the only endogenous variable in the model. This is not surprising, since we built the trade instrument based on the aid partnership characteristics, so it might be a weak instrument for trade flows, which does not concern us too much. The two first-stages are presented in columns (2) and (3) of Table 2. They reveal mixed results regarding trade but provide further evidence that our aid instrument is valid.

**Exclusion restrictions - Aid budgets** One last point concerns the aggregation procedure we use as the final step to build our instrument. We might have argued so far that we have a valid instrument for aid *shares*, but since we then multiply the shares with donors' aid budgets in order to obtain aid *quantities*, we might reintroduce endogeneity through them. For example, a negative shock for an important partner country might induce a donor to inflate the total aid budget for that year, even if the specific recipient's instrumented aid share is not affected. That would particularly be the case for emergency aid, that responds to food crises or natural disasters. For this reason, the aid definition we use excludes emergency relief. Alternatively, we can think of other factors affecting at the same time donors' total budgets and growth outcomes in the recipient countries. For example, a boom year for one or more donor countries can lead to larger aid budgets and at the same time larger trade flows; if some of the recipients are also trade partners, which is often the case, we might erroneously attribute to aid the beneficial effects that come in fact from other channels. In previous empirical analyses, there is no consensus on the procyclicality of aid budgets (Pallage and Robe (2001), Minoiu et al. (2010)). In our own data, aid budgets do not respond to a weighted average of recipients' GDP, where the weights are the aid shares of each recipient, as shown in Table 6. Although there is covariation between donors' GDP and aid budgets, we think that year effects do a good job controlling for potential confounding influences related to cyclical properties of the aid budgets. Even if we conceded that aid budgets are infact endogenous, other recent contributions in the litrature have used instruments composed interacting one exogenous and one potentially endogenous component (Dreher and Langlotz, 2015; Nunn and Qian, 2014). They argue that the resulting regression resembles a difference-in-difference set up. Given that the effect of the potentially endogenous components is controlled for, as we do through country and year fixed effects, the interaction can be interpreted as being exogenous (Bun and Harrison, 2018; Nizalova and Murtazashvili, 2016).

To further alleviate concerns about the importance of aid budgets in our predicted aid variable, we explore in Table 7 the respective importance of the shares and budget components in our instrument. We build two similar instruments, in one case keeping constant the donors' budget over time, multiplied with time-varying aid shares, and in the other case multiplying the actual donors' budget with time-invariant average aid shares, in order to suppress the variation in each component in turn. These instruments are then used in our growth-aid specification, detailed in the following section. In Table 7, we only present the coefficient on our instrument in the first stage, alongside some second-stage statistics that inform about instrument strength (the Angrist and Pischke test and the Kleibergen and Paap $F$ statistics). Our focus is indeed not on the second-stage coefficients, which are uninformative with these voluntarily curtailed instruments, but rather on how strong they are. Although neither of them performs well, the relative performance of the first one is much stronger, as assessed by both the coefficient of the first stage (meaning that it is this component which is relevant to predict actual aid) and the diagnostic statistics of the second stage. The $p$-value of the Angrist and Pischke test is very large when only the variation in the aid budgets is considered, showing that the instrument is weak. Similarly, the Kleibergen and Paap statistics is extremely low. With this we want to show that the part of variation that strongly instruments for aid comes indeed from the shares.

Table 6: Impact of donors' and recipients' GDP on official development assistance disbursements and trade volume

| | (1) Donor aid budget | (2) Trade flows | (3) Donor aid budget | (4) Trade flows |
|---|---|---|---|---|
| Donor GDP | 6.23** | .013 | | |
| | (2.52) | (.24) | | |
| Recipient GDP | | | .076 | -.026 |
| | | | (.26) | (.040) |
| Observations | 149 | 149 | 159 | 156 |

Note: The dependent variable is the donors' total aid budget in columns (1) and (3) and the total trade volume in column (2) and (4). GDP, aid and trade are measured in constant 2006 UDS millions. Recipient GDP is a weighted average of the GDP of all the recipient countries in each donor's portfolio, weighted with the recipients aid share. All the variables are differences in logs, so the coefficients express elasticities. All the regressions include year effects and country-specific time trend. Robust standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Relative importance of aid shares and aid budgets in the instrument

| | (1)<br>First stage<br>Aid budget constant | (2)<br>First stage<br>Aid share constant |
|---|---|---|
| Predicted aid, shares | .98*** | |
| | (.34) | |
| Predicted aid, budget | | .26 |
| | | (.39) |
| Countries | 57 | 57 |
| $R^2$ | .31 | .25 |
| Observations | 339 | 339 |
| Second-stage statistics | | |
| AP test ($p$-val) | .0058 | .51 |
| KP $F$ stat | 8.21 | .45 |

