

Aid Quality and Donor Rankings

Stephen Knack
F. Halsey Rogers
Nicholas Eubank

The World Bank
Development Research Group
Human Development and Public Services Team
May 2010



Abstract

This paper offers new measures of aid quality covering 38 bilateral and multilateral donors, as well as new insights about the robustness and usefulness of such measures. The 2005 Paris Declaration on Aid Effectiveness and the follow-up 2008 Accra Agenda for Action have focused attention on common donor practices that reduce the development impact of aid. Using 18 underlying indicators that capture these practices—derived from the OECD-DAC's Survey for Monitoring the Paris Declaration, the new AidData database, and the DAC aid tables—the authors construct an overall aid quality index and four coherently defined sub-indexes on aid selectivity, alignment, harmonization, and specialization. Compared with earlier indicators used in donor rankings, this indicator set is more comprehensive and representative of the range of donor practices addressed in

the Paris Declaration, improving the validity, reliability, and robustness of rankings. One of the innovations is to increase the validity of the aid quality indicators by adjusting for recipient characteristics, donor aid volumes, and other factors. Despite these improvements in data and methodology, the authors caution against overinterpretation on overall indexes such as these. Alternative plausible assumptions regarding weights or the inclusion of additional indicators can still produce marked shifts in the ranking of some donors, so that small differences in overall rankings are not meaningful. Moreover, because the performance of some donors varies considerably across the four sub-indexes, these sub-indexes may be more useful than the overall index in identifying donors' relative strengths and weaknesses.

This paper—a product of the Human Development and Public Services Team, Development Research Group—is part of a larger effort in the department to determine how aid can more effectively promote development. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The authors may be contacted at sknack@worldbank.org and hrogers@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.

Aid Quality and Donor Rankings

Stephen Knack, F. Halsey Rogers, and Nicholas Eubank¹

¹ Knack, Rogers, and Eubank are, respectively: Lead Economist, Development Research Group, World Bank; Senior Economist, Development Research Group, World Bank; and PhD student, Graduate School of Business, Stanford University. Corresponding author: sknack@worldbank.org. We are grateful to the World Bank's Knowledge for Change Program (KCP) for funding the valuable research assistance from Afroza Chowdhury and Lodewijk Smets, to Ryan Powers for clarifying technical issues regarding AidData, and to David Roodman, Mark Sundberg, and Adam Wagstaff for providing very helpful comments. This paper was originally prepared for the March 2010 Oxford Conference on "Aid Transparency and Donor Finance". The findings, interpretations, and conclusions expressed in this chapter are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Introduction

How should aid quality be evaluated and compared across donors? A number of recent papers have proposed rankings of bilateral and multilateral donors, on the assumption that donor performance can be measured and ranked globally in a meaningful and robust way (Mosley 1985; Roodman 2006 and 2009; Center for Global Development 2007; Easterly and Pfutze 2008). These public rankings are intended to inspire both bilateral and multilateral donors to improve their effort, effectiveness, and efficiency in ways that will move them up in the rankings. And there is evidence that donors do in fact pay attention to these rankings and care about public perceptions (see, for example, Ramankutty et al. 2009). A sustained and focused “peer pressure” campaign within the DAC appears to have contributed to a marked decline in the share of aid that is tied to purchases of good and services in donor countries (OECD 2009a).²

Whether this approach is successful or not is likely to depend on the perceived validity and robustness of the rankings. While the ranking of aid agencies on well-defined individual components is hard to argue with, this paper investigates the robustness of these overall rankings to inclusion of new indicators and to different weightings of the components. If the various indicators of donor performance that have been put forward are not highly correlated with each other, then the rankings will be sensitive to decisions on how to weight them in the index. Indeed, an example from Easterly and Pfutze (2009) shows how weightings of different components of the index – in this case, policy selectivity vs. poverty selectivity – can lead to an institution being ranked differently:

[The Global Environment Facility] has a comparatively low share of its ODA going to corrupt or non-free countries, but at the same time spends relatively little on the poorest countries. As we freely admit in the paper, *one has to trade off one aspect of selectivity against the other* [emphasis added] . . .

In assessing aid quality by donors, this paper extends and enlarges upon the work of Roodman (2006 and 2009) and Easterly and Pfutze (2008) in several respects. First, it includes a more comprehensive set of indicators that have been used to assess donor performance, combining existing indicators with new ones that reflect the content of the Paris Declaration in a more comprehensive and representative way. In particular, we draw heavily on the OECD-DAC Survey on Monitoring the Paris Declaration (SMPD), a rich new source of indicators on the use of country systems, aid predictability, and coordination with other donors in country (OECD 2008). We also include an indicator on bilaterals' share of aid funneled through multilateral donors. One advantage of including more indicators is that an overall ranking of donors should be less sensitive to the inclusion of (or weight assigned to) any one indicator.

Second, most of our aid quality indicators adjust for important factors that are not directly under the control of the donor agencies, since one goal of the ranking is to assess how effectively the aid agencies are managing. For example, from a management perspective, it is pointless to castigate the EBRD for not focusing on the world's poorest countries, given that the institution is allowed to operate only in the largely middle-income Europe and Central Asia region. So for our ranking exercise, we measure poverty selectivity using only the set of countries in which the EBRD is mandated to work. In effect, we are therefore assessing whether within its mandated region, the

² This approach has also been tried at the country level; in Rwanda, for example, donors participate in a Donor Performance Assessment Framework that rates their progress toward certain targets (Rwanda Development Partners 2009).

EBRD is targeting its efforts at poorer countries. The methodological sections below describe in greater detail how and why we adjusted for constraints on the donor agencies.

Third, in computing indicators of sectoral fragmentation and project proliferation, we take advantage of the new AidData database produced by researchers at the College of William and Mary and Brigham Young University. AidData is based on the CRS data from the OECD-DAC used previously for this purpose by Easterly and Pfutze (2008) and Roodman (2006 and 2009), but includes data on additional donors, recipients and projects obtained directly from donor agencies.

Finally, we provide a database of our aid quality indicators for 2007, including the adjusted indicators and the indexes and sub-indexes we construct. This database will be updated with indicators for 2009, when data from the third round of the Paris Declaration monitoring survey become available. Note that our sample of donors is similar to the Easterly-Pfutze sample, in that it includes all DAC bilaterals and the most significant multilaterals and global funds. We exclude any donors missing data on more than half of the indicators or on more than one sub-index; those excluded tend to be newer, smaller, and non-DAC donors. For consistency with the Paris Declaration monitoring survey—the source of 8 of our 18 indicators—we treat the UN as a single donor, aggregating data over its various agencies.

