

Building or Bypassing Recipient Country Systems

Are Donors Defying the Paris Declaration?

Stephen Knack

The World Bank
Development Research Group
Human Development and Public Services Team
April 2013



Abstract

The 2005 Paris Declaration on Aid Effectiveness sets targets for increased use by donors of recipient country systems for managing aid. It also calls for donors to be more responsive to the quality of recipient country systems: the optimal level of their use, in terms of maximizing the development effectiveness of aid, is believed to vary with their quality. This study investigates the degree to which donors' use of country systems is in fact positively related to their quality, using indicators explicitly endorsed for this purpose by the Paris Declaration and covering the 2005–2010 period. The results of these tests strongly confirm a positive and significant relationship, and show it is robust to corrections for potential sample selection, omitted variables, or endogeneity bias. The result holds even when

estimates are informed only by variation over time within each donor-recipient pair in use and quality of country systems. Moreover, donor-specific tests show that use of country systems varies positively with their quality for the vast majority of donors. These findings contradict several other studies that claim there is no relation and imply that donors in this respect are failing to live up to their commitments under the Paris Declaration. The author's interpretation of the available evidence on use of country systems is more favorable: donors' behavior over the measurement period is largely consistent with their commitments in this area. In this respect, at least, donors appear to have modified their aid practices in ways that build rather than undermine administrative capacity and accountability in recipient country governments.

This paper is a product of the Human Development and Public Services Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The author may be contacted at sknack@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Building or Bypassing Recipient Country Systems: Are Donors Defying the Paris Declaration?

Stephen Knack
Lead Economist, DECHD

JEL: F35, H61, O10, O19

Keywords: Aid effectiveness, Paris Declaration, corruption, use of country systems, public financial management

This research was supported by the World Bank's Knowledge for Change Program (KCP). The conclusions of this paper are not intended to represent the views of the World Bank, its Executive Directors, or the countries they represent. Valuable comments, suggestions or other assistance were provided by Nicholas Eubank, Daniel Lederman, Adam Wagstaff and Claudia Berg. The author assumes full responsibility for any remaining errors.

1. Introduction

In response to concerns regarding the quality of development aid, the Paris Declaration on Aid Effectiveness (in 2005) moved implementation issues to the top of the international aid effectiveness agenda. Donors committed to making their aid practices more consistent with the principal of country “ownership,” including conformity with partner countries’ “national development strategies, institutions and procedures.”

The importance of delivering aid in ways consistent with long-run institutional-strengthening is a major theme in the Paris Declaration (PD) agenda. It calls for increased use of recipient systems in managing aid, but it explicitly acknowledges that weak country systems make aid less effective. Recipients, with technical assistance from donors, commit to strengthening their public financial management (PFM) systems and formulating a credible national development strategy where one does not exist. In the meantime, using those systems, despite their flaws, is believed to strengthen them: “Donors can help build capacity and trust by using country systems to the fullest extent possible, while accepting and managing the risks involved...” (OECD, 2009a: 27). Ellmers (2011: 18) further notes: “Improved country systems, in their turn, improve the transparency and accountability of all public spending beyond aid.”

Donors’ incentives to use country systems (or alternatively to micro-manage aid using their own parallel systems) depend in part on their perceived quality. Where recipient aid management systems are stronger, corruption scandals tarnishing the donor agency’s reputation are less likely to occur, and aid-funded programs are more likely to be selected and implemented more efficiently. Although the Paris Declaration commits donors to use country systems and procedures “to the maximum extent possible,” it also recognizes that weaknesses in country systems sometimes justify donors’ decisions to bypass them. Aid effectiveness is furthered by

using country systems “when these provide assurances that aid will be used for agreed purposes.” Despite the lack of a consensus regarding the developmentally-optimal level of donors’ use of country systems overall, there is broad agreement that the optimal level varies positively with the quality of those systems. Jansen (2009: 23) provides an illustration of how weaknesses in a country’s financial management system can lead to rampant corruption and failure to achieve development objectives.

The existing evidence base is insufficient to identify the developmentally optimal level of use of country systems, conditional on varying degrees of the quality of those systems. However, the Paris Declaration (and its 2008 follow-up “Accra Agenda for Action”) reflects a consensus belief that overall, donors’ use of country systems and their sensitivity to the quality of systems are both sub-optimal. In other words, development effectiveness would be enhanced if the mean use of country systems, and the elasticity of their use with respect to their quality, were both increased.

The three rounds of the Paris Declaration Monitoring Survey (PDMS) were designed to provide quantitative evidence of progress toward these goals and others set by the PD. Data were collected from donors and recipients pertaining to the years 2005 (as a baseline), 2007 (for measuring interim progress), and 2010 (the target year).

This study assesses donor behavior in terms of these premises underlying the PD agenda. Other studies report a weak relationship between quality of systems and donors’ use of them, casting doubt on whether donors are living up to one of their commitments under the PD. In contrast, this study finds a strong, significant and positive relationship between quality of PFM systems and donors’ reliance on them. Moreover, donors’ use of PFM systems overall increased over the 2005-2010 period far more than can be explained solely by improvements in the quality

of PFM systems in recipient countries. In donor-specific tests, a positive time trend in use of country PFM systems is observed for most donors, controlling for quality of systems and other factors. Although donors' aid practices doubtless are driven primarily by other considerations (e.g. domestic politics, diplomatic or commercial objectives, and beliefs on how global poverty is best reduced), the "peer pressure" components of the PD campaign may have succeeded in getting some donors to tolerate somewhat greater short-term risks in their aid management practices. However, if this increased risk tolerance is a one-off change, or proves to be temporary, future increases in donors' use of country PFM systems are likely to depend on gradual and sustained improvements in their quality.

Taken together, the findings from this study imply that donors' behavior over the measurement period is largely consistent with their commitments in this area under the PD. In this respect, at least, donors appear to have modified their aid practices in ways that build rather than undermine administrative capacity and accountability in recipient country governments.

The remainder of the study is organized as follows. Section 2 briefly reviews the existing evidence on this question. Section 3 describes the data and methods used in the analysis. Detailed results are presented for the full sample in section 4. A variety of robustness tests in that section address potential concerns regarding sample selection, omitted variables and endogeneity bias. Donor-specific tests are presented in section 5, and the final section summarizes and concludes.

2. Related Literature

The OECD-DAC has issued several studies (OECD, 2011a, 2009a, 2009b) reporting progress with respect to the Paris Declaration's quantitative targets. The final report (OECD,

2011a: 50) echoes the earlier ones in concluding that “the relationship between the quality of a country’s PFM systems and donors’ use of them is at best weak.” This conclusion is based on simple bivariate analyses of data on use of PFM systems from the PDMS, and the indicator of PFM quality specified in the Paris Declaration (from the World Bank’s Country Policy and Institutional Assessments, or CPIA).

Based on similar methods but using somewhat different data, IEG (2011) and Ellmers (2011) arrive at a similar conclusion. According to IEG (2011), there is no association between the World Bank’s use of PFM systems and their quality in low-income countries, as measured by countries’ scores from PEFA (Public Expenditure and Financial Accountability) assessments.¹ The analysis was limited to only 17 countries, however, because it used PEFA to measure quality of PFM systems instead of CPIA, which is available for many more countries. Ellmers (2011) concludes there is no correlation between donors’ use of country procurement systems (as measured by the 2007 PDMS) and quality of PFM systems as measured by the CPIA. However, that study is limited to only 9 countries in which Eurodad has also sponsored case studies of procurement systems. Ellmers (2011), IEG (2011) and OECD (2011a) all present simple bivariate correlations, each based on a single year of data.

Using data from all three rounds of the PDMS, Knack (2013) showed that a given donor’s use of country PFM systems was positively related to (1) its share of the total aid provided to a recipient, (2) public opinion in support of aid provision in the donor country, and (3) the trustworthiness of recipient country systems. In measuring the latter, however, Knack (2013) did not employ the CPIA indicator of PFM quality specified in the Paris Declaration. Rather, trustworthiness of country systems was measured primarily using the Control of Corruption indicator from the Worldwide Governance Indicators (WGI) project.

¹ See www.pefa.org for more information on PEFA assessments and indicators.

The OECD (2011a) provides a mixed picture on progress toward greater donor use of country systems, apart from the issue of responsiveness to system quality. For the 32 countries included in both the 2005 and 2010 surveys, it reports that average use of country PFM systems increased from 40% to 48%, and increased from 40% to 44% for country procurement systems. These improvements may appear impressive for the short 5-year time span covered, but as the report notes, the 48% figure fell far short of the 55% target value. This “failure” should be interpreted in its broader context. The PD targets – as with the Millennium Development Goals – were designed to be ambitious, to motivate action more effectively. In these same 32 countries, for example, the number of parallel project implementation units (PIUs) declined by nearly a third (from 1,696 to 1,156) between 2005 and 2010, but the target was for an even more dramatic decline of two-thirds (to only 565).

This study follows the methodological approach of Knack (2013), namely by applying panel data methods to the three-round PDMS dataset, controlling for other factors that influence use of country systems, and correcting for possible selection and endogeneity bias. However, it differs from Knack (2013) in several important respects. First, it directly tests the importance of PFM systems quality for their use by donors, measuring quality with the CPIA indicator specified in the PD. The CPIA indicator is more appropriate for a second reason: it is designed specifically to assess public sector systems for managing public funds including aid. Other well-known “governance” indicators are designed for commercial (e.g. assessing risks to foreign investors) or advocacy (democracy, human rights, transparency) purposes. Second, it addresses the related issue of quality and use of recipient government procurement systems. Finally, this study goes beyond Knack (2013) in reporting not only results for the full three-dimensional (donor/recipient/year) panel, but also donor-specific results. Findings support the hypothesis that

use of country systems by donors is sensitive to their quality, not only for all donors collectively, but also for the vast majority of individual donors. This conclusion does not apply to a few large donors, however, namely the USA, France and the UN system.