Note: AP: Angrist-Pischke. KP: Kleibergen-Paap. The dependent variable is aid quantity. The instruments for aid are built from fitted values of the preliminary stage estimated at the bilateral level, and then aggregated at the country level, but keeping constant either the bilateral share in column (2) or the donors' budget in column (1). All the regressions include the variables listed in Table 1, with country and year effects. Robust standard errors clustered at the country level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# 6 Robustness

## 6.1 GMM estimators

In a working paper version of this work, Frot and Perrotta (2010), we apply two GMM estimators in order to adress the dynamic panel bias, the correlation between lagged income and the error term, due to the strong persistence in income and the individual specific component in the error term. However, diagnostics on these estimations lead us to infer that the GMM approach may not improve on the fixed effects specification. If our external instrument for aid is truly exogenous, hence not correlated with the lagged income, it should not be affected by the presence of the dynamic panel bias. GMM estimates can however be requested to the authors.

## 6.2 Aid as a share of GDP

We depart from the aid effectiveness literature by measuring aid in constant dollars, while past research traditionally used aid as a share of GDP. This departure was done to avoid introducing additional endogeneity in the aid variable. It is indeed peculiar how this literature strives to remove reverse causality from GDP to aid by using instrumental variables only to later re-introduce GDP as a denominator. We prefer to instead use aid quantities.[22] This offers also the advantage that the log-log specification directly estimates the elasticity of GDP per capita with respect to aid. Nevertheless, and despite the fact that instrumentation is likely to be more problematic, we feel that we cannot completely ignore the past convention and, in Table 8, we present results where the aid variable is expressed as GDP share.

In this case the problems caused by the dynamic panel bias are more serious when it comes to the aid coefficient.[23] This measure of aid, with GDP at the denominator, is more correlated with the lagged GDP per capita regressor than the aid level. This makes the use of GMM estimators, that remove the bias induced by the dynamic nature of the specification, more relevant, despite their drawbacks documented in the working paper version of this article (Frot and Perrotta, 2010). Hence we present results using the GMM system estimator when aid as a share of GDP is the independent variable of interest. The 2SLS estimator in column (2) is small and not statistically significant. However the IV estimate in column (3) is large. This leads us to suspect some outliers may considerably bias the 2SLS estimate. In column (4), the 2SLS specification is estimated without the outliers identified by the Hadi procedure. Only 11 observations are removed from the regression sample but these indeed strongly influenced the estimate. It is now identical to the IV estimate and significant.[24]

The aid coefficient in column (5), which corresponds to the GMM estimator, is very similar in size and the Hansen J test fails to reject the validity of the GMM approach, while the instrument count is no too high. The Hansen test of exogeneity of the difference instruments for the level equation also fails to reject their validity in this case, which leads us to be somewhat confident about the GMM estimator.[25]

[22] To our knowledge, only one other paper, Arndt et al. (2010), uses aid quantities, and only in the robustness checks.

[23] The dynamic panel nature of 5 biases the estimation of the coefficient on lagged GDP. Any other coefficient in the equation will also be estimated with a bias to the extent that the variable is correlated with the problematic lagged GDP. Therefore, to the extent that our instrument is exogenous, this would not be a problem for the aid coefficient.

[24] Although the Hadi procedure finds 11 outliers, it is sufficient to remove only two observations corresponding to Gambia to find an aid coefficient equal to 0.24, with a standard error of 0.086.

[25] While the GMM approach in this case seems to be appropriate, we remain cautious about this

Table 8: Aid as a share of GDP

| | (1) FE OLS | (2) FE 2SLS | (3) FE IV | (4) FE 2SLS outliers | (5) GMM |
|---|---|---|---|---|---|
| Aid as share of GDP, lagged | .033 | .080 | | .32*** | .38** |
| | (.036) | (.10) | | (.12) | (.16) |
| Predicted aid, share of GDP, lagged | | | .32 | | |
| | | | (.36) | | |
| GDP per capita, lagged | -.052*** | -.049*** | -.052*** | -.037*** | .033** |
| | (.0090) | (.010) | (.0098) | (.010) | (.013) |
| Population | -.069*** | -.069*** | -.070*** | -.062*** | .013*** |
| | (.021) | (.021) | (.023) | (.024) | (.0042) |
| Inflation | -.014** | -.014** | -.014** | -.014** | -.031*** |
| | (.0065) | (.0065) | (.0064) | (.0069) | (.0095) |
| Money | .00012 | 9.9e-05 | .00012 | 1.1e-05 | -.00018 |
| | (.00020) | (.00019) | (.00017) | (.00017) | (.00021) |
| Schooling | .0010 | .0026 | .00086 | .0099 | -.0011 |
| | (.0057) | (.0061) | (.0058) | (.0070) | (.0096) |
| Institutional quality | .00070 | .00074 | .00078 | .00085 | -.0012 |
| | (.00089) | (.00087) | (.00092) | (.00089) | (.0015) |
| Openness | .0045 | .0039 | .0045 | .0037 | .0090 |
| | (.0038) | (.0039) | (.0044) | (.0049) | (.0070) |
| Instruments | | | | | 20 |
| Countries | 58 | 58 | 58 | 57 | 58 |
| Hansen J test ($p$-val) | | | | | .71 |
| Hansen test ($p$-val), level | | | | | .49 |
| *AR(1)* | | | | | .00096 |
| *AR(2)* | | | | | .64 |
| AP test ($p$-val) | | 6.0e-14 | | 4.7e-8 | |
| KP $F$ stat | | 98.5 | | 40.0 | |
| $R^2$ | .33 | .32 | .33 | .22 | |
| Observations | 338 | 338 | 338 | 327 | 338 |