The results from this exercise in collecting and combining aid quality indicators are both encouraging and instructive. First, we show that it is possible to group these indicators into logical sub-indexes that capture key aspects of aid quality: selectivity, alignment with country systems, harmonization in country, and specialization. Each of these aspects has been a focus of efforts to improve quality, and we develop sub-indexes of indicators that are sufficiently highly correlated with each other that they appear to be measuring a common underlying aspect of quality. Donor rankings within these sub-indexes should be useful in and of themselves, by helping donors to identify areas of relative strength and weakness.

Second, we offer a new overall ranking of donor quality, constructed from these sub-indexes. This ranking has the advantage of incorporating more aspects of quality than other rankings, so that it is likely to be a fairer representation of overall donor quality.

Third, we show that weightings do matter to the rankings: because donor rankings vary among the different sub-indexes, the ultimate overall ranking will be sensitive to the weightings of those sub-indexes. We argue that ranking exercises need to take into account this sensitivity to weightings and should make their weighting choices explicit, to allow the reader to decide whether he or she agrees with them. In the interest of transparency, we aggregate our sub-indexes several different ways and show how the rankings change with each reweighting.

Ranking donors: Challenges, motivations, and methods

Any effort to rank donors on aid quality faces a number of challenges. First, most indicators of donor performance are based on plausible but largely untested beliefs about best practices in aid management. Policy selectivity has received the most attention in the literature (World Bank 1998; Burnside and Dollar 2000). Nevertheless, the evidence linking the quality of policies to the effectiveness of aid has been disputed by some studies (Hansen and Tarp 2000; 2001; Easterly, Levine, and Roodman 2003; 2004), while others have questioned the criteria for assessing selectivity (Amprou, Guillaumont, and Guillaumont Jeanneney 2007). In the case of other aid quality indicators, such as donor use of country systems, there is an even stronger consensus in the donor community on their importance, despite an even weaker base of empirical evidence. Despite some

recent progress, the conclusion reached by Klein and Harford (2005) several years ago remain largely true:

Clearly, it is difficult to settle on a convincing measure of aid quality without better theories and evidence about what makes aid effective . . . [T]he caveats are as important as the conclusions, because the indexes are based either on narrow empirical research or on no research at all. One consequence is that indexes that would intuitively be expected to show similar things produce very different messages. The real value of this embryonic work on aid quality is to highlight the importance of the issue and the scale of our ignorance. Hopefully, future measures of aid quality and donor performance will be able to draw on rigorous measures of aid effectiveness.

In selecting plausible indicators of aid quality, we follow the approach of Easterly and Pfutze (2008), who write:

What should an ideal aid agency look like? The academic aid policy literature and the aid agencies themselves agree on many elements of “best practice” . . . By taking this consensus as our standard, we are asking in effect if aid agencies operate the way they themselves say they should operate.

Second, the rankings implicitly assume that not only do these indicators measure aid characteristics that are associated with better outcomes, but that this relationship is monotonic over the relevant range. That is, either they assume that more of the characteristic is always better (for example, harmonization or selectivity), or at least that all donors are currently operating to the left of the optimum and are unlikely to exceed it in the near future. In making this assumption, we follow the existing ranking efforts.

Third, some of the indicators included in the rankings ignore interactions among donors, which could be important. For example, if most donors “herded” toward the poorest countries, it would improve those donors’ poverty-selectivity ratings, but would not necessarily improve outcomes at the margin. Even the stronger advocates of aid’s effectiveness tend to assume that the marginal effect of aid on growth diminishes with aid and eventually may turn negative. In this situation, the marginal donor could have a greater impact on poverty if it avoided the stampede and instead allocated its aid to countries that are slightly less poor.

This paper largely sets those issues to one side and asks instead the following question: Even if we assume that the individual measured donor behaviors are important for development, does it follow that one can aggregate them into a single index of donor quality that is valid and reliable? Or are the component measures sufficiently uncorrelated that any overall ranking of donors on aid quality will be sensitive to subjective judgments regarding weighting or grouping into sub-indexes?

In ranking donors, it is important to match the methodology to the purpose. Donor-ranking exercises to date have been “name and shame” exercises: their goal has been to motivate lagging donor agencies to improve the quantity and quality of the aid they deliver by benchmarking donors against each other. Both Roodman/CGD and the Easterly and Pfutze ranking exercises have made this aim explicit, although their level of focus is different. Easterly and Pfutze see themselves as describing the “behavior of aid agencies”, in the areas of transparency, fragmentation, selectivity, use of ineffective aid channels, and overhead costs. They are clearly aiming to motivate agencies to improve on these measures, as for example in the case of transparency:

If, as a result of our paper, aid agencies become more transparent and start making quality data available for future research, our expectations will have been exceeded. Some day it might then even become possible to rank agencies based on outcomes (Easterly and Pfutze 2009).

The Roodman/CGD exercise is also intended to motivate change, in areas such as aid-tying, project proliferation/fragmentation, and aid quantities. The aid rankings are published as part of the annual Commitment to Development Index, and the CDI website explains that

[w]e want to inspire a race to the top, so "winners" should be proud of their achievements. Yet there is room for improvement in all rich countries. Almost all countries score below average in at least one area . . . (Center for Global Development 2010)

As with the Easterly and Pfutze rankings, an aid agency could improve its ranking on the CGD's "naming and shaming" lists by changing its practices, so the ranking exercise could meet the goal of having aid agencies improve their behavior.

What does this goal imply for methodologies? To assess how well aid agencies are delivering on their mandate, we believe it is important to control econometrically for factors that are outside the agencies' control. Above, we noted the example of the EBRD with its limited geographic mandate. Another example is the African Development Bank, which would inevitably rank poorly on policy selectivity if we calculated it for a global sample, rather than for the African sample of countries to which it is limited by its mandate. By adjusting for limited mandates, we are able to ask the more relevant question for aid agency management: given its mandate, is the agency reducing poverty as effectively as possible, for example by targeting based on policy and poverty within its assigned geographic area? For other indicators, we adjust for donors' total aid budgets; for example, larger donors will typically "proliferate" their aid across more countries and sectors, other things being equal, and should not be penalized for this. And in measuring use of country systems, we adjust for differences in risk among donors' country portfolios, to avoid unduly penalizing donors that provide a large share of their aid to recipients that have weaker public financial management systems.³ All of these adjustments enter into our calculation of the core ranking below.