3. Data and Empirical Strategy

Measuring use of country PFM and procurement systems

We measure use of country PFM systems with data from the OECD DAC's Paris Declaration Monitoring Survey (PDMS). The PDMS was designed to assist in measuring progress toward the Paris Declaration's objectives between 2005 and 2010. We measure use of country PFM systems with PD Indicator 5a. This indicator is constructed as a simple average of three sub-indicators from the survey: use of national (i) budget execution procedures, (ii) financial reporting procedures, and (iii) auditing procedures, each as a percentage of aid for the government sector.² Detailed criteria for these three dimensions of use-of-PFM-systems are provided in Appendix 1. Correlations among them average .73 (ranging from .54 to .77). Findings presented below change very little if any one of its three components is analyzed instead of Indicator 5a. Mean use of PFM systems among the 2,089 observations appearing in our baseline multivariate regression is 33.5%.³

Use of country procurement systems is measured using PD Indicator 5b, defined as the share of aid to the government sector that is not subject to additional or special requirements by

² Aid for the government sector accounts for about 83% of total aid disbursements reported in the PDMS. It excludes aid disbursed to NGOs, parastatals or private companies unless it is provided "in the context of an agreement with officials authorized to act on behalf of central government" (OECD, 2008). Empirical results reported below are unaffected if government sector aid is replaced with total aid in the denominator of the use-of-PFM-systems indicator.

³ Appendix Table A1 provides summary statistics for all variables in the analysis.

donors for procurement of works, goods and services. Mean use of procurement systems in the main sample of observations is 37.7%.

The PDMS data form an unbalanced three-dimensional panel: surveys were conducted in 2006, 2008 and 2011, and were designed to obtain information on aid flows for the previous calendar year (2005, 2007 and 2010, respectively) for each donor-recipient pair. In each successive survey round, the comprehensiveness and quality of data were improved. The number of recipient countries choosing to participate in the survey increased from 34 in 2006, to 54 in 2008, and 78 in 2011. These 78 recipients account for about three-fourths of total aid worldwide (OECD, 2011a). Guidance on definitions was strengthened and comparability of data reported by donors and by recipients was enhanced in later rounds (OECD, 2011a; OECD 2008).

Coverage of aid within aid-recipient countries is comprehensive with respect to donors. Most of the “emerging” bilateral donors and a few of the smaller multilateral donors and vertical funds are dropped from the analysis, due to missing data on other donor variables from the DAC database. Dropped donors collectively account for a total of only 1.4% of aid reported in the survey. The 34 donors remaining include the 22 traditional DAC donors, Korea (a new DAC member in 2011), Turkey, 7 multilateral organizations, and 3 vertical funds.⁴

The panel is unbalanced in multiple respects. Not all recipients are included in all three years. In addition, in each survey year where aid flows for a donor-recipient pair are 0, the dependent variable is undefined so the observation is treated as missing. These features of the data create potential selection problems, considered in more detail below.

Independent variables

⁴ Multilaterals in the sample include the African, Asian and Inter-American Development Banks, the World Bank, IMF, EU and UN. Vertical funds include the GAVI Alliance, IFAD, and the Global Fund to Fight AIDS, Tuberculosis, and Malaria. The sample includes all donors with sufficient data for listing in the appendix of the OECD (2011) survey report, plus Greece.

The most direct measure of the quality of country PFM systems, and the one with the most comprehensive data coverage, is question 13, titled “Quality of Budgetary and Financial Management,” from the World Bank’s Country Policy and Institutional Assessments (CPIA). For brevity we re-name this variable “PFM Quality.” Higher ratings reflect a comprehensive and credible budget linked to policy priorities, effective financial management systems to ensure that the budget is implemented as intended, and timely and accurate accounting and fiscal reporting.

In a robustness test, we use a second CPIA indicator (question 16), titled “Transparency, Accountability and Corruption in the Public Sector.” For brevity, we name it “Transparency.” It assesses the extent to which the executive can be held accountable for its use of funds and the results of its actions by the electorate and by the legislature and judiciary, and the extent to which public employees within the executive are required to account for the use of resources, administrative decisions, and results obtained. Weaker country systems as measured by lower scores on PFM Quality and Transparency reflect increased risks to donors of corruption scandals, diversion of funds to lower priority uses, or inability to account for how funds were spent. The Paris Declaration explicitly asserts that corruption “inhibits donors from relying on partner country systems.” We therefore expect coefficients to be positive for both PFM Quality and Transparency, in our use-of-country-systems regressions.

The CPIA indicators are produced annually by World Bank staff for aid allocation purposes, for approximately 135 developing countries. Assessments are on a 1 to 6 scale, including half-point increments.⁵ We use a one-year lag of the CPIA ratings in our tests, to reduce the potential for capturing reverse causality. In our main sample PFM Quality ranges from 2 to 5.5, with a mean of about 3.4. The mean value of Transparency is about 3.0, with a

⁵ For example, if a country meets some of the criteria for a rating of 3, but also meets some of the criteria for a rating of 4, it will be assigned a rating of 3.5.

minimum value of 1.5 and maximum of 5.0.

The Freedom House civil liberties indicator is used as a supplementary indicator measuring other aspects of the quality of recipients' aid management systems. Civil liberties scores are higher in countries with a free and independent media, freedom of assembly, freedom for NGOs, and greater personal autonomy and individual freedoms.⁶ The use of donor and other public funds is likely to be more transparent in a more open society. Moreover, the principle of "country ownership" may have a stronger appeal to donors operating in countries with stronger mechanisms for public voice and government accountability to broader constituencies. In short, donors will tend to trust country systems more when aiding recipients that look more like most donor countries, i.e. those that are relatively open and democratic. Where there is reason to think local elites are not committed to development objectives and will divert aid to maintain their political and economic power (Bjornskov, 2010; Angeles and Neandis, 2009), donors are less likely to provide budget support and to trust country PFM and procurement systems. Civil Liberties can measure the potential for elite capture of aid flows.

Other recipient characteristics included as control variables are log of per capita income, recent growth in per capita income (average annual increase in the two preceding years), log of population, and log of per capita aid. These variables can capture the effects of real or perceived quality of country systems, separate from what is measured by PFM Quality and Civil Liberties.⁷ Donors may sometimes infer that wealthier or more rapidly growing countries have stronger capacity for managing aid, in the absence of more direct evidence. If there are economies of scale in aid management systems, their quality may tend to be higher in larger countries. For a

⁶ An index scored from 1 to 7 is available going back to 1972, but we use a more finely-grained version scored from 0-100 that is available for 2003 onward. The political freedoms indicator also produced by Freedom House is correlated with civil liberties at .95, and results are similar if it is used instead.

⁷ Anecdotal evidence from country evaluations suggests that donors sometimes bypass country systems because they are "slower and more cumbersome" as well as vulnerable to corruption (Wood et al., 2011).

given level of expected quality, uncertainty about quality is likely higher in smaller countries, where donors have obtained relevant information from fewer projects and have undertaken fewer diagnostic and other analytic studies. If donors tend to be risk averse, this higher uncertainty can reduce use of country systems.

Use of country systems may be higher in more aid-dependent countries for two reasons. First, delivering higher volumes of aid may be accomplished in part by disbursing a greater percentage of it in the form of direct budget support, which by definition uses country PFM systems. Second, the impact of donor decisions to use (or bypass) country systems on strengthening (or undermining) government capacity and accountability is likely to be greater where aid volumes are higher. These potential impacts may also be greater in lower-income countries, perhaps explaining in part why governments of many middle-income countries receiving modest levels of aid show relatively little concern over donors' choices of aid management practices (Wood et al., 2011). The net effect of income on use of country systems is therefore ambiguous.

Use of country systems is expected to increase with the quality of recipient countries' national development strategies, as measured by PD Indicator 1. These qualitative assessments were based on three criteria (OECD, 2008; World Bank, 2006): (1) existence of an authoritative country-wide development policy; (2) realism of the development policy with clearly-identified priorities; and (3) explicit links with policies in the budget.⁸ This "Strategy" indicator is scored on an A-E scale. Nearly half of the observations in the sample receive the middle score of C; only 2% have the highest grade of A and 1% have the lowest grade of E.

⁸ Scores were assigned by World Bank staff based in part on information provided by government officials responding to a PDMS questionnaire. These assessments are not done by the same Bank staff responsible for the CPIA. Clist, Isopi and Morrissey (2012) show that countries with an explicit national development strategy are more likely to receive aid in the form of budget support from the EC and World Bank, but they do not differentiate strategies by quality on the basis of the relevant PD indicators.

A second qualitative indicator (PD Indicator 11) assesses the quality of recipients’ “results-based frameworks.” It considers (1) the quality of the information generated; (2) stakeholder access to information; and (3) the extent to which the information is utilized within a country level monitoring and evaluation system (OECD, 2008a). The highest grade observed in the sample for this “Results” indicator is B (about 14% of observations), and the most frequent grade is C (about 55%).

“Strategy” and “Results” scores are not available for some of the middle-income country observations in the sample. We therefore do not control for them in every test, but merely show that other key findings are robust to their inclusion.