Note: AP: Angrist-Pischke. KP: Kleibergen-Paap. The dependent variable is the annualized growth rate in GDP per capita over five years. In column (5), the Blundell and Bond (1998) system GMM estimator is used. Twice lagged log GDP is used as instrument in the differences equation; lagged differenced log GDP is used as instrument in the levels equation. All the regressions include year and country fixed effects. Robust standard errors clustered at the country level in parentheses. Bootstrapped standard errors in column (3), 10,000 repetitions. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Our conclusions are therefore mostly robust to the change in aid measurement. When properly instrumented for, aid has a positive and significant effect on GDP. Our 2SLS estimate with this new variable is 0.32. It can be related to our former estimates, using a back-of-the-envelope calculation. If $\gamma_1$ and $\gamma_2$ are the aid coefficients using log aid and aid as a share of GDP, then computing marginal effects[26], $\gamma_1 = \frac{A_{t-1}}{Y_{t-1}}\gamma_2$. The mean of aid to GDP in the regression sample is 0.12, so that the corresponding $\gamma_1$ is 0.038. The actual estimate from Table 1 is 0.04, so the two specifications lead to very similar results.

## 6.3 Definition of growth

Another departure from the literature in our variable definitions is the growth rate. As indicated in the appendix, growth is defined over five-year periods and then annualized. The aid effectiveness literature traditionally measures growth as the average yearly growth rate during the time period, i.e. as $\frac{1}{5}\sum_{i=0}^{4} \frac{y_{\tilde{t}+i+1}-y_{\tilde{t}+i}}{y_{\tilde{t}+i}}$. The two growth rates are highly correlated so we do not expect this change to affect the results. On the other hand, we want to ensure that our results are not driven by this modification, and for greater comparability with the existing literature, we here replicate some of our results with growth defined as the five-year average of yearly rates.

Column (1) of Table 9 is the within groups estimator with aid instrumented. The coefficients on aid in columns (1) and (2) are still significant, and similar in size to the coefficient in Table 1. This confirms that our findings are not driven by our alternative definition of growth. Given the serial correlation that averaging introduces (see Acemoglu et al. (2008)), we suspect that the dynamic bias problem might be accentuated, so in column (3) we report results using the system GMM estimator. The aid coefficient is much reduced, but the coefficient on lagged income is way out of the acceptable range[27], so we consider this last estimation to be flawed.

result, as the working paper version of this article shows that the conditions required for the GMM approach to be valid are more often violated than not in the aid effectiveness context.

[26]With the log of aid levels, the derivative of growth with respect to aid is equal to $\frac{\gamma_1}{A_{t-1}}$. With the aid share specification, the derivative is equal to $\frac{\gamma_2}{Y_{t-1}}$.

[27]Lagged income is positively correlated with the errors, and hence its coefficient is upward biased in the basic OLS setting. When the fixed effect transformation is applied, this positive correlation induces a negative correlation between the lagged income and the error term, which biases downward instead the coefficient. We hence know that the true coefficient lies somewhere in this range; implicitly, we can therefore evaluate the performance of the GMM estimator checking that the income coefficient lies in this acceptable range.

Table 9: Growth as an average

| | (1) FE 2SLS | (2) FE IV | (3) GMM |
|---|---|---|---|
| Aid, lagged | .027*** | | .0066 |
| | (.0064) | | (.011) |
| Predicted aid, lagged | | .033*** | |
| | | (.0076) | |
| GDP per capita, lagged | -.052*** | -.056*** | .010 |
| | (.012) | (.0092) | (.015) |
| Population | -.095*** | -.077*** | .0019 |
| | (.026) | (.021) | (.0075) |
| Inflation | -.0049 | -.0093 | -.020** |
| | (.0086) | (.0062) | (.010) |
| Money | 8.0e-06 | .00014 | -8.5e-05 |
| | (.00022) | (.00016) | (.00022) |
| Schooling | .0048 | .0027 | -.0031 |
| | (.0079) | (.0055) | (.012) |
| Institutional quality | -.00059 | .00094 | .00044 |
| | (.00088) | (.00083) | (.0017) |
| Openness | .0036 | .0054 | .015*** |
| | (.0054) | (.0042) | (.0046) |
| Instruments | | | 22 |
| Countries | 59 | 59 | 59 |
| Hansen J test ($p$-val) | | | .060 |
| Hansen test ($p$-val), level | | | .044 |
| *AR(1)* | | | .0020 |
| *AR(2)* | | | .69 |
| AP test ($p$-val) | 7.5e-07 | | |
| KP $F$ stat | 31.0 | | |
| $R^2$ | .14 | .38 | |
| Observations | 341 | 341 | 341 |