Note that other ranking goals could imply different methodological choices. For example, a ranking could be devised to help donor-country governments determine how best to allocate their aid resources. Should aid funds be channeled through the country's bilateral agency, or through multilateral agencies and development banks? If through multilaterals, then which ones? For this second goal, not all of our adjustments would be appropriate. If the goal is to allocate the marginal aid dollar most effectively, the donor country's decision should take into account any constraints that affect that goal. For example, if the donor country wants to direct its marginal resources toward poor countries, then it may not want to channel those resources through an agency that is constrained by mandate to work only in a middle-income region—even if that agency scores highly on poverty selectivity *within* the region. We therefore present below an alternative "allocation index" that omits these adjustments and that would therefore be more useful for such allocation decisions.

Data and methodology: Constructing the sub-indexes

As discussed above, one of the main contributions of this paper is to base donor rankings on a more comprehensive set of indicators of aid quality than have been used in the past. This section

³ The policy selectivity indicator already penalizes donors once for aiding countries with weaker public management.

describes how we have interpreted these indicators, adjusted them where appropriate to enhance validity, and aggregated them into several sub-indexes.

In constructing our index, we identified four distinct dimensions of aid quality: selectivity, alignment, harmonization, and specialization. Measurement theory suggests that indicators measuring similar aspects of quality should be grouped together, which is one reason for constructing these sub-indexes. The four dimensions that we chose emerge from the aid effectiveness literature and from international agreements on aid, but we find that they also hold up empirically. When we categorized indicators using common-sense assignment to the four sub-indexes, we almost inevitably found that the intra- and inter-sub-index correlations confirmed those assignments. In the Results section below, we discuss why these sub-indexes may be useful in their own right, even beyond their utility in building the overall index of aid quality.

Aid selectivity

Aid is widely believed to have greater development impact where it is needed most – that is, where there are large numbers of poor people – and where the policy and institutional environment is favorable to growth and development. The intuition for this belief is difficult to dispute, even if the empirical results in Burnside and Dollar (2000) have not proven very robust. For this reason, both the CGD and Easterly and Pfutze rankings incorporate measures of policy and poverty selectivity in their rankings, and Dollar and Levin (2006) focus solely on these variables in assessing likely trends in aid effectiveness.

For our aid selectivity variables, we compute measures of *policy and poverty selectivity* using the method of Dollar and Levin (2006). Specifically, we regress the log of aid from donor i to recipient j on three variables for the recipient j : the log of its population, its GDP per capita (adjusted for purchasing power parity), and its overall score on the World Bank's annual "Country Policy and Institutional Assessment" (CPIA). The CPIA consists of 16 indicators “representing the different policy and institutional dimensions of an effective poverty reduction and growth strategy.”⁴ These 16 indicators, formulated by Bank staff with input from an external panel of prominent development specialists, are grouped into 4 "clusters": macro and fiscal policy, structural policies, social policies, and public sector management (including corruption). The CPIA ratings are produced for the purpose of allocating the World Bank's concessionary (IDA) lending.

This selectivity regression is run separately for each donor. Policy selectivity is measured simply by the partial regression coefficient on the CPIA score, while poverty selectivity is measured by the partial regression coefficient for per capita income (multiplied by -1 so that more positive values indicate better targeting of aid to low-income recipients).⁵

Aid disbursements used in these regressions are for 2007, and the three independent variables are measured with a one-year lag, to reflect more accurately the information donors possess when they make their allocations. For most donors, the sample of recipients used in the regressions includes 125 countries. For multilaterals with region-specific mandates, we restrict the sample to eligible countries.⁶

⁴ See <http://siteresources.worldbank.org/IDA/Resources/73153-1181752621336/CPIA2008questionnaire.pdf>.

⁵ “Poverty” selectivity is measured using GDP per capita as a proxy, because it is available for most aid recipients on an annual basis. Actual poverty rates are estimated from surveys that are not conducted annually, if at all, in most aid recipients.

⁶ The African, Asian, Inter-American and European regional banks respectively have 50, 33, 23, and 18 eligible aid recipients. Nielson's (2010) comparisons across donors of the responsiveness of sectoral aid to indicators of sectoral

Not all aid recipients have data for 2007 on population, GDP, or CPIA scores. Most notably, Iraq and Afghanistan accounted for more than 27% of U.S. ODA in 2007, but are missing data on GDP and CPIA. The U.S. ranks relatively low on policy selectivity regardless, but likely would fall even further if those two recipients were in the sample.

We follow Dollar and Levin (2006) in computing an Aid Selectivity sub-index by equally weighting the policy and poverty elasticity estimates for each country, after standardizing them to mean 0 and standard deviation equal to 1. (These z-scores are listed in Appendix Table A1.) The correlation between policy and poverty selectivity is .53, indicating that donors that are selective on one dimension are also (*ceteris paribus*) selective on the other.

Aid selectivity for all aid—that is, aid summed over all donors—is remarkably stable over time. Our GDP and CPIA elasticities (respectively) for 2007 are -.52 and +1.91, reflecting only marginal improvements over the 2002 estimates of -.49 and +1.76 reported in Dollar and Levin (2006). Because of the difficulties of establishing a robust empirical link between selectivity and recipient-country growth or poverty reduction, there is no way of knowing for sure whether aid in the aggregate is yet selective “enough” in either the poverty or policy dimension to maximize its impact on poverty reduction. The assumption implicit in the various aid-donor ranking exercises is that it is not, since all donors are still rewarded at the margin for increases in selectivity.

As Dollar and Levin (2006) note, the use of the World Bank-constructed CPIA index as the measure of policy quality potentially biases the selectivity rankings in favor of the World Bank’s IDA. They therefore replicate their analysis substituting a combination of the International Country Risk Guide’s (ICRG) “rule of law” indicator and the Freedom House (FH) democracy index for the CPIA. Using this variable, several donors move up or down quite a few places in the rankings, suggesting some sensitivity to the choice of policy variable. Nevertheless, rankings on policy selectivity and overall aid selectivity (respectively) constructed with the ICRG/FH policy variable are correlated with their original rankings at .65 and .77. Moreover, the alternative measures have serious shortcomings: most (like the ICRG) cover only a few of the many development-related policy areas measured in the CPIA, and the few with its breath of policy coverage (like the Economic Freedom of the World Index) do not match its coverage of aid recipient countries.