Donor characteristics, such as differing mandates of bilateral and multilateral donors, may also affect use of country systems (Knack, 2013). Here we control for those influences with donor-year fixed effects.

A final control variable is donor aid share, defined as the percentage of total aid (inclusive of aid not to the government sector) to recipient j accounted for by donor i . A higher aid share gives a donor a stronger reputational stake in a recipient’s long-run development, strengthening incentives for the donor to build instead to bypass country systems. Higher donor aid shares are thus predicted to increase use of country systems (Knack, 2013). The sample mean for donor aid share is 7%, with a minimum value of 0.0014% (Switzerland in Egypt in 2010) and a maximum of 83.6% (World Bank in St. Vincent in 2010).⁹

We estimate regressions of the general form:

$$UCS_{ijt} = \beta Z_{ijt} + \delta X_i Y_t + \eta M_{jt-1} + u_{ijt}$$

⁹ Larger donors at the global level do not always have higher average aid shares at the recipient level, as some donors concentrate their aid in fewer countries. Portugal’s average aid share (23.5%) exceeds that of the U.S. (12.4%), although the U.S. provided more than 100 times as much total aid to countries in the sample as Portugal.

where UCS_{ijt} is the share of donor i 's aid to recipient j in year t that is managed by recipient PFM systems, Z_{ijt} are regressors that vary by donor, recipient, and year; M_{jt} are regressors that vary by recipient and year but are donor-invariant, and $X_i Y_t$ are donor-year fixed effects. All recipient variables are lagged by one year.

The dataset can be treated as an unbalanced panel, with anywhere between 1 and 80 observations per donor. With the inclusion of donor-year fixed effects estimates are informed only by cross-recipient variation in the data within each of the 102 (= 34 x 3) donor-year panels. The analysis follows a highly conservative approach to determining statistical significance of estimates. All tests correct standard errors for two-way clustering (Cameron, Gelbach and Miller, 2011), for non-independence of errors within clusters of observations pertaining to each donor and to each recipient.

4. Results for All Donors

Equation 1.1 of Table 1 shows a positive and highly significant bivariate relationship between use of country PFM systems and PFM quality. Each 0.5-point increase in the 1.0-6.0 scale for PFM Quality is associated with an increase of more than 6 percentage points in use of country PFM systems. This finding is not consistent with the conclusions of OECD (2011a), IEG (2011), and Ellmers (2011) that donors' decisions on use of country systems are insensitive to their quality. Moreover, the relationship is even stronger when the sample is limited to the subset of countries eligible for the World Bank's concessionary loans and grants, i.e. IDA recipients. This is the same sample examined in OECD (2011a), because CPIA ratings are not public for non-IDA countries. As shown in equation 1.2, in IDA countries each 0.5-point increase in PFM Quality is associated with an increase of about 3 percentage points in use of

country PFM systems. If donors' decisions to use country PFM systems are influenced by the PD-endorsed measure of PFM Quality, then we should observe a stronger relationship in a sample limited to recipients with publicly-available CPIA ratings than in a sample that includes a mixture of public and non-public ratings.¹⁰ Results from equations 1.1 and 1.2 are strongly consistent with this prediction.

The explanatory power of these regressions ($R^2 = .05$ in 1.1 and $.08$ in 1.2) suggests that donors' decisions are influenced by many other factors, potentially including other recipient characteristics. The OECD (2011: 50, 52) report and Ellmers (2011) both emphasize the large variation in use of country systems across recipients, for a given level of PFM Quality, that is reflected in these low R^2 values. Nevertheless, this variation should not obscure the reality of a large, positive and significant relationship between the two variables.

Donor-specific factors such as tolerance for risk undoubtedly contribute to the large variation in use of systems at a given level of quality (Knack, 2013), but they cannot account for the bulk of variation in the data. A full set of donor-year dummy variables (in the absence of any other regressors) statistically explains only 20% of the variation in use of country PFM systems, compared to 23% for a full set of recipient-year dummies. Recipient characteristics thus appear to matter somewhat more than donor characteristics.¹¹ Henceforth, we control for donor-year fixed effects and for several additional recipient-level variables, in testing the robustness of our initial results from equations 1.1 and 1.2. Estimates are therefore informed only by cross-

¹⁰ We cannot entirely rule out the possibility that some donor agency officials may have obtained the confidential ratings for non-IDA countries through unofficial channels.

¹¹ In some instances, particularly where line ministries have a larger role in managing aid and interacting directly with donors, recipient country officials may actually prefer using donor systems over their own central government systems (OECD, 2011a: 52).

recipient variation in the data within each donor-year panel of observations. This method effectively controls for any omitted donor characteristics, even if they vary by survey year.¹²

Equation 1.3 presents results from the baseline multivariate model. Coefficients for PFM Quality and Civil Liberties are positive and significant at the .001 level. Each 0.5-point increase in the 1.0-6.0 scale for PFM Quality is associated with an increase of about 3.8 percentage points in use of country PFM systems. Each 10-point increment in the civil liberties index (scaled from 0 to 100) is associated with an increase in use of country PFM systems by 5.5 percentage points. Results in equation 1.3 also show that donors use country systems more in larger and more aid-dependent recipients; both of these relationships are significant at the .001 level. The level of per capita income is not significant, but the recent growth rate of income has a positive, marginally significant coefficient. In countries where a donor has a larger share of the aid market, its aid is more likely to be managed via country PFM systems. The coefficient of .77 indicates that for each 10 percentage point increase in the donor's share of aid in a country, its use of country systems increases by 7.7 percentage points.

Equation 1.4 reports the same specification but limited to the World Bank's IDA countries, i.e. those eligible for its concessionary loans and grants. As was the case in the bivariate tests of equations 1.1 and 1.2, donors' use of country PFM systems is more responsive to PFM Quality in IDA countries than in the larger sample that included middle-income countries. The coefficient for Civil Liberties is also slightly larger in 1.4 than in 1.3, while the growth coefficient is somewhat smaller and not significant in 1.4.

¹² Our dependent variables are bounded by 0% and 100%, so tobit estimation might be considered more appropriate than OLS. However, fixed-effects tobit estimates are biased, so using tobit would come at the expense of greater omitted variables bias associated with dropping the donor-year fixed effects. Moreover, the value added of tobit is relatively small: empirical findings typically are very similar using tobit or OLS (Angrist and Pischke, 2009). In the tables we therefore report OLS fixed-effects regressions rather than tobit. However, the coefficient for PFM Quality is significant at the .001 level if equations 1.1 and 1.2 are run using tobit, random-effects tobit, or with OLS but dropping all observations equal to either 0% or 100%. (Results available on request.)

Coefficients on PFM Quality from Table 1 can help shed light on the degree to which the upward trend in use of country PFM systems over time, as reported in OECD (2011a), may be attributed to any upward trend in quality. For 31 countries with data for 2005 and 2010, on both PFM Quality and use of PFM systems, the latter averaged 32.6% in the initial year, increasing to 37.5% by 2010. Mean PFM Quality increased only from 3.56 to 3.61. Based on the coefficient of 7.5 in equation 1.3, improvements in PFM Quality account for an increase of only about 0.4 percentage points in use of PFM systems, less than one-tenth of the actual increase. Even based on the larger coefficient in equation 1.2, improvements in quality statistically explain only about one-fifth of the observed increase in use. Performing the same exercise for the 49 countries with the necessary data available in the shorter time span 2007 to 2010, the small average improvement in PFM Quality (from 3.35 to 3.39) is even less effective in accounting for the 4.4 percentage point increase in use of PFM systems (from 30.0% to 34.4%). Civil Liberties is even less effective in accounting for the observed improvement in use of PFM systems, as the mean value for Civil Liberties actually worsens slightly over time in these samples. The overall increase in donors' use of country PFM systems, for a given set of countries and controlling for other factors, suggests that the PD campaign may have succeeded in encouraging donors to accept more (short-term) risk under the premise that bypassing country systems undermines them, thereby slowing development in the long run.

The remainder of this section reports robustness tests, mostly designed to address three potential problems with the basic results reported in Table 1. First, those estimates are potentially affected by sample selection bias. Second, there may be omitted variables correlated with quality and use of country systems, despite the inclusion in 1.3 and 1.4 of various control variables and donor-year fixed effects. Third, coefficients for PFM Quality may be biased

upward, if use of country systems strengthens them. Results turn out to be robust to correcting for each of these potential problems in turn.

Sample selection issues

There are two sources of potential selection bias in the data. First, aid-recipient governments self-select into (or out of) the PDMS. Of the 151 eligible aid recipients on the official DAC list, only 80 participated in one or more rounds of the survey, and fewer than half of those 80 participated in all three rounds. Second, not all donors choose to provide aid to all recipients in all years (or, in some cases, recipients may choose not to accept aid). Pooling the three survey years, nearly 64% of potential observations are censored by recipient non-participation in the PDMS (although the majority of these would be censored anyway due to zero aid flows), and more than 60% of the remaining pairs are censored by donors' provision of zero aid.

Moreover, the censored observations differ systematically from the uncensored observations, potentially creating a selection bias in our tests. Use of country systems averages 38.3% for 2010 among all observations corresponding to the initial (2005) set of PDMS recipient participants. The average for 2010 is only 28.5% for recipients that joined the survey in 2007, and is only 26.5% for recipients that did not participate until 2010. Very few recipients dropped out, but many were added over the three rounds, and use of country systems is markedly higher in the earlier volunteers. Uncensored observations differ significantly from censored observations in many important respects: they are larger, poorer, and more aid-dependent.