Note: KP: Kleibergen-Paap. AP: Angrist-Pischke. In column (3), twice lagged log GDP is used as instrument in the differences equation; lagged differenced log GDP is used as instrument in the levels equation. Log predicted aid is used as an instrument for aid in all the regressions. All the regressions include year effects. Robust standard errors clustered at the country level in parentheses. Bootstrapped standard errors in columns (2), 10,000 repetitions. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Moreover, the two Hansen tests reject the validity of the GMM instruments. This estimation is an illustration of the risks one encounters when using GMM estimators in this context.

# 7 Conclusions

In this article, we proposed a new instrument for identifying the causal effect of aid on growth. This instrument takes the supply side approach, that relates to the aid allocation decision, a step further, using a source of variation that is not just external but exogenous to growth. As far as possible, the instrument is shown to be valid and strong. We claim that this is an improvement from a stream of papers that relied on weak and non-exogenous instruments.

When it comes to the estimation strategy and the choice of estimator, we make simple and clear methodological choices, explain and motivate them step by step and probe their validity as best we can.

The effects uncovered by our identification strategy indicate an elasticity of GDP with respect to aid of about 0.04. These effects imply that the average developing country would have been about 30% poorer had aid not been disbursed between 1963 and 2007. The average developing country citizen would have been 15% poorer. The median citizen would have been 7% poorer.

We voluntarily shy away from the debate about how the aid system can be improved, or, as some would argue, whether any improvement is feasible at all. Our goal is more modest. We want to answer the question: given available data, what can be said about aid effectiveness in the past 40 years, using appropriate econometric techniques that take instrumentation seriously? Our answer is clear, and robust across specifications: the effect of aid has been, on average, positive. Whether an average 30% GDP per capita reduction (or a 7% reduction if one thinks the median citizen is a more appropriate measurement unit) justifies the disbursed aid quantities is out of the scope of this article. It is nevertheless our hope that this result will usefully inform the current debate about reforming the aid system.

# 8 Funding

This work was supported by the Handelsbankens Forskningsstiftelser, [P2009-0262:1], to E. F.

## Acknowledgments

The authors wish to thank Ethan Kaplan, Jakob Svensson, Rajeev Dehejia, Philippe Aghion, Pamela Campa, Martin Berlin, Steven Davis, Stephen Knack, Carl-Johan Dalgaard, David Weil, participants at the IIES and Economics Department seminars at Stockholm University and three anonymous referees. Editorial assistance by Christina Lönnblad is gratefully acknowledged.

# References


Acemoglu, D., S. Johnson, and J. A. Robinson (2008). Income and democracy. *The American Economic Review 98*(3), 808–42.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Arndt, C., S. Jones, and F. Tarp (2010). Aid, Growth, and Development: Have We Come Full Circle? *Journal of Globalization and Development 1*(2), 1–27.

Arndt, C., S. Jones, and F. Tarp (2014). Assessing foreign aidâĂŹs long run contribution to growth and development. *World Development*.

Barro, R. J. and J.-W. Lee (2010). A new data set of educational attainment in the world, 1950-2010. Working Paper 15902, NBER.

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

Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics 87*(1), 115–143.

Brückner, M. (2013). On the simultaneity problem in the aid and growth debate. *Journal of Applied Econometrics 28*(1), 126–150.

Bun, M. J. and T. D. Harrison (2018). Ols and iv estimation of regression models including endogenous interaction terms. *Econometric Reviews*, 1–14.

Burnside, C. and D. Dollar (2000). Aid, policies, and growth. *The American Economic Review 90*(4), 847–868.

Clemens, M. A., S. Radelet, R. R. Bhavnani, and S. Bazzi (2012). Counting chickens when they hatch: Timing and the effects of aid on growth*. *The Economic Journal 122*(561), 590–617.

Dalgaard, C.-J., H. Hansen, and F. Tarp (2004). On the empirics of foreign aid and growth. *Economic Journal 114*(496), 191 – 216.

Deaton, A. (2013). *The great escape: Health, wealth, and the origins of inequality.* Princeton University Press.