We can further test for a World Bank bias by using the CPIA ratings produced by the African and Asian Development Banks. Those ratings use the same scale and the same questionnaire as the World Bank’s CPIA, but they reflect the judgments of African and Asian Development Bank staff rather than World Bank staff. We ran the selectivity regression for the World Bank’s allocations to the 73 countries rated by either of those two regional development banks, first using the World Bank’s own CPIA ratings, and second using the AfDB/AsDB ratings. The World Bank scores better on both policy and poverty selectivity using the other development bank’s ratings than it does using its own ratings, suggesting that our use of the World Bank CPIA ratings does not bias our selectivity rankings in favor of the Bank.⁷

need similarly restricts the samples for the regional banks. Dollar and Levin (2006) apparently did not truncate the sample of recipients for regional development banks in this manner.

⁷ In the case of the AfDB, policy selectivity improves marginally and poverty selectivity worsens marginally when we calculate them using the institution’s own ratings for 50 countries instead of the World Bank’s CPIA. The AsDB scores better on both poverty and (especially) policy selectivity using its own ratings in place of the World Bank’s.

Aid alignment

The Paris Declaration commits donors to "base their overall support on partner countries' national development strategies, institutions and procedures". Through the follow-up Accra Agenda, donors pledged to support country ownership of development programs "by respecting countries' priorities, investing in their human resources and institutions, making greater use of their systems to deliver aid, and increasing the predictability of aid flows" (OECD 2009a). These statements reflect a consensus that donors have often delivered aid in ways that undermine country ownership and institutions.

The Paris Declaration specifies several indicators of the extent to which donors' aid is aligned with country policies and systems. These include: use of country public financial management systems, use of country procurement systems, avoiding parallel project implementation units (PIUs), aid predictability, untying aid, and coordinating technical assistance with national development strategies. The OECD-DAC donors have been reporting on whether or not their aid is tied for many years. The other indicators, however, have been measured only in the last few years through the DAC's Survey on Monitoring the Paris Declaration (SMPD), which was designed to assess progress toward the Declaration's numerical targets.

We use data from the 2008 SMPD, which reports data for the calendar year 2007. A sizeable fraction of our aid quality indicators are taken from this survey, so we describe its methods and content in some detail below. For comparability with the survey-derived indicators, the indicators we compute from other sources are also based on 2007 aid data, even where 2008 data are already available.

The SMPD excludes humanitarian aid, debt relief, and administrative costs. It covers primarily ODA, but non-concessionary loans from multilateral development banks are also included. The DAC survey includes only 54 aid recipients, but coverage on the donor side is fairly comprehensive. All DAC donors, bilateral and multilateral, are included, as are vertical funds (such as the Global Fund and the GAVI Alliance) and several non-DAC bilateral donors (such as Poland and Hungary). The various UN agencies are treated as a single donor in the survey, with their data collected and reported by the UNDP. Aid volumes to the 54 recipients as reported in the DAC survey closely approximate the corresponding aid totals in the DAC's official aid database, compiled from an entirely separate reporting system.

Recipient governments decide whether or not to volunteer their country for the survey, which could raise concerns about their representativeness. And in fact the 54 self-selected recipients differ significantly from other aid recipients in several respects: they are larger (in population), more aid-dependent (measured by aid per capita), and poorer (measured by GDP per capita). Countries that have completed the HIPC process for debt relief are also over-represented in the survey. Controlling for these four variables, survey participation is unrelated to political openness, quality of economic policies and governance, colonial heritage, or region.⁸

Despite the non-random (self-)selection of recipients, there is no obvious reason to believe the sample is biased with respect to measuring the performance of any particular donor. However, the survey may capture a slightly larger share of the aid delivered by some donors than by others (for example, by undersampling donors that allocate more of their aid to smaller or less-poor recipients), increasing the random measurement error on performance indicators for the latter group. Among

⁸ Although 63% of aid-recipient countries in Sub-Saharan Africa are in the SMPD, compared to only 26% from other regions, this difference is accounted for by their income, aid, and HIPC status. An Africa dummy, if added to the selection regression, produces a small and insignificant coefficient.

reporting countries, non-concessionary loans from multilateral development banks were sometimes but not always included in the survey (World Bank 2010). We also cannot entirely rule out the possibility that some donors interpreted the SMPD definitions more liberally than others. The 2008 survey guarded against this by substantially strengthening the definitions, guidance, and survey process, based on experiences from the 2006 first-round survey, and by minimizing donor discretion to the extent practicable. These changes did improve the comparability of data reported by donors and by recipients (OECD, 2008; 2007). Nevertheless, donors were able in some cases to exercise their own judgment—for example, in defining whether their aid in particular countries constituted “support to a program-based approach”—and this discretion could affect their ratings on some indicators (World Bank 2010).

The performance of donors on the alignment and harmonization indicators in the SMPD can be influenced systematically by characteristics of the countries they aid (Knack and Eubank 2009). It is likely that all donors may rely less on country systems plagued by corruption or other weaknesses. For example, the SMPD records that in DR Congo, donors collectively channeled 0% of aid through the recipient’s country public financial management (PFM) systems, compared with 79% of aid in Morocco. Donor performance on the unadjusted “use of country systems” indicator could be driven substantially by differences in risk in their country portfolios, rather than differences in how they operate within any given recipient country. Penalizing donors for bypassing country systems without adjusting for recipient country mix therefore penalizes them for operating in risky environments—which in effect means double-counting the policy selectivity indicator.

For this reason, to generate our measures of donor performance, we adjust the SMPD indicators for the characteristics of the countries in which donors are operating. The “use of country systems” and other indicators are measured at the donor-recipient level, so we can control for differences in their country portfolios using recipient fixed effects. Specifically, we regress use of PFM systems (and other SMPD indicators) on a full set of recipient and donor dummies. For these variables, our donor performance indicators are therefore the coefficient estimates on the donor dummies, which capture the donors’ underlying propensity to use country systems.