Correcting for potential selection bias using the Heckman approach requires identifying exogenous variables that affect selection by recipients into the PDMS, and by donors and recipients into active aid relationships. Selection regressions are reported in Appendix Table A2.

The sample in Equation A2.1 is the full set of 15,402 pairings of the 34 donors and 151 eligible DAC recipients for each of the three survey years. Selection variables include population, HIPC status, a colonial tie dummy, and distance from the donor capital or headquarters city (in log of kilometers).¹³ Coefficient estimates presented in Table A2 represent marginal effects from a probit regression, evaluated at the means of all of the other regressors.

The selection variables are all highly significant in equation A2.1, which includes data from all three survey years. The probability of a donor-recipient pair being uncensored increases by 20 percentage points if they have a colonial tie, by 26 points if the recipient had reached the HIPC completion point (for full irrevocable debt cancellation), and by 11 points if it had reached the HIPC decision point (for interim debt relief). Greater geographic distance separating the donor from the recipient significantly increases the probability of censoring, as expected. Larger recipients are less likely to be censored. Year dummies confirm that censoring has diminished with each subsequent round of the PDMS.

The three remaining equations in Table A2 report results for 2005, 2007 and 2010 respectively. Country size, colonial ties and distance matter more in the later years, while HIPC status matters more in the earlier years. The pseudo- R^2 for this selection model drops from .23 in 2005 to .20 in 2007 and .11 in 2010, suggesting that the PDMS sample becomes more representative as additional recipient countries join later rounds of the survey.

To test and correct for potential sample selection bias in our use of country systems regressions, we compute inverse Mills' ratios from equations A2.2-A2.4, allowing the impact of selection variables to vary by survey year (Wooldridge, 2002). Equation 2.1 of Table 2 adds the inverse Mills' ratio (IMR) to the baseline specification of equation 1.3. The IMR is inversely

¹³ Knack (2013) provides more detail, including theoretical justifications for selection variables. Other regressors from equation 1.3 are not included in selection regressions because they would reduce the sample size, defeating the purpose of testing for sample selection bias.

related to the predicted probability of selection, so its negative coefficient estimate in equation 2.1 implies that use of country systems is positively related to the selection probability. It is not significant at conventional levels, however, so there is only weak evidence of selection bias. Moreover, the addition of the IMR to the model has very little effect on results for PFM Quality and other variables. For simplicity, therefore, we do not correct for potential selection bias in the remaining tests reported in Tables 2 through 5. Note that these tests all include donor-year fixed effects, and the potential for selection bias is reduced (although not eliminated) in the presence of fixed effects (Wooldridge, 2002).

Omitted variables

Equations 2.2-2.4 test the robustness of results for PFM Quality and other recipient characteristics to the inclusion of additional regressors. The quality of national development strategies and results-based frameworks are controlled for in equation 2.2. Use of country PFM systems is significantly higher in countries with better scores on the Results indicator: each 1-grade improvement from E to B is associated with an average increase of 4.5 percentage points. The ordinal-level Strategy indicator is insignificant (results not shown), but a dummy variable for the few countries with the top score of A (Tanzania and Rwanda in 2010) was significant, as shown in Table 2, with a coefficient of 14.7. Both Strategy and Results are significantly correlated (at .46 and .40 respectively) with PFM Quality, so it is not surprising that the latter's coefficient is about a percentage point lower in equation 2.2 relative to 1.3 (dropping from 7.5 to 6.6). However, PFM quality remains significant at the .001 level in equation 2.2.

Equation 2.3 adds the Transparency indicator from the World Bank's CPIA. Transparency and PFM Quality are correlated at .59 in our sample. The latter's coefficient declines from 7.5 in equation 1.3 to 5.5 in equation 2.3, but is still significant at the .05 level.

Transparency itself is also significant at the .05 level, and each 0.5-point improvement on its 1-6 scale is associated with an increase of 4 percentage points in use of country PFM systems.

Equation 2.4 adds a dummy variable for 12 “fragile states” (listed in OECD, 2011a: 104). In 2007 the OECD-DAC adopted a set of “Fragile States Principles” reflecting a view that they had different problems requiring special treatment in comparison with better-performing aid recipients. One of the 10 principles calls for “a focus on state building as the central objective,” suggesting that use of country PFM systems may be higher in fragile states than otherwise expected, based on income, size, etc. On the other hand, donors might rely less on country systems in fragile states, if these countries are perceived to involve extra risks to donors, beyond the risks measured by PFM Quality and Civil Liberties. Results in equation 2.4 suggest that the latter effect dominates. Use of country PFM systems is 10 percentage points lower on average in fragile states, other things equal, and this effect is significant at the .05 level. Controlling for the fragile states dummy, coefficients for PFM Quality and Civil Liberties are somewhat smaller than in the baseline specification of equation 1.3, but they remain significant at the .01 level.

Equation 3.1 adds the percentage of aid for each donor-recipient pair that is in the form of direct budget support (DBS), as reported in the PDMS. Budget support (which comprises 12.5% of aid on average in our sample) uses country PFM systems, by definition. When this variable is included, the regression in effect is testing for relationships between the regressors and use of PFM systems for project aid and other non-DBS aid. Coefficients for PFM Quality and Civil Liberties are again somewhat smaller in 3.1 than in 1.3, but they remain highly significant.

Results for PFM Quality are potentially weaker for donors other than the World Bank. The CPIA assessments are conducted by Bank staff, so they are particularly likely to affect decisions by the Bank on aid implementation. Equation 3.2 therefore replicates the base

specification of equation 1.3, but dropping the 145 World Bank observations. The coefficient for PFM Quality actually increases slightly, from 7.5 to 7.7, suggesting that judgments of Bank staff on quality of PFM systems tend to be consistent with judgments of most other donors.¹⁴

Technically, PFM Quality is an ordinal rather than a cardinal indicator. Equation 3.3 tests whether or not its effects are approximately linear and monotonic, as assumed by its use as a cardinal indicator, by substituting a set of dummy variables for PFM Quality. Results in 3.3 are supportive. The omitted category is the set of observations with PFM Quality = 2. Relative to this omitted group, coefficients for dummy variables representing higher values for PFM Quality are broadly consistent with the cardinality assumption. Specifically, they increase with higher values of PFM Quality, and at a roughly equal rate, although there is some evidence of a diminishing marginal effect towards the upper end of the scale. Model fit is not improved by substituting the more flexible specification represented by these dummy variables: the R^2 is .31 in both equations 1.3 and 3.3. Thus, treating PFM Quality as a cardinal indicator – for greater simplicity and conservation of space in tables - does not come at the expense of a significant loss of information in the data.

Equation 3.4 substitutes donor-recipient fixed effects and year dummies for the donor-year fixed effects. In this specification, estimates are informed only by variation over time within each donor-recipient pair in use of country PFM systems, PFM Quality, etc. This is therefore a particularly strong test, as Civil Liberties and other recipient characteristics may exhibit little variation over the three periods, covering only a 5-year time span. The donor-recipient fixed effects implicitly control for the effects of any pairwise historical, commercial or

¹⁴ Similarly, Clist, Isopi and Morrissey (2012: 280) report that the Bank's CPIA indicators on Public Sector Management (including PFM Quality) matter more for the EC's decisions to provide general budget support than they do for the Bank's own budget support decisions. Table 5 below reports donor-specific tests of the responsiveness of use of country PFM systems to PFM Quality.

other ties as well as for any donor-specific biases regarding quality of systems, to the extent they are invariant over the three periods. Despite the stringent nature of this test, several of the variables significant in previous tests remain significant in equation 3.4. Most notably, PFM Quality is significant at the .05 level, with a coefficient only slightly smaller than in many of the other regressions. Year dummies are not statistically significant, although the coefficient on the 2010 year dummy can be interpreted as a 2.6 percentage point secular increase in use of country PFM systems relative to the 2005 base year, controlling for the other variables in the model.

In Table 4 we exploit our larger dataset to test the robustness of findings from IEG (2011) and Ellmers (2011). Equation 4.1 substitutes an index of PEFA indicators for the CPIA measure of PFM Quality. The coefficient is positive -- a one-standard-deviation increase in the index value is associated with a 1.6 percentage point increase in use of country PFM systems -- but not significant. This finding is consistent with the IEG's (2011) study of World Bank aid to low-income countries, where PEFA is not significantly related to use of country PFM systems.

Even in our larger study, more than two-fifths of the observations are lost by using PEFA instead of CPIA. However, when the CPIA indicator of PFM Quality is tested on the PEFA sample in equation 4.2, its coefficient is positive, highly significant and of about the same magnitude as in the larger sample. The PEFA and CPIA indicators are correlated at .77, but they differ enough that they perform very differently in our tests, even in the same restricted sample. The findings from equations 4.1 and 4.2 confirm that the empirical link between quality and use of recipient PFM systems is sensitive to the measure of quality. There is no way of ascertaining which indicator, if either, is a more accurate or relevant measure of PFM quality, so it would be more reassuring if use of PFM systems were strongly related to both of them. However, it is meaningful that of the two, the one specified in the PD goals is the one that does significantly

predict use of PFM systems. This finding is consistent with the view that donor decisions are influenced by the PD indicators, as part of a serious effort to live up to the PD commitments.