Doucouliagos, H. and M. Paldam (2009). The aid effectiveness literature: The sad results of 40 years of research. *Journal of Economic Surveys 23*(3), 433–461.

Doucouliagos, H. and M. Paldam (2014). Finally a breakthrough? the recent rise in the size of the estimates of aid effectiveness. Technical report.

Dreher, A. and S. Langlotz (2015). Aid and growth. new evidence using an excludable instrument. Technical report, CEPR Discussion Papers.

Easterly, W., R. Levine, and D. Roodman (2004). Aid, policies, and growth: Comment. *The American Economic Review 94*(3), 774–780.

Frot, E. (2009). Early vs. late in aid partnerships and implications for tackling aid fragmentation. Working Paper No.1, SITE.

Frot, E. and M. Perrotta (2010). Aid Effectiveness: New Instrument, New Results? Working Paper No. 11, SITE.

Hadi, A. S. (1992). Identifying multiple outliers in multivariate data. *Journal of the Royal Statistical Society 54*(2), 761 – 771.

Hansen, H. and F. Tarp (2001). Aid and growth regressions. *Journal of Development Economics 64*(2), 547 – 570.

Kleibergen, F. and R. Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*(1), 97–126.

Lof, M., T. J. Mekasha, and F. Tarp (2014). Aid and income: Another time-series perspective. *World Development*.

Manski, C. F. and E. Tamer (2002). Inference on regressions with interval data on a regressor or outcome. *Econometrica 70*(2), 519–546.

Minoiu, C., L. Zanna, and E. Dabla-Norris (2010). Business Cycle Fluctuations, Large Shocks, and Development Aid: New Evidence. *IMF Working Papers*.

Nizalova, O. Y. and I. Murtazashvili (2016). Exogenous treatment and endogenous factors: Vanishing of omitted variable bias on the interaction term. *Journal of Econometric Methods 5*(1), 71–77.

Nowak-Lehmann, F., A. Dreher, D. Herzer, S. Klasen, and I. Martínez-Zarzoso (2012). Does foreign aid really raise per capita income? a time series perspective. *Canadian Journal of Economics/Revue canadienne d'économique 45*(1), 288–313.

Nunn, N. and N. Qian (2014). Us food aid and civil conflict. *American Economic Review 104*(6), 1630–66.

Pallage, S. and M. Robe (2001). Foreign aid and the business cycle. *Review of International Economics 9*(4), 641–672.

Qian, N. (2014). Making progress on foreign aid. Technical report, National Bureau of Economic Research.

Rajan, R. G. and A. Subramanian (2008). Aid and growth: What does the cross-country evidence really show? *Review of Economics and Statistics 90*(4), 643–665.

Ravallion, M. (2014). On the role of aid in the great escape.

Rigobon, R. and T. Stoker (2009). Bias from censored regressors. *Journal of Business and Economic Statistics 27*(3), 340–353.

Roeder, P. G. (2001). Ethnolinguistic fractionalization (ELF) indices, 1961 and 1985. *Department of Political Science, University of California at San Diego*.

Roodman, D. (2007). The anarchy of numbers: aid, development, and cross-country empirics. *The World Bank Economic Review 21*(2), 255–277.

Sachs, J. D., A. Warner, A. Åslund, and S. Fischer (1995). Economic Reform and the Process of Global Integration. *Brookings papers on economic activity 1*(1995), 1–95.

Wacziarg, R. and K. H. Welch (2008). Trade Liberalization and Growth: New Evidence. *The World Bank Economic Review 22*(2), 187–231.

# Appendix

## No-aid counterfactual

This section details how the actual GDP per capita is compared to its no-aid counterfactual. Consider equation (5), but written with the aid share as a separate independent variable:

$$\Delta lny_{it} = \beta_1 lny_{it-5} + \beta_2 \frac{a_{it-5}}{y_{it-5}} + \boldsymbol{x_{it}\gamma} + \epsilon_{it} \tag{7}$$

where regressors other than lagged GDP per capita and aid form the matrix $\boldsymbol{x_{it}}$. Equation (7) is equivalent to

$$lny_{it} = (1+\beta_1)\, lny_{it-5} + \beta_2 \frac{a_{it-5}}{y_{it-5}} + \boldsymbol{x_{it}\gamma} + \epsilon_{it} \tag{8}$$

Assume that aid is first disbursed in period $t=0$, while in the counterfactual scenario $\frac{a_{i0}}{y_{i0}} = 0$, so that the counterfactual GDP per capita $\tilde{y}_{i5}$ is

$$ln\tilde{y}_{i5} = lny_{i5} - \beta_2 \frac{a_{i0}}{y_{i0}} \tag{9}$$

In the next period (after 5 years) and from (7) and (9),

$$\begin{aligned} ln\tilde{y}_{i10} &= (1+\beta_1)\, ln\tilde{y}_{i5} + \boldsymbol{x_{i10}\gamma} + \epsilon_{i10} && (10)\\ &= lny_{i10} - \beta_2 \frac{a_{i5}}{y_{i5}} - (1+\beta_1)\,\beta_2 \frac{a_{i0}}{y_{i0}} && (11)\end{aligned}$$