Table 1 shows how variation in the SMPD indicators is determined jointly by donor and recipient characteristics. Each indicator is regressed on, alternatively, (1) a full set of donor dummy variables, (2) a full set of recipient dummies, and (3) both sets together. Recipient characteristics in some cases matter more than donor characteristics. Recipient dummies alone explain 28 percent of the variation in use of PFM systems, compared to only 21 percent for donor dummies. Adjusting for recipient effects as described above will make little difference in the measured performance of donors, if the recipient and donor dummies are uncorrelated—in other words, if donors do not differ significantly in the degree to which they aid “good” or “bad” recipients in the survey. However, for some indicators—such as use of PFM systems—the dummies in column 1 of Table 1 do explain some of the same variation in the data as the dummies in column 2, as shown by the fact that the R-squared coefficient in column (3) is much less than the sum of the R-squares in columns 1 and 2. It is these indicators for which the adjustment procedure is likely to have a larger impact on donor rankings. On use of PFM systems, the adjustment makes little or no difference in the ranking of many donors: for example, Ireland and the Inter-American Development Bank still rank first and second, and the World Bank climbs only from seventh place to sixth. But the IMF vaults from 16th to 5th, and Greece from 26th to 14th. In the other direction, Denmark drops from 3rd to 8th; Norway from 6th to 12th, Finland from 9th to 17th, and Canada from 10th to 20th.

Using this adjustment procedure, we construct an “alignment with country systems” sub-index of aid quality from the SMPD indicators along five dimensions: (1) use of country public financial management (PFM) systems, (2) use of country procurement systems, (3) avoidance of parallel

project implementation units (PIUs), (4) aid predictability, and (5) coordination of technical assistance with national development strategies. We also include: (6) the Paris indicator of untied aid. This indicator is computed (following standard practice) from the DAC's Creditor Reporting System (OECD 2009b).

Use of country PFM systems is measured as the average of three sub-indicators on the share of donors' aid to the government that uses recipients' budget execution, financial reporting, and auditing systems.⁹ *Use of procurement systems* is similarly computed as a share of donors' aid. *Coordinated technical assistance* is computed as a share of the total (coordinated plus uncoordinated) technical assistance provided by the donor. Donor performance indicators for use of PFM systems, procurement systems, and coordinated technical assistance adjust for differences in recipient characteristics in fixed-effects regressions using the SMPD data on all donor-recipient pairs with positive values of aid. The *number of PIUs* adjusts similarly for recipient fixed effects, but also for donors' aid volumes in each country, as more aid not surprisingly is associated with more PIUs.¹⁰ In constructing the alignment sub-index, all variables are standardized to have a mean of 0 and standard deviation of 1. For PIUs, the standardized value is multiplied by -1, so that larger values indicate better performance as in the case of all other variables in the sub-index.

Aid predictability is measured as the share of aid donors scheduled for disbursement in 2007 that donors report as actually disbursed within the year. In many cases, actual disbursements exceed rather than fall short of planned disbursements. We follow World Bank (2009) in taking the absolute deviation between the two values, expressed as a share of the larger of the two. Finally, unlike the World Bank paper, we adjust this value for recipient fixed effects, to account for any characteristics of the recipient that could induce all donors operating in that country to exceed or fall short of planned disbursements. Note that this indicator measures predictability only over the short term. Kharas (2008) addresses the important question of aid volatility over the medium term and the contribution of aid from some large donors to volatility of aid receipts, but most donors in our sample and the other donor rankings are not included in that study.

Untied aid is measured as the share of aid commitments for 2007 reported to the CRS that is recorded as neither tied nor partially tied to purchases from the donor country (OECD 2009b). Tying aid has been estimated to increase costs of procurement by 15 to 30 percent (Jepma 1991). Many countries do not report the tied status of some of their aid activities. Rather than omitting this aid from the denominator, we follow Easterly (2002) in treating aid with unreported tying status as tied. Such aid is disproportionately technical assistance, which is more often tied than other aid categories. Non-reporting of tying status is likely motivated in part by the desire to avoid showing a high ratio of tied aid (OECD 2009b); thus omitting unreported aid from both numerator and denominator would overstate the untied aid share for donors that do not report fully. Forced to choose between an indicator that overstates the share of untied aid for less-transparent donors or understates it, we prefer the latter. Of the two alternatives, our preferred one also has the benefit of being more highly correlated with our other aid alignment indicators, which improves the internal reliability of the sub-index. It is debatable whether to give the non-bilateral donors credit for

⁹ One might think that aid effectiveness might be enhanced by bypassing the weakest PFM systems. In that case, it could be argued that donors should not be penalized for failing to use PFM systems where they are below a minimum threshold of quality. As it turns out, applying a rule like this does not affect donor scores in our data. Specifically, we replicated the use of country PFM systems indicator after dropping from the SMPD the 8 countries with the lowest-quality PFM systems, as measured by the Paris Declaration indicator on reliability of country systems (the CPIA indicator on PFM quality). The resulting indicator is correlated with the original at .99.

¹⁰ Specifically, we regressed the log of (the number of PIUs + 1) on log of donor aid in the country, and on donor and recipient dummies. Donor aid volumes are highly significant (at .001 level), and the number of PIUs is estimated to increase by about 10% with a doubling of aid.

untying all of their aid. They are unable to tie it, and were in fact established to avoid such problems as tied aid. Untying all of its aid is an achievement for the U.K., for example, and the untying reveals more about its aid performance. Therefore, in constructing the alignment sub-index and an overall index of donor performance, we code untied aid as missing data for the multilateral donors.¹¹

Scores on the sub-index for aid alignment—and donor rankings—produced by aggregating these six indicators are reported in Table 2. The mean inter-item correlation among the six standardized sub-index components is .27, producing an alpha coefficient of internal reliability for the sub-index of .68.

Adjusting indicators for recipient effects is intended to improve their validity as indicators of aid quality by donor. It will not necessarily improve reliability, however, in the sense of increasing correlations among the several indicators in the aid alignment sub-index. If the adjustments remove random measurement error that is uncorrelated across indicators, then internal reliability of the index should increase. But if the adjustments instead primarily have the effect of removing a common source of bias in the indicators, reliability could decline. In this case, it turns out that improved validity does not come at the expense of reliability: mean inter-item correlations and the alpha coefficient either increase or are unchanged when the four sub-indexes are constructed using adjusted instead of unadjusted indicators.