Equation 4.3 tests the robustness of Ellmers' (2011) finding of no significant relationship between PFM Quality and use of country procurement systems. In this vastly expanded sample, we find a positive, highly significant and large effect. Each 0.5-point increase in PFM Quality is associated with a 3.4 percentage-point increase in use of procurement systems. The CPIA indicator of PFM Quality is not designed specifically to measure the quality of country procurement systems. We therefore provide a supplementary test in equation 4.4, where the PEFA indicator on procurement system quality (from question 19 in the PEFA questionnaire) is substituted for PFM Quality.¹⁵ Although this substitution cuts the sample size by nearly half, the procurement quality coefficient is positive and highly significant. Each 1-unit increment in the 7-point scale is associated with a 3 percentage point increase in use of country procurement systems.¹⁶

Endogeneity

If country systems are strengthened when used, and weakened when bypassed in favor of donors' parallel (OECD, 2011a; 2008), then PFM Quality is endogenous. We address this potential problem in several ways. First, PFM Quality is lagged by a year in our tests. Second, results from equations 4.1 and 4.2 are consistent with a causal interpretation that donors follow the PD's prescriptions in assessing quality of country PFM systems using the CPIA indicator in deciding whether to trust in them or bypass them. The PEFA indicator should be expected to

¹⁵ The "official" PD indicator of procurement systems quality is the OECD-DAC's "Methodology for Assessing Procurement Systems," or MAPS. However, data are available for only 5 countries in 2010, 17 other countries for 2007, and none in 2005. "Given the small sample size and the one-off nature of the assessment... it is difficult to draw any conclusions" from them (OECD, 2011a: 47).

¹⁶ We found no evidence of heterogeneous effects of PEFA indicators in explaining use of PFM and procurement systems. Specifically, coefficients on PEFA indicators are very similar for two groups of donors: the seven (EC, IMF, World Bank, France, Norway, Switzerland and the UK) that created and fund the PEFA initiative and Secretariat, and the other 27.

have weaker predictive power than the CPIA indicator, because the latter but not the former is included in the PD monitoring framework. To the extent that using systems strengthens them, this reverse effect should matter equally for either measure of PFM quality, whether PEFA or the CPIA. Therefore, if causation is mostly from PFM quality to use, the CPIA should be more strongly associated than PEFA with use of country systems. If causality is mostly the other way around, however, CPIA and PEFA should not differ much in the strength of their association with use of country systems. Results in 4.1 and 4.2 clearly show a stronger relationship for CPIA, consistent with the view that PFM Quality has a causal effect and its positive, significant coefficient is not merely the product of endogeneity bias.

Finally, in another approach to the endogeneity issue, we test in Table 5 for the presence of heterogeneous effects among donors that should be present – or absent - if coefficients on PFM Quality are biased upward to a significant degree by reverse causation. Our basic identification strategy is premised on differences among donors. First, we distinguish donors by size. The impact of a percentage point increase in a donor’s use of country systems should vary with a donor’s aid share. Each \$1 increase should have the same impact, so 1 percentage point will imply a larger impact for larger donors. A sufficiently small donor could safely take system quality as given in its decisions. In contrast, there is no particular reason to expect the causal impact of PFM Quality on use of country systems to vary by donor size. Therefore, if coefficients on PFM Quality in Tables 1 and 2 mostly reflect endogeneity bias, the strength of association should vary positively with donor size. We add the appropriate interaction term in Table 5, equation 5.1, to test this hypothesis.¹⁷ Results are not consistent with reverse causation:

¹⁷ The interaction term is equal to the product of the deviations of PFM Quality and donor aid share from their respective means. The coefficient on PFM Quality thus indicates its marginal effect conditional on the mean value of donor aid share (rather than the out-of-sample value of 0).

the coefficient on the interaction term is not only insignificant, but is negative, not positive as implied by an argument for reverse causation.

In our second variant of this identification strategy, we distinguish donors by selectivity in aid allocations. Specifically, we posit that donors that are more sensitive to quality of country policies and systems when allocating aid will also be more sensitive to the quality of systems when deciding whether to rely on them for managing their aid. In contrast, there is no particular reason to expect the strength of any reverse causation from use of country systems to vary with donors' sensitivity to policy in allocating aid. An empirical finding of heterogeneous impacts in this instance can be explained better by the causal argument than by reverse causality (from use of systems to system quality). Equation 5.2 includes an interaction term, to test the hypothesis that PFM Quality will affect use of country systems more for donors that are more policy selective in allocating aid among recipients.¹⁸ For this purpose, we use the donor Policy Selectivity scores from Knack, Rogers and Eubank (2011), based on aid disbursements in 2007. The interaction coefficient is positively signed, as hypothesized, and significant at the .01 level. These heterogeneous impacts are more consistent with a causal explanation from PFM Quality to use of PFM systems than with an argument for reverse causation from use to quality.

Endogeneity bias cannot be entirely ruled out by the results in Tables 4 and 5, or of course by the use of lagged values of PFM quality in all of the regressions. However, these results are all consistent with a direction of causality running predominantly from PFM Quality to use of country PFM systems, and much harder to square with an argument of reverse causality.

¹⁸ The interaction term is equal to the product of the deviations of PFM Quality and Policy Selectivity from their respective means. The coefficient on PFM Quality thus indicates its marginal effect conditional on the mean value of Policy Selectivity. Because Policy Selectivity is donor-invariant, its coefficient is subsumed by the donor-year fixed effects.

5. Donor-specific Results

Table 6 disaggregates the data by donor. Separate regressions are reported for each of them. The specification is similar to the baseline from equation 1.3, but with year (rather than donor-year) fixed effects. For space reasons, results are shown only for PFM Quality, Civil Liberties and year effects.¹⁹ Donor regressions are sorted by the magnitude of coefficients on PFM Quality, from the largest to the smallest. Of the 24 regressions, the PFM Quality coefficient is positive in 19 and negative in only 5. In all 8 cases where the coefficient is significant at the .10 level or better, it is positive. Norway, Finland, the UK, and Canada have highly significant (.05 or better) coefficients of 20 or more, indicating that their use of country PFM systems increases by 10 or more percentage points for each 0.5-point increment in PFM quality.

For 21 of the 24 donors, the coefficient on Civil Liberties is positive, including in all 9 cases where it is significant at the .10 level or better. The 16 bilateral donors account for all 3 negative coefficients on Civil Liberties, and the 8 multilaterals account for 4 of the 9 positive and significant coefficients. No donor has a negative coefficient on both PFM Quality and Civil Liberties. The UK, Canada and the Netherlands have positive and significant coefficients on both variables. These findings indicate that not only donors overall, but most of them individually, are responsive to the quality of country systems in deciding whether to rely on them.

The PD envisions increased reliance over time on country PFM systems by donors for two reasons. First, it sets a goal of improving the quality of systems in recipients, and there is some favorable evidence of gradual progress in this respect (OECD, 2011a: ch. 3). Coupled with the positive relationship between quality and use of PFM systems demonstrated in Tables 1-6,

¹⁹ Full results are available on request.

this gradual improvement in quality should be reflected in gradual increases in their use over time. Second, a premise of the PD campaign and goals is that donors have been too risk averse in using country systems, and that donors should strengthen them by relaying on them more, for any given level of system quality. If most donors are acting on these exhortations, it should be reflected in coefficients on the year dummies for 2007 and 2010 in the Table 6 regressions, where the base category is 2005, and PFM Quality already controls for any changes in quality of systems over time.

The majority of donors in fact appear to be increasing their use of country PFM systems over time, controlling for PFM Quality and other variables in the baseline specification. Only 11 of the 24 coefficients for the year=2007 dummy variable in Table 6 are positive. However, most of them are not statistically significant, and all 4 that are significant (Denmark, Spain, World Bank and Japan) are positive. By 2010, 5 years instead of 3 had elapsed since the 2005 baseline year, so the difference between 2005 and 2010 is a more meaningful indicator of trends. Of the 24 coefficients for the year=2010 dummy, 16 are positive, including 6 of the 8 that are statistically significant.

Donors that tend to use country PFM systems at above-average levels are not necessarily the same ones that are most responsive to system quality. Figure 1 portrays each of the 24 donors in Table 6 on two dimensions: regression coefficients for PFM Quality are plotted on the Y axis, and mean use of country systems (listed toward the right side of Table 6) is plotted on the X axis. It turns out there is a strong positive correlation (correlation = .58) between them. About half of the donors are located in the upper-right quadrant of the scatterplot, representing above-average values of both responsiveness and mean use. Denmark (in the lower right quadrant) stands out as a donor that tends to use country PFM systems at above-average rates in both more

and less risky environments. The USA, conversely, stands out as one using country systems at very low rates, regardless of risk levels. It is the donor that best exemplifies Ellmer's (2011: 19) assertion that "some bilateral donors have extremely low tolerance to fiduciary risks which makes it very difficult for them to use country systems." Belgium (in the upper left quadrant) exemplifies yet another very different pattern: of the numerous donors that are highly responsive to risk, it is the only one that tends not to rely much on country PFM systems.

6. Summary and Conclusions

The 2005 Paris Declaration (PD) exhorts donors to deliver aid in ways more consistent with recipient country priorities and institution-building. Among other specific provisions, it urges donors to increase their use of country PFM systems, particularly where systems are more trustworthy so that their use does not entail inordinate reputational, fiduciary or other risks to donors. The 2008 Accra Agenda for Action (AAA) asserts that "even where there are good-quality country systems, donors often do not use them," and commits donors to use them "to the maximum extent possible" and to ensure that the internal incentives for donor agency staff members are aligned with this goal. The PD and AAA acknowledge however that use of country systems may not be feasible in some cases, and numerical targets for their use were set at more ambitious levels where quality of systems is stronger, as measured by PFM Quality.