Because GDP per capita in period $t$ depends on period its lag, which is itself modified by the change in aid quantities, the total impact of aid in 2007 depends on aid shares in all previous periods. More specifically, simple iteration of (11) yields, for any $t > s$

$$ln\tilde{y}_{it} = lny_{it} - \beta_2 \sum_{k=1}^{(t-t0)/5} (1+\beta_1)^k \frac{a_{it-5k}}{y_{it-5k}} \tag{12}$$

Table **??** presents ratios of counterfactual to actual GDP per capita, which are

$$\frac{\tilde{y}_{it}}{y_{it}} = exp\left(-\beta_2 \sum_{k=1}^{(t-t0)/5} (1+\beta_1)^k \frac{a_{it-5k}}{y_{it-5k}}\right) \tag{13}$$

This result uses the five-year growth as the dependent variable in equation (7), whereas in our estimations the dependent variable is the *annualized* five-year growth rate. In order to use (13) instead of its non-convenient equivalent with annualized growth rates, we ran the same regression as in Table 8, column (4), but with the five-year growth as the dependent variable. Because growth rates are, on average, quite small in our sample, this is almost equivalent to, using a first-order approximation, multiplying our dependent variable by 5. Indeed, the coefficient on the aid variable with this new specification is found to be 1.59, with a standard error of 0.64, while it is 0.32 in Table 8, with a standard error of 0.12.

## Normalized aid shares

Donor portfolio expansion implies that aid shares are bound to fall on average. Donors disbursed aid to a handful of countries in the 1960s but nowadays often have more than 100 countries in their recipient portfolio. In order to make aid shares neutral with respect to portfolio size, we define normalized aid shares $\sigma_{ijt}$:

$$\sigma_{ijt} = s_{ijt} - \frac{1}{N_{jt}} \tag{14}$$

where $N_{jt}$ is the number of recipients that have received aid from donor $j$ at least once before year $t$. Normalized shares are hence deviations from an equal sharing rule among all recipients. These normalized shares show some interesting behavior over time.

## Outliers

Easterly et al. (2004) showed how aid effectiveness results could be sensitive to the exclusion of a few outliers. In Table A.1, we make use of the Hadi (1992) procedure to exclude outliers from the sample. The procedure fails to find any outliers when trade is included as a control, and the results (not reported) are identical to those in Table 5. When trade is not included, only three outliers are detected, and excluding them hardly affects the estimates of the aid coefficient.

## Sample size

Our main specifications include control variables that are commonly found in the aid effectiveness literature. However, limited data availability for some of them sharply reduces the sample size. Our dataset contains 125 countries but the regressions rely on 58 countries at most. Larger sample size comes at the cost of omitting some growth determinants and hence, potentially biases the aid coefficient. On the other hand, the aid instrument, if truly exogenous, should remove the correlation between aid and the error term even in the presence of omitted variables. This provides an indirect test of instrument validity, in addition to extending the estimation to many more countries. The most parsimonious specification with only lagged GDP as control allows us to use data on 106 countries, a dramatic increase. The effect of aid is smaller in this augmented sample but it is still significant when aid is instrumented, see Table A.2. However, the coefficient becomes insignificant when predicted aid is directly included as a regressor, in column (4).

Table A.1: Excluding outliers

| | (1)<br>FE 2SLS | (2)<br>FE IV |
|---|---|---|
| Aid, lagged | .030** | |
| | (.012) | |
| Predicted aid, lagged | | .038*** |
| | | (.011) |
| GDP per capita, lagged | -.052*** | -.056*** |
| | (.012) | (.0095) |
| Population | -.090*** | -.068*** |
| | (.028) | (.021) |
| Inflation | -.0059 | -.010 |
| | (.0094) | (.0065) |
| Money | -3.2e-05 | .00011 |
| | (.00022) | (.00016) |
| Schooling | .0071 | .0048 |
| | (.0082) | (.0056) |
| Institutional quality | -.00050 | .0012 |
| | (.00099) | (.00092) |
| Openness | .0026 | .0049 |
| | (.0056) | (.0042) |
| Countries | 57 | 57 |
| AP $p$-value | .0031 | |
| KP $F$ stat | 9.57 | |
| Observations | 337 | 337 |