Harmonization in country

The Paris Declaration commits donors to implement common arrangements and simplify procedures, with the goal of reducing transactions costs imposed by donors on recipient governments. It includes three indicators of harmonization of donor activities around country-led programs, measured through the SMPD. The first indicator is the share of aid delivered through *program-based approaches (PBAs)*, including not only direct budget support but also project aid, if it is part of a government-led sector-wide approach (SWAp). This harmonization indicator obviously overlaps conceptually with the issue of alignment of aid under country-owned development programs, and all direct budget support qualifies as aid using PBAs and as aid using country PFM and procurement systems.¹²

The two other harmonization indicators in the Paris Declaration and measured through the SMPD are the share of *missions* and *country analytic studies* that are coordinated with other donors. Scores on the sub-index for aid harmonization generated by aggregating these three indicators are reported in Table 2. The mean inter-item correlation among the three standardized sub-index components is .30, producing an alpha coefficient of internal reliability for the sub-index of .56.

Specialization

It is widely believed that most donors fragment their aid across too many countries, sectors, and projects (World Bank 2003). A proliferation of donors and projects overstretches and undermines the capacity of recipient governments to manage aid and their relations with donors (Knack and Rahman 2007).¹³ Donors' reluctance to specialize also reduces their average level of expertise in the countries and sectors in which they operate, which may reduce aid's effectiveness even in recipients

¹¹ In one variant of our overall index mentioned above, the "Allocation Index," untied aid is coded as 100% for the multilaterals.

¹² Direct budget support was debated for inclusion as a Paris Declaration indicator, but it was opposed by the US.

¹³ Djankov, Montalvo, and Reynal-Querol (2009) find that in recipients with aid fragmented among more donors, corruption tends to be higher and aid tends to be less beneficial to growth.

that do not suffer themselves from a proliferation of donors and projects. Our specialization sub-index is thus conceptually distinct from our sub-index of in-country harmonization practices.

The Paris Declaration calls for a more effective division of labor among donors, declaring that “[e]xcessive fragmentation of aid at global, country or sector level impairs aid effectiveness.” Donors commit to delegating authority to lead donors where appropriate on the basis of comparative advantage. No indicators are specified in the declaration, however, nor are any included in the SMPD. We construct several indicators for the year 2007 from the DAC’s databases, following the methods of Acharya, Fuzzo de Lima, and Moore (2006), Knack and Rahman (2007), Easterly and Pfütze (2008), and Roodman (2006 and 2009).

First, we calculate the *geographic concentration of aid* for each donor’s 2007 ODA disbursements after subtracting out debt relief, humanitarian aid, and administrative costs. This indicator is calculated as a Herfindahl index, equal to the sum of the squared shares of the donor’s aid going to each of its recipients. Larger donors and those with more expansive geographic mandates tend to have less concentrated aid, probably for justifiable reasons as they are more likely to have sufficient expertise in many countries (and sectors). Geographic concentration is correlated with (the log of) ODA at .49, and using the unadjusted indicator as a performance measure would arguably reward small donors unduly simply for being small.¹⁴ For our donor performance indicator, we therefore take the residuals from a regression on two variables: (1) the log of donor total ODA and (2) the proportion of all aid recipients covered by the donor’s geographic mandate. For most donors this proportion equals one, but for the regional development banks (AFDB, ASDB, IDB, and EBRD), it is less than one.

We similarly compute a Herfindahl-type measure of *aid concentration by sector* for each donor using commitments data for 2007 from AidData.¹⁵ A donor with lower concentration, by this measure, does not necessarily contribute more to fragmentation measured from the recipient standpoint, because it may be aiding many small countries that receive little or no aid from other donors. However, Acharya, Fuzzo de Lima, and Moore (2006) show that any such cases are the exception and not the rule. Donors that fragment their aid across recipients the most are significantly more likely to be suppliers of ODA to recipients receiving aid from many donors (and hence more fragmented aid, from the recipient’s perspective).

AidData incorporates the DAC’s CRS database, used by Easterly and Pfüte in calculating sectoral fragmentation, but adds information from other donor sources. For our indicator of donor performance indicator we take the residuals from a regression of sectoral concentration on two variables: (1) the log of total 2007 aid commitments by the donor reported in AidData, and (2) a dummy for donors with a limited sectoral mandate (namely, the global or “vertical” funds). The resulting indicator is positively but modestly correlated (at .23) with the performance indicator on geographic concentration.

These concentration indexes capture problems associated with fragmentation very imperfectly, for at least two reasons. First, their values typically decline only slightly if a donor adds recipients or sectors to its portfolio but allocates them only small fractions of its total aid. For example, if a donor with \$1 billion in annual disbursements gave 20 new recipients \$0.5 million each, it would hardly affect that donor’s concentration index, but it could substantially increase the administrative costs for the donor and the recipient countries. Second, a given value for sectoral concentration

¹⁴ The specialization index in Easterly and Pfüte (2008) ranks Greece, Portugal, and other small donors as top performers, because it does not adjust sectoral and geographic fragmentation for donor size.

¹⁵ Sectors are aggregated to the two-digit level for this purpose. For example, we treat “Health” as a single sector, rather than treating medical research, health education, health infrastructure etc. as separate sectors.

calculated at the global level for a donor can be consistent with very different patterns at the recipient level, with very different implications for aid quality. Suppose Donor A's aid portfolio consists of \$1 million in education aid to each of 10 recipients and \$1 million in health aid to each of 10 other recipients, while Donor B's consists of \$0.5 million in education and \$0.5 million in health aid for each of 20 recipients. Sectoral concentration for each donor is 0.5—even though Donor A assists only 1 sector in each recipient while Donor B aids two sectors in each of its recipients, and Donor B is therefore contributing more to the problems of proliferation of projects and administrative burdens on recipient governments.

To address these shortcomings with the concentration measure, we therefore computed from the PLAID dataset a donor-level indicator of the *average number of sectors aided* per recipient receiving aid. We then took as a performance indicator the residuals from a regression of the log of this average on the log of total aid commitments by the donor and on a dummy for donors with a limited sectoral mandate. A similar indicator was generated based on the *number of recipients aided* by each donor in 2007 (using the DAC disbursements data), adjusting for donors' aid volumes and geographic mandates. The number of sectors and countries aided are correlated at .48 before removing the common effect of aid volumes; even after adjusting for aid volumes and limited mandates, they are correlated at .37.