Based on a three-dimensional panel dataset constructed from all three rounds of the PD Monitoring Survey, this study provides empirical evidence supporting the proposition that use of country systems is strongly related to their quality. This relationship holds not only for donors collectively, but also for most donors individually. This finding is contradictory to the conclusions of other (less rigorous) analyses that quality of PFM systems has little or no effect

on donors' use of them (IEG, 2011; OECD, 2011a, 2009a, 2009b; Ellmers 2011). The relationship between PFM Quality (and Civil Liberties) and use of country PFM systems remains strong and significant when controlling for donor-year fixed effects, where estimates are informed only by variation in the data over recipients within a particular year. Results are also robust to a variety of corrections for possible sample selection, omitted variables and endogeneity bias. However, PFM Quality is not the sole determinant of donors' decisions. Use of country PFM systems increases with Civil Liberties, population, and total aid per capita in the recipient country, and with a donor's share of the aid market in the country. There is a positive but modest time trend in use of country PFM systems, controlling for PFM Quality and other variables. The 5-year time span under study is too short, however, to assess whether the PD and AAA agenda on use of country systems will have positive medium and long-run impacts on donor behavior.

Despite domestic political and other pressures on many of the DAC donors, the PD campaign and associated "peer pressure" resulting from it has likely had some positive, impact on donors' use of country PFM systems. Any such effects may be one-off or even temporary, however: in the medium and long term, donors' decisions to use country systems are likely to depend more on the level of risk that their domestic constituents find tolerable, and the level of risk incurred in trusting recipient countries' PFM systems. Slow growth in most donor countries in the last several years is creating pressure to reduce aid volumes (OECD, 2012; Dang, Knack and Rogers, forthcoming), and perhaps also to minimize fiduciary and other risks in aid management in an effort to protect aid budgets. Further sizeable increases in the use of country PFM systems may thus be contingent on a gradual, sustained improvement in the quality of systems in aid recipient countries.

Because successful reform of PFM systems tends to occur only over long time periods, and involves political as well as technical challenges (OECD, 2009b; Allen, 2009; IEG, 2008), it is probably not realistic to expect dramatic improvements. Improvements over time in PEFA indicators are observed more often for aspects of PFM where higher scores can be attained merely by adopting new laws, regulations or technical tools (such as the existence of an annual budget calendar), or which require changing the behavior of more officials or agencies (e.g., an indicator on whether a debt sustainability analysis has been done in the last three years). Where higher scores require evidence of actual implementation, and/or changing the behavior of numerous officials or agencies (e.g. deviation in actual from budgeted expenditures, or extent of unreported extra-budgetary expenditures), progress is observed in fewer cases (PEFA, 2011).

The Paris Declaration itself emphasizes the importance of strengthening weak country systems, and that technical assistance from donors can contribute importantly to this objective.²⁰ In many cases, however, the binding constraint for PFM reform is not donor resources or even government capacity, but rather resistance from certain public officials who benefit from maintaining less-transparent systems and procedures (OECD, 2012; 2011b; 2009b; Allen, 2009). Donors have an important role in the aid effectiveness agenda. However, the PD and AAA not only commit recipient countries to strengthen their systems, but also place primary responsibility on them for their own capacity development, with donors playing only a supporting role.

There are no current plans for a fourth round of the PDMS. Any “peer pressure” effects it had on donor behaviors may dissipate over time. Alternative indicators of donor performance produced by academic researchers and development NGOs can at least partially fill the gap

²⁰ The multi-donor Public Expenditure and Financial Accountability program (www.pefa.org) replaces numerous uncoordinated donor projects supporting diagnostic and analytic work, financing for reform, and technical support for reform implementation in PFM. An empirical study by de Renzio, Andrews and Mills (2010) finds a modest but statistically significant link between donor support for PFM reform and quality of PFM systems (as measured by PEFA indicators).

created by ending the PDMS. The most prominent of these are the Aid Transparency Index produced by the NGO “Publish What You Fund”²¹ and the Quality of Official Development Assistance (QuODA) indexes produced by the Center for Global Development and the Brookings Institution (Birdsall and Kharas, 2011). These indicators borrow from the PD, but go beyond it to incorporate other aid transparency and quality indicators. They also differ from the PD in implicitly assigning full responsibility for quality of aid to donors. The PD indicators on use of country systems, included in the QuODA indexes, represent a mixture of donor and recipient country performance, consistent with the PD’s message that aid effectiveness is a shared responsibility. As mentioned above, recipient fixed effects statistically explain a slightly larger part of the variation in use of country PFM systems in our donor-recipient observations than do donor fixed effects. In measuring the quality of aid and ranking donors, it is therefore important to adjust for characteristics of the recipient countries in which donors are operating. Otherwise, donors with larger shares of their aid portfolios allocated to riskier countries are implicitly penalized in the rankings. The one-off ranking of donor performance by Knack, Rogers and Eubank (2011) includes such adjustments, but the more prominent and permanent QuODA project does not (despite being impressively designed and executed in most other respects), and thus risks creating the wrong incentives for donors.

²¹ See <http://www.publishwhatyoufund.org/index/2013-index-changes/>

REFERENCES

- Allen, Richard, 2009. The Challenge of Reforming Budgetary Institutions in Developing Countries. IMF Working Paper WP/09/96. Washington DC: International Monetary Fund.
- Angrist, Joshua D., Pischke, Jorn-Steffen Pischke, 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton.
- Angeles, Luis, Neanidis, Kyriakos C., 2009, Aid Effectiveness: The Role of the Local Elite. *Journal of Development Economics* 90(1), 120-134.
- Birdsall, Nancy, Kharas, Homi, 2011. Quality of Official Development Assistance Assessment. Center for Global Development, Washington, DC.
- Bjornskov, Christian, 2010. Do Elites Benefit From Democracy and Foreign Aid in Developing Countries? *Journal of Development Economics* 92(2), 115-124.
- Cameron, A. Colin, Gelbach, Jonah B., Miller, Douglas L., 2011. Robust Inference With Multi-Way Clustering. *Journal of Business and Economic Statistics* 90(3), 414-27.
- Clist, Paul, Isopi, Alessia and Morrissey, Oliver, 2012. Selectivity on Aid Modality: Determinants of Budget Support from Multilateral Donors. *Review of International Organizations* 7(3): 267-84.
- Dang, Hai anh, Knack, Stephen, Rogers, F. Halsey, forthcoming. "International Aid and Financial Crises in Donor Countries." *European Journal of Political Economy*.
- De Renzio, Paolo, Andrews, Matt and Zach Mills, 2010. Evaluation of Donor Support to Public Financial Management (PFM) in Developing Countries: Analytical Study of Quantitative Cross-Country Evidence. Overseas Development Institute, London.
- Elmers, Bodo, 2011. *How to Spend It: Smart Procurement for More Effective Aid*. Brussels: Eurodad (European Network on Debt and Development).
- Frot, Emmanuel, Perrotta, Maria, 2011. Aid Effectiveness: New Instruments, New Results? Stockholm Institute of Transition Economics. Working Paper No. 11.
- IEG (Independent Evaluation Group), 2011. *World Bank Progress in Harmonization and Alignment in Low-Income Countries: An Evaluation*. Washington, DC: Independent Evaluation Group, the World Bank.
- IEG (Independent Evaluation Group), 2008. *Public Sector Reform: What Works and Why*. Washington, DC: Independent Evaluation Group, the World Bank.
- Jansen, Erik G., 2009. Does Aid Work? Reflections on a Natural Resources Programme in Tanzania. U4 Issue 2009:2. U4 Anti-Corruption Resource Centre Issue, Bergen.
- Knack, Stephen, 2013. Aid and Donor Trust in Recipient Country Systems. *Journal of Development Economics* 101: 316-29.
- Knack, Stephen, Rogers, F. Halsey, Eubank, Nicholas, 2011. Aid Quality and Donor Rankings. *World Development* 39(19), 1907-17.

- OECD, 2012. Aid to developing countries falls because of global recession.
<http://www.oecd.org/dac/stats/developmentaidtodevelopingcountriesfallsbecauseofglobalrecession.htm>.
- OECD, 2011a. *Aid Effectiveness 2005-10: Progress in Implementing the Paris Declaration*. OECD, Paris.
- OECD, 2011b. *Supporting Capacity Development in PFM: A Practitioner's Guide. Task Force on Public Financial Management*. OECD, Paris.
- OECD, 2009a. *Aid Effectiveness: A Progress Report on Implementing the Paris Declaration*. OECD, Paris.
- OECD, 2009b. *Managing Development Resources: The Use of Country Systems in Public Financial Management*. OECD, Paris.
- OECD, 2008. *2008 Survey on Monitoring the Paris Declaration: Effective Aid by 2010?* OECD, Paris.
- PEFA, 2011. *Monitoring Report 2010: An Analysis of Repeat Assessments Including Changes in PFM Systems Performance Measured by Means of PEFA Indicators*. PEFA Secretariat, World Bank, Washington DC.
- Wood, Bernard, 2011. *The Evaluation of the Paris Declaration Phase 2: Final Report*. Danish Institute for International Studies, Copenhagen.
- Wooldridge, Jeffrey, 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge MA.
- World Bank, 2006. *Integrating Global Partnership Programs With Country-Led National Programs*. Background Document for the International Workshop on Global Programs and the Paris Agenda. World Bank, Washington DC.