Note: AP: Angrist-Pischke. KP: Kleibergen-Paap. The dependent variable is the annualized growth rate in GDP per capita over five years. Log predicted aid used as instrument in all the regressions. All the regressions include year effects. Robust standard errors clustered at the country level in parentheses in columns (1) and (2). Bootstrapped standard errors in column (3) and (4), 10,000 repetitions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: OLS and IV regressions, large sample

| | (1) OLS | (2) FE | (3) FE 2SLS | (4) FE IV |
|---|---|---|---|---|
| Aid, lagged | .00042 | .0024 | .016*** | |
| | (.0012) | (.0023) | (.0049) | |
| Predicted aid, lagged | | | | .0032 |
| | | | | (.0046) |
| GDP per capita, lagged | .0022 | -.039*** | -.039*** | -.039*** |
| | (.0014) | (.0065) | (.0057) | (.0057) |
| Countries | 106 | 106 | 106 | 106 |
| AP test ($p$-val) | | | 5.0e-7 | |
| KP $F$ stat | | | 28.8 | |
| $R^2$ | .072 | .22 | .15 | .22 |
| Observations | 628 | 628 | 628 | 628 |

Note: AP: Angrist-Pischke. KP: Kleibergen-Paap. The dependent variable is the annualized growth rate in GDP per capita over five years. All the regressions include year effects. Robust standard errors clustered at the country level in column (1)-(3). Bootstrapped standard errors in column (4), 10,000 repetitions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Because we include as few controls as possible in these regressions, there may be strong outliers in these specifications. We put this result to the test of excluding outliers, once more following the Hadi procedure. Table A.3 shows that outliers do not strongly bias coefficients. However the standard error of the 2SLS estimate is larger, which makes its statistical significance fall. On the other hand, the IV estimate is now similar to the 2SLS aid coefficient, and is now significantly different from zero. Excluding outliers therefore confirms that the aid coefficient in the whole sample lies around 0.15, about half its size when controls are added to the regression, and that it remains significant.

## Data appendix

*Time periods.* Observations for all variables except GDP, aid and trade are five-year arithmetic averages. Time period 1 represents years 1963-1967. The last period (period 9) is 2003-2007.

*Log aid.* Aid is Official Development Assistance (ODA) and comes from the Donor Assistance Committee (DAC) database of the OECD, Table 2a. Because predicted aid is built from predicted aid shares, net ODA, which is the usual aid variable in the

Table A.3: OLS and IV regressions excluding outliers, large sample

| | (1) OLS | (2) FE | (3) FE 2SLS | (4) FE IV |
|---|---|---|---|---|
| Aid, lagged | .0015 | .0032 | .014 | |
| | (.0013) | (.0031) | (.0096) | |
| Predicted aid, lagged | | | | .013** |
| | | | | (.0061) |
| GDP per capita, lagged | .0020 | -.038*** | -.038*** | -.039*** |
| | (.0013) | (.0057) | (.0051) | (.0055) |
| Countries | 106 | 106 | 101 | 101 |
| AP test ($p$-val) | | | .0029 | |
| KP $F$ stat | | | 9.29 | |
| $R^2$ | .075 | .24 | .16 | .25 |
| Observations | 621 | 619 | 608 | 608 |

Note: AP: Angrist-Pischke. KP: Kleibergen-Paap. The dependent variable is the annualized growth rate in GDP per capita over five years. All the regressions include year effects. Robust standard errors clustered at the country level in column (1)-(3). Bootstrapped standard errors in column (4), 10,000 repetitions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

aid effectiveness literature and which is potentially negative, cannot be used. Aid is defined as gross ODA, minus gross debt relief. The latter is excluded because it artificially inflates aid numbers in very recent years, where large debt cancellations were granted. Aid is not averaged, but summed up over the time period. It is expressed in millions of 2006 USD. Aid from all donors whose activity is reported by DAC and to all developing countries, according to DAC definition, is considered.

*Log trade.* Trade at the bilateral level is defined as the sum of imports and exports. At the recipient country level, it is summed across donor countries. Data in current USD millions from the International Trade dataset, version 2.01, of the Correlates of War Project. It is converted in 2006 USD by deflating it with the Consumer Price Index of the US Bureau of Labor Statistics.

*Aid and trade as shares of GDP.* Data in current USD is divided by GDP in current USD.

*Log GDP.* GDP in 2000 USD is from the World Development Indicators. GDP is not averaged, but measured every fifth year. We do so instead of averaging to avoid introducing serial correlation.

*GDP per capita.* In 2000 USD. Source: World Development Indicators.

*Growth.* Growth is defined as the difference $d = ln(y_{\tilde{t}}) - ln(y_{\tilde{t}-5})$, where $y_{\tilde{t}}$ is

GDP per capita in year $\tilde{t}$. Note here that $\tilde{t}$ indexes year and not time periods. The annual equivalent to growth rate $d$ is then computed as $g = (1+d)^{\frac{1}{5}} - 1$.

*Log population.* Population is measured in millions. Source: World Development Indicators.