A fifth indicator on aid specialization is based on the *average size of projects* in AidData, using 2007 data on commitments.¹⁶ Following Roodman (2006 and 2009) and Knack and Rahman (2007), the premise is that, in general, a proliferation of distinct aid projects imposes unnecessary transactions costs on recipients and unduly taxes their administrative capacity. Roodman uses particular assumptions about project production technology to estimate the efficiency losses associated with breaking aid up into an excessive number of projects. For example, he assumes that the optimal size of projects for a given aid volume is larger in better-governed countries because the marginal benefit from oversight is lower. We are able to take a more flexible approach with the data, because in contrast to Roodman we do not need to generate plausible estimates of efficiency losses to apply a discount to aid volumes. We simply regress the size of project i (in log of current US dollars) on a full set of donor, recipient, and sector dummies and on (the log of) total aid by the donor to the recipient for projects other than project i .¹⁷ If all donors tend to break aid up into smaller pieces in certain recipient countries, whether because of poor governance or other factors, the recipient dummies will control for any such effects. Similarly, the sector dummies will control for any systematic differences in donors' involvement in activities that have a larger optimal scale. The donor dummy coefficients then reflect donor-specific effects not attributable to either the characteristics of the recipients or sectors in which they happen to be operating.

We find that project size is significantly and positively correlated with the sum of other aid provided by the donor to the recipient. Despite netting out this scale effect, the donor dummy coefficients are still subject to scale effects at the global level. Specifically, even when controlling for donor's aid volumes in the country where the project is located, projects are significantly larger for donors with larger global aid budgets. We therefore make one further adjustment to our project size indicator,

¹⁶ We treat Ireland as missing data on project size, because its records in CRS (which are incorporated in AidData) correspond to expenditure items rather than to projects or programs. There are undoubtedly other significant discrepancies in the way donors report their aid activities in the CRS, but this appears to be the most extreme case of non-comparability in data reporting. The average size of Ireland's "projects" in the CRS is smaller than for any other donor, so its score on the specialization sub-index (already its worst among the four sub-indexes) would decline if this indicator were not coded as missing data.

¹⁷ Following Kilby (2010), we exclude project i to avoid endogeneity in testing for scale effects. Sector dummies are based on the more detailed 5-digit sector codes in AidData, so average differences between (say) health infrastructure and health education programs can be captured.

regressing the donor dummy coefficient estimates on the log of that donor's total aid commitments for 2007 in AidData, and taking the residuals as a donor performance indicator for average project size. Our resulting project size indicator is correlated with Roodman's "size weight" indicator for 2007 at .73.

The Paris Declaration calls for donors to delegate aid efforts to other donors in some countries and sectors more often. Neither the SMPD nor any other survey systematically collects data on delegation of this sort, but from a broader perspective, delegating aid delivery to multilateral donors can be viewed as consistent with this goal. Channeling aid through multilaterals can reduce the problems associated with aid fragmentation described above. Multilaterals have particular sectoral and geographic mandates, so channeling aid through them will often reduce the number of donor-recipient-sector combinations of existing aid relationships. Multilaterals' particular mandates also tend to give them greater expertise in many countries or sectors, compared to most bilaterals (especially the newer or smaller donors). As a sixth aid specialization indicator, we therefore take *contributions to multilateral donors* (other than the EC¹⁸) as a share of total ODA for each bilateral donor. This indicator is coded as missing data for the multilateral donors themselves.

A final indicator of aid specialization is *administrative costs* as a share of aid, as reported by donors to the DAC. Donors that fragment their aid more excessively by recipient, sector, project and other dimensions are likely to incur higher administrative costs.¹⁹

We construct a fourth sub-index on Aid Specialization from these seven indicators, by standardizing them, reversing the signs where appropriate (for administrative costs, number of recipients aided, and average number of sectors aided), and taking the mean. The mean inter-item correlation among the seven standardized sub-index components is .26, producing an alpha coefficient of internal reliability for the sub-index of .71 (Table 2).

Other indicators

We have chosen not to include among our aid quality indicators several measures proposed by other researchers. Generosity of repayment terms for aid—with grants as the "best" form of aid—is one of several aid quality indicators in Mosley (1985) and in White and Woestman (1994), and Bulow and Rogoff (2005) urge development banks to give aid in the form of grants instead of loans, on transparency and other grounds. We agree, however, with Easterly (2007: 666) that grants "may have other problems such as not requiring the kind of long-run investment that would generate returns to repay loans." Grants versus loans is therefore in our view an overly ambiguous indicator of aid quality. It is more easily viewed as an aid quantity issue, and Roodman (2006 and 2009) simply deals with it by netting out both interest and principal repayments from aid volumes prior to implementing his quality adjustment procedures.

Easterly and Pfütze include food aid and technical assistance as indicators of bad aid practices, largely on grounds that they tend to be tied to purchases from the donor country. They (and we) already include a direct measure of untied aid, so the fact that food aid and technical assistance are often (but not always) tied is a weak rationale for counting them against donors. They note that technical assistance is "often condemned as reflecting donor rather than recipient priorities", but

¹⁸ The EC in important respects is more like a bilateral than a multilateral agency. For example, Martens and others (2002: 192) report that "national representatives in EC foreign aid decision-making committees spend a considerable part of their time exploring and pursuing contract opportunities for national suppliers."

¹⁹ If fixed costs of establishing an aid program are large, administrative costs could be a higher share of ODA for smaller donors. No such relationship is found in the data, however, so we do not adjust the administrative costs indicator for donor size.

that issue is covered separately in our ranking by the SMPD indicator on the share of technical assistance that is aligned with recipients' programs. We exclude both the food aid and technical assistance measures.

Apart from the problems of tying, and weak impact of uncoordinated technical assistance, it is difficult to make the case that donors should be penalized for providing a larger share of their aid in this form. In their model of aid's growth effects, Annen and Kosempel (2009) argue that technical assistance is superior to other forms of aid, because it is much more likely to transfer knowledge. The Paris Declaration and Accra Agenda for Action commit donors to help recipients strengthen their statistical capacity and their public financial management, procurement, and information systems. Donor aid in these areas largely takes the form of technical assistance.

Results: Donor scores and ranks

Table 3 gives the scores and ranks on each of the sub-indexes, as well as the overall score and rank, for each donor. The Appendix provides a summary list mapping the 18 underlying indicators to the relevant sub-indexes, and Appendix Table A1 lists the z-scores for all 18 indicators included in the overall index.

Ranks on sub-indexes

Before discussing the overall rankings, it is worth reviewing the sub-index rankings. By identifying where donors are performing relatively well or poorly, the sub-indexes may provide more “actionable” indexes of performance than the overall rankings, which are based on aggregates of behaviors in very different areas.