Table 1
Basic tests

Equation	1.1	1.2	1.3	1.4
Model variation	Bivariate	IDA only	baseline	IDA only
PFM Quality	12.600*** (4.85)	17.583*** (6.64)	7.530*** (3.22)	10.915*** (4.33)
Civil liberties in R			0.545*** (4.30)	0.605*** (4.51)
Log of R per capita Income			-2.174 (-1.11)	-0.141 (-0.06)
Income growth rate in R			0.682* (1.75)	0.467 (1.27)
Log of R population			7.801*** (4.29)	7.778*** (4.05)
Log of aid per capita in R			8.634*** (2.85)	7.856** (2.41)
D share of aid in R (%)			0.771*** (3.94)	0.618*** (2.63)
Constant	-10.066 (-1.17)	-25.072 (-2.96)		
No. of observations	2136	1717	2089	1670
No. of Donors, Recipients	34, 78	34, 59	34, 76	34, 57
R ²	.05	.08	.31	.36

Dependent variable is percentage of donor aid to government managed by recipient government PFM systems. Equations 1.3-1.4 include donor-year fixed effects. T-statistics, reported in parentheses below point estimates, are based on standard errors adjusted for non-independence of errors within both donor and recipient clusters of observations, with *** p<0.01, ** p<0.05, * p<0.1.

Table 2
Robustness to correction for sample selection and omitted variables

Equation	2.1	2.2	2.3	2.4
Model variation	Heckman selection	Results, Strategy	Transparency	Fragile states
PFM Quality	7.337*** (3.26)	6.600*** (2.67)	5.517** (2.15)	5.782*** (3.02)
Civil liberties in R	0.517*** (4.05)	0.607*** (5.14)	0.300 (1.62)	0.485*** (4.19)
Log of R per capita Income	-1.396 (-0.66)	-3.239* (-1.69)	-2.809 (-1.44)	-2.951 (-1.49)
Income growth rate in R	0.820** (2.04)	0.598 (1.30)	0.557 (1.45)	0.472 (1.14)
Log of R population	6.805*** (3.57)	6.196*** (3.42)	7.267*** (4.24)	7.829*** (4.50)
Log of aid per capita in R	7.676*** (2.94)	6.806** (2.38)	7.915*** (2.79)	8.998*** (3.29)
D share of aid in R (%)	0.786*** (3.97)	0.748*** (3.98)	0.757*** (3.84)	0.786*** (3.98)
Inverse Mills ratio	-4.109 (-0.85)			
Results		4.532** (2.12)		
Strategy (score=A)		14.699*** (4.64)		
Transparency			7.977** (2.01)	
Fragile state dummy				-10.001** (-1.96)
No. of observations	2089	1903	2089	2089
No. of Donors, Recipients	34, 76	34, 75	34, 76	34, 76
R ²	.31	.34	.32	0.32

Dependent variable is percentage of donor aid to government managed by recipient government's public financial management systems. All equations include donor-year fixed effects. T-statistics, reported in parentheses below point estimates, are based on standard errors adjusted for non-independence of errors within both donor and recipient clusters of observations, with *** p<0.01, ** p<0.05, * p<0.1.

Table 3
Robustness to sample and specification changes

Equation	3.1	3.2	3.3	3.4
Model variation	Direct budget support	WB obs. dropped	Ordinal PFM quality	D-R fixed effects
PFM Quality	5.456*** (2.87)	7.701*** (3.31)		5.359** (1.98)
Civil liberties in R	0.388*** (3.30)	0.533 (4.42)	0.778*** (4.34)	0.803 (1.64)
Log of R per capita Income	-1.042 (-0.58)	-2.058 (-1.03)	-1.976 (-1.06)	20.416 (0.91)
Income growth rate in R	0.468 (1.38)	0.577 (1.57)	0.656* (1.73)	0.521 (1.35)
Log of R population	6.799*** (4.23)	7.993*** (4.28)	7.257*** (3.91)	69.647 (1.49)
Log of aid per capita in R	6.570*** (2.87)	8.964*** (3.12)	7.661*** (2.57)	8.872** (2.30)
D share of aid in R (%)	0.537*** (2.70)	0.904*** (4.14)	0.771*** (3.92)	0.532*** (2.62)
Direct budget support	0.542*** (11.13)			
PFM quality = 2.5			7.754 (1.07)	
PFM quality = 3			13.438** (2.49)	
PFM quality = 3.5			19.354*** (3.48)	
PFM quality = 4			21.458*** (3.58)	
PFM quality = 4.5+			25.788*** (3.91)	
Year=2007				-0.553 (-0.21)
Year=2010				2.577 (0.57)
No. of observations	2058	1944	2089	2089
No. of Donors, Recipients	34, 76	34, 76	34, 76	34, 76
R ²	0.41	0.32	0.31	0.78

Dependent variable is percentage of donor aid to government managed by recipient government's public financial management systems. Equations 3.1-3.3 include donor-year fixed effects and equation 3.4 includes donor-recipient fixed effects. T-statistics, reported in parentheses below point estimates, are based on standard errors adjusted for non-independence of errors within both donor and recipient clusters of observations, with *** p<0.01, ** p<0.05, * p<0.1.

Table 4
PEFA indicators and Use of procurement systems

Equation	4.1	4.2	4.3	4.4
Model variation	PEFA PFM quality	PEFA and CPIA	Use of procurement	PEFA procurement
PFM Quality		8.743*** (2.78)	6.861*** (2.320)	
Civil liberties in R	0.778*** (4.34)	0.678*** (4.25)	0.484*** (3.51)	0.644*** (3.86)
Log of R per capita Income	-3.203 (-1.59)	-2.855 (-1.44)	-1.461 (-0.75)	-1.688 (-0.76)
Income growth rate in R	1.596** (2.52)	1.749*** (2.85)	0.274 (0.66)	1.713*** (2.85)
Log of R population	7.373*** (3.74)	6.662*** (3.69)	5.927*** (2.90)	2.746 (1.17)
Log of aid per capita in R	8.725*** (2.65)	8.464*** (2.67)	6.631** (1.96)	4.726 (1.18)
D share of aid in R (%)	0.642*** (3.13)	0.663*** (3.20)	0.808*** (4.09)	0.555** (2.30)
PEFA PFM quality	1.592 (0.58)	-5.29 (-1.60)		
PEFA procurement score				2.990*** (3.74)
PFM quality x Donor aid share				
PFM quality x Policy selectivity				
No. of observations	1235	1219	2091	1140
No. of Donors, Recipients	34, 66	34, 65	34, 76	34, 64
R ²	0.35	0.36	0.30	0.34

Dependent variable is percentage of donor aid to government managed by recipient government's public financial management systems in equations 4.1 and 4.2, and percentage of donor aid to government managed by recipient government's procurement systems in equations 4.3 and 4.4. All regressions include donor-year fixed effects. T-statistics, reported in parentheses below point estimates, are based on standard errors adjusted for non-independence of errors within both donor and recipient clusters of observations, with *** p<0.01, ** p<0.05, * p<0.1.

Table 5
Endogeneity of PFM Quality

Equation	5.1	5.2
	PFM	PFM
Model variation	Quality x aid share	Quality x selectivity
PFM Quality	7.528*** (3.23)	7.572*** (3.16)
Civil liberties in R	0.546*** (4.29)	0.571*** (4.37)
Log of R per capita Income	-2.168 (-1.11)	-2.127 (-1.06)
Income growth rate in R	0.684* (1.76)	0.721* (1.80)
Log of R population	7.815*** (4.29)	8.170*** (4.34)
Log of aid per capita in R	8.662*** (3.00)	9.088*** (3.08)
D share of aid in R (%)	0.767*** (3.81)	0.752*** (3.80)
PFM Quality x Donor aid share	-3.205 (-0.19)	
PFM Quality x Policy selectivity		0.763*** (2.89)
No. of observations	2089	2028
No. of Donors, Recipients	34, 76	33, 76
R ²	0.31	0.30