*Inflation.* Natural logarithm of 1+consumer price inflation rate. Source: World Development Indicators.

*Money.* Ratio of M2 to GDP. Source: World Development Indicators.

*Schooling.* Average years of primary schooling attained. Source: Barro and Lee (2010).

*Institutional Quality.* Variable between 0 and 16, defined as the sum of "Corruption", "Law and Order", and "Bureaucracy Quality", from the International Country Risk Guide (ICRG) of the PRS Group. Data is not available before 1984. For earlier years, data from the first available year is used. By doing so we follow the practice in the literature (see Roodman (2007)).

*Openness.* Index constructed by Sachs et al. (1995) and Wacziarg and Welch (2008).

*Ethnic fractionalization.* Ethnolinguistic Fractionalization index. Source: Roeder (2001).

*Regional dummies.* Dummies for East Asia and Pacific, and Sub-Saharan Africa. Region definitions are from the World Development Indicators.

*Colony.* Dummy variable equal to 1 if the pair has ever had a colonial link. Source: CEPII.

*Distance.* Distance in thousands of kilometers between the two main cities of the two countries. Source: CEPII.

Table A.4: Summary statistics and source of variables

| Variable | Mean | Std. dev. | Unit |
|---|---|---|---|
| *Aid efficiency variables* | | | |
| Growth | .010 | .033 | Annualized five-year growth |
| Aid | 530.7 | 726.0 | Five-year disbursements, constant 2006 USD mil. |
| Aid, GDP share | .079 | .092 | Five-year share |
| GDP per capita | 1659.7 | 2003.2 | Constant 2000 USD |
| Population | 27.7 | 91.6 | Millions |
| Inflation | .15 | .28 | percentage change |
| Openness | .27 | .43 | 0-1 index |
| Money | 34.0 | 23.7 | M2 as percentage of GDP |
| Trade | 11.7 | 28.1 | Five-year quantity, constant 2006 USD bn |
| Schooling | 1.35 | 0.89 | Years |
| Institutional quality | 6.84 | 2.50 | 1-16 continuous variable |
| East Asia | .11 | .32 | Identifier |
| Sub-Saharan Africa | .40 | .49 | Identifier |
| Ethno-linguistic frac. | .55 | .27 | Index (0 to 1) |
| *Predicted aid variables* | | | |
| Aid share | 1.18 | 3.86 | percentage |
| Entry | 7.42 | 7.42 | Years |
| Length | 17.6 | 11.7 | Years |
| Colony | .038 | .19 | Index (0 to 1) |
| Distance | 8.41 | 3.84 | Thousands of km |

Note: The summary statistics are based on the large sample, except for the predicted aid variables where it is based on all the available donor-recipient pairs.

Table A.5: List of countries

| Sample with controls | | | Large sample | |
|---|---|---|---|---|
| Algeria | Malawi | | Angola | Micronesia, Federated States |
| Argentina | Malaysia | | Barbados | Namibia |
| Bangladesh | Mali | | Belize | Nepal |
| Bolivia | Mexico | | Benin | Nigeria |
| Botswana | Morocco | | Bhutan | Oman |
| Brazil | Mozambique | | Burkina Faso | Palestinian Adm. Areas |
| Cameroon | Nicaragua | | Burundi | Rwanda |
| Colombia | Niger | | Cambodia | Samoa |
| Congo, Republic | Pakistan | | Cape Verde | Saudi Arabia |
| Costa Rica | Panama | | Central African Republic | Sierra Leone |
| Côte d'Ivoire | Papua New Guinea | | Chad | Solomon Islands |
| Dominican Republic | Paraguay | | Chile | St. Lucia |
| Ecuador | Peru | | Comoros | St.Vincent & Grenadines |
| Egypt | Philippines | | Congo, Democratic Republic | Sudan |
| El Salvador | Senegal | | Djibouti | Suriname |
| Gabon | Sri Lanka | | Equatorial Guinea | Swaziland |
| Gambia | Syria | | Ethiopia | Tonga |
| Ghana | Tanzania | | Fiji | Vanuatu |
| Guatemala | Thailand | | Grenada | Vietnam |
| Guyana | Togo | | Guinea | |
| Haiti | Trinidad and Tobago | | Guinea-Bissau | |
| Honduras | Tunisia | | Laos | |
| India | Turkey | | Lebanon | |
| Indonesia | Uganda | | Lesotho | |
| Iran | Uruguay | | Liberia | |
| Jamaica | Venezuela | | Madagascar | |
| Jordan | Yemen | | Maldives | |
| Kazakhstan | Zambia | | Mauritania | |
| Kenya | Zimbabwe | | Mauritius | |

Note: The large sample corresponds to the regressions where the only controls are lagged log GDP, lagged log aid, and log population. Note that in addition to including more countries, the "large" sample also includes more observations for some countries than the sample with controls.