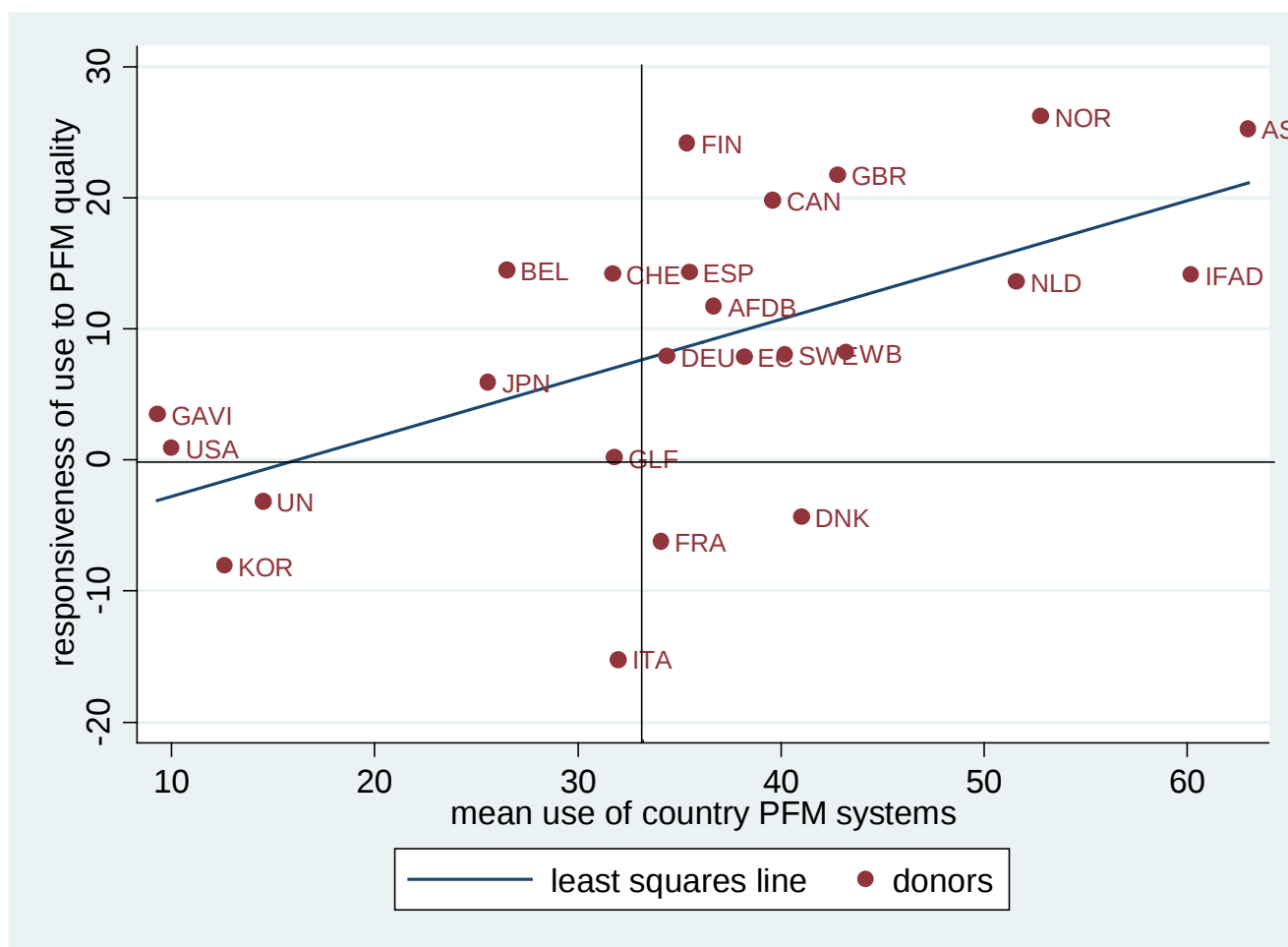
Dependent variable is percentage of donor aid to government managed by recipient government's public financial management systems. Regressions include donor-year fixed effects. T-statistics, reported in parentheses below point estimates, are based on standard errors adjusted for non-independence of errors within both donor and recipient clusters of observations, with *** p<0.01, ** p<0.05, * p<0.1.

Table 6
Donor-specific regressions

Donor	PFM Quality	Civil liberties	2007	2010	R ²	Mean UCS
Norway	26.22**	0.13	4.1	4.7	.41	52.8
Asian Devp. Bank	25.29	0.30	-12.1	29.7	.36	63.0
Finland	24.23***	-0.04	13.7	-4.8	.67	35.4
UK	21.76**	0.85*	-8.3	-8.9	.56	42.8
Canada	19.85**	1.63***	8.4	12.7	.44	39.6
Belgium	14.50*	0.51	-0.4	-7.1	.28	26.5
Spain	14.35	-0.10	19.7*	38.7***	.20	35.5
Switzerland	14.20*	0.39	-16.5	-34.1***	.42	31.7
IFAD	14.15	0.19	-0.1	11.0	.18	60.2
Netherlands	13.61**	1.79***	-5.0	-1.3	.47	51.6
African Devp. Bank	11.78	0.02	15.3	14.1	.20	36.7
World Bank	8.25	0.76**	10.5*	23.4***	.24	43.2
Sweden	8.04	1.73*	5.1	7.9	.48	40.2
Germany	7.95	0.28	3.1	18.3***	.15	34.4
EC	7.89	0.67**	-5.3	13.1*	.18	38.2
Japan	5.97	0.13	9.3*	17.7***	.40	25.6
GAVI	3.49*	0.01	-1.45	-28.9***	.84	9.3
USA	0.90	-0.07	-2.5	-4.7	.07	10.0
Global Fund	0.23	0.63*	-2.4	3.3	.17	31.8
Denmark	-4.37	0.11	27.2***	23.0**	.46	41.0
UN	-3.12	0.64***	-2.9	2.5	.14	14.5
France	-6.22	0.93**	3.1	5.8	.26	34.1
Korea	-8.07	0.55	-33.7	-36.2	.44	12.6
Italy	-15.21	0.88	-14.9	6.6	.21	32.0

Dependent variable is percentage of donor aid to government managed by recipient government's public financial management systems. Regressions are based on equation 1.3 but only coefficient estimates for selected independent variables are shown, with *** p<0.01, ** p<0.05, * p<0.1. Standard errors are corrected for clustering by recipient countries. Donors are ordered from largest to smallest PFM quality coefficient estimates. Only donors with more than 30 observations are included.

Figure 1
Mean use of PFM systems and responsiveness to their quality



Appendix

Paris Declaration Monitoring Survey Indicator Definitions

Use of country PFM systems (PD indicator 5a): Simple average of following three sub-indicators, as a percentage of all aid for the public sector.

Use of national budget execution procedures

Three of these four criteria must be met to qualify:

- 1) Funds are included in the annual budget approved by the legislature;
- 2) Funds are subject to established country procedures for authorization, approval and payment of funds;
- 3) Funds are deposited and disbursed through the established treasury system;
- 4) Opening of separate bank accounts for donor funds is not required

Use of national financial reporting procedures

Both criteria must be met to qualify:

- 1) No separate accounting system is required to satisfy donor's reporting needs;
- 2) No separate chart of accounts is required to record the use of donor funds

Use of national auditing procedures

To qualify, funds are subject to audits by the country's Supreme Audit Institution using its auditing cycle and standards, and additional auditing arrangements are not requested in normal circumstances.

Use of national procurement procedures

Donors do not make additional or special requirements on governments for the procurement of works, good and services.

Direct budget support

To qualify, funds must be transferred to the national treasury and be managed using national budgetary procedures, and not earmarked for specific uses. They may be nominally earmarked for a broadly-defined sector such as education (sector budget support).

PFM Quality (CPIA question 13)

This criterion assesses the extent to which there is: (a) a comprehensive and credible budget, linked to policy priorities; (b) effective financial management systems to ensure that the budget is implemented as intended in a controlled and predictable way; and (c) timely and accurate accounting and fiscal reporting, including timely audit of public accounts and effective arrangements for follow up.

Transparency (CPIA question 16)

This criterion assesses the extent to which the executive, legislators, and other high-level officials can be held accountable for their use of funds, administrative decisions, and results obtained. Accountability is generally enhanced by transparency in decision-making, access to relevant and timely information, public and media scrutiny, and by institutional checks (e.g., inspector general, ombudsman, or independent audit) on the authority of the chief executive. The criterion covers four dimensions: (a) the accountability of the executive and other top officials to effective oversight institutions; (b) access of civil society to timely and reliable information on public affairs and public policies, including fiscal information (on public expenditures, revenues, and large contract awards); (c) state capture by narrow vested interests; and (d) integrity in the management of public resources, including aid and natural resource revenues.

Detailed criteria with descriptions of each ratings level are available at <http://www.worldbank.org/ida/papers/CPIAcriteria2011final.pdf>.

Table A1
Summary statistics

	mean	Std. dev.	Min.	Max.	N
Use of PFM systems (%)	33.5	34.9	0	100	2089
PFM quality	3.47	0.60	2	5.5	2089
Civil liberties in R	32.0	10.5	6	53	2089
D share of aid in R (%)	7.04	8.98	0.001	83.6	2089
Log of R per capita income	6.30	0.90	4.48	8.37	2089
Income growth rate in R	2.95	2.73	-7.43	9.88	2089
Log of R population	16.55	1.37	11.55	19.30	2089
Log of aid per capita in R	3.80	0.84	1.46	6.07	2089
D share of aid in R (%)	7.04	8.98	0.001	83.6	2089
Strategy (score=A)	0.02	0.14	0	1	1903
Results	2.96	0.63	1	4	1903
Transparency	2.98	0.55	1.5	5.0	2089
Direct budget support	12.46	22.79	0	1	2058
Fragile state dummy	0.11	0.31	0	1	2089
PEFA PFM quality	0.00	0.52	-1.29	1.50	1235
PEFA procurement quality	3.66	1.50	1	7	1140

Table A2
selection regressions (probit)

Equation	A2.1	A2.2	A2.3	A2.4
Years	All	2005	2007	2010
Log of R population	0.034*** (5.17)	0.022 (3.68)	0.040*** (4.73)	0.040*** (4.79)
Colonial tie	0.175*** (8.41)	0.107*** (6.04)	0.186*** (7.57)	0.221*** (7.19)
Distance from donor HQ city	-0.031*** (-2.31)	-0.013 (-1.16)	-0.038*** (-2.73)	-0.045** (-2.36)
HIPC completion	0.262*** (8.61)	0.314*** (8.02)	0.274*** (8.11)	0.221*** (5.96)
HIPC decision point	0.115** (2.55)	0.087** (1.96)	0.213*** (3.72)	0.006 (0.09)
Year=2007	-0.040*** (-3.65)			
Year=2005	-0.087*** (-6.25)			
No. of observations	15402	5134	5134	5134
No. of Donors, Recipients	34, 151	34, 151	34, 151	34, 151
Pseudo R ²	0.18	0.23	0.20	0.11

Dependent variable is selection, coded 1 for donor-recipient pairs if positive aid flows are recorded in the PDMS and 0 otherwise. Coefficients are marginal effects, evaluated at the means of all other independent variables. T-statistics, reported in parentheses below point estimates, are based on standard errors adjusted for non-independence of errors within recipient clusters of observations, with *** p<0.01, ** p<0.05, * p<0.1.