First, on our selectivity sub-index, the World Bank ranks first, followed by the Asian Development Bank (ADB), with the Netherlands third overall and top-ranked among bilaterals.²⁰ Our donor rankings for 2007 aid disbursements are broadly consistent with the 2002 rankings given by Dollar and Levin (2006). Good performers in both years include the World Bank, IMF, Denmark, and the Netherlands, while France, the United States, Portugal, and Greece appear near the bottom of both rankings.²¹ Some of the regional development banks are notable exceptions: because of our adjustment procedure, which adjusts for their limited geographic mandates, the Asian and European Development Banks perform markedly better in our selectivity rankings than in those of Dollar and Levin (2006).

On alignment, Ireland is ranked first, followed by the Asian Development Bank, the GAVI Alliance, and Denmark. Japan ranks high (at #8) on this sub-index despite ranking in the lower half on the other three sub-indexes. The UN, US, Austria, Portugal, and Greece are ranked near the bottom, along with a couple of newer bilateral donors. Portugal and Greece have the largest share of aid for which tying status is unreported, so their rankings are potentially sensitive to our decision to treat such aid as tied. If we instead assume that the unreported aid is tied in the same proportion as aid for which tying status is reported, Greece would move up from 32nd to 28th on the alignment sub-index, but Portugal would remain 37th out of the 38 donors ranked. To this point, we have followed

²⁰ It ranks fourth and first respectively on policy- and poverty-responsive allocations. Nielson (2010) finds the World Bank's IDA aid also tends to be the most responsive to need at the sectoral level, e.g. in allocating health aid toward countries with lower life expectancy, energy aid to countries with low electricity consumption, etc.

²¹ Throughout this section, we focus on rankings rather than scores, primarily because the scores are ordinal and do not generally have cardinal interpretations. However, a reader who wants a sense of how agencies are grouped in terms of performance—for example, whether the top-ranked agencies are really far out in front of the others—can look either at the scores in Tables 3 and 4 (and Appendix Table A1) or at Figure 1.

the OECD-DAC convention in treating partially tied aid as tied rather than untied (OECD-DAC, 2009). Partially tied aid differs from fully tied aid in excluding procurement of goods and services from other developed countries, but not from developing countries. Most tied aid for most donors is fully tied, so if we re-classify partially tied aid as untied, rankings on the alignment sub-index are unaffected with one notable exception. All EC aid reported in the CRS is partially tied, so its untied aid changes from 0% (worst of all donors) to 100% under this more generous definition. Its ranking on the sub-index would accordingly improve from 31st to 21st.

On the harmonization sub-index, Ireland is again top-ranked, followed by Denmark and GAVI. The UK, Australia, and (especially) the UN rank higher on this sub-index than on the other three. The ADB ranks only 14th, despite ranking first or second on the other three sub-indexes. The bottom nine spots are held by smaller bilaterals plus Japan (35th).

On the specialization sub-index, five of the top seven donors are multilaterals, while the other two are new, small bilaterals (Poland and Hungary). Specialization is a relative weakness for three bilaterals that rank high on the other three sub-indexes: Denmark (14th), the Netherlands (18th) and Ireland (23rd). It is a relative strength for several large donors: the EC (10th), France (12th) and the U.S. (17th). Several new, smaller bilaterals (Greece, Turkey, Korea, and the Czech Republic) rank at the bottom of this sub-index, along with the EBRD.

Overall rankings and sensitivity to alternative weightings

The “overall” score in Table 3 is calculated as the unweighted average of the four sub-indexes, after standardizing them.²² With this weighting, the top five donors in terms of overall rank are the Asian Development Bank (ADB), the World Bank, Denmark, Ireland, and GAVI, while the bottom five are Portugal, the Czech Republic, Korea, Greece and Austria. Large donors that earn relatively low overall scores include the United States (ranked 32) and the United Nations agencies (ranked 30).

Bilateral donors (including the European Commission) that are members of the OECD-DAC account for 22 of the 38 donors in our rankings. The mean overall index score for this group is -.08, significantly higher ($p=.06$) than the mean of -.63 for the six non-DAC bilaterals (including Korea, a member only since January 1, 2010). The mean for the six multilateral development banks, including the IMF, is .78, significantly higher ($p=.01$) than the mean for the DAC bilaterals. As a group, the four remaining donors (UN, IFAD, Global Fund, GAVI) do not differ significantly from the DAC bilaterals.

The final column in Table 3 provides a summary indicator of the variability in each donor’s ranking across the four sub-indexes. Specifically, we take the difference (in absolute value) between a donor’s ranking on one sub-index and on each of the other three, and compute the mean of these rank differences over all six pairwise comparisons of sub-indexes. There are cases of low variability among both high-ranking (for example, the World Bank and IFAD) and low-ranking (for example, Greece) donors. Cases of high variability across sub-index rankings are accounted for mostly by newer and smaller bilaterals such as Hungary and Poland, but also some multilaterals, including the AfDB, EBRD, and UN. The most obvious interpretation of this variability indicator is that some donor agencies indeed perform differently across four distinct aspects of aid quality. Alternatively, the indicator could be viewed as a measure of relative uncertainty or imprecision in measuring the broader underlying concept of aid quality for each donor; in this interpretation, there is more uncertainty regarding the “true” ranking of donors that have higher variability across sub-indexes.

²² For the one donor that is missing data on one sub-index (the GAVI Alliance, which has no selectivity rating), the score is calculated as the unweighted average of the other three sub-indexes.

One potential concern is that rankings could be overly sensitive to findings from a single data source, the SMPD, as it provides all three indicators in the sub-index on harmonization and 5 of the 6 indicators in the alignment sub-index. However, the untied aid indicator is correlated almost as strongly (at .60) as the 5 SMPD indicators are on average (.63) with the alignment sub-index scores. Moreover, the two SMPD-dominated sub-indexes are not as strongly correlated (at .77) with the overall index scores as the selectivity sub-index is (at .83). The single indicator most strongly correlated with the overall index is not one of the 8 Paris Survey indicators, but rather poverty selectivity (at .82). As mentioned above, we cannot entirely rule out the possibility that some donors may have interpreted the survey guidance more liberally than others to improve their apparent performance, but if so, any resulting bias is unlikely to affect the overall ranking very much.

In Table 4, we further investigate the sensitivity of the overall rankings to alternative assumptions about indicator weightings or regarding the purpose of an overall index. The first rankings column replicates the “Overall” rankings from Table 3, while the next three columns use other weighting schemes. Index_18 and Index_20 dispense with the sub-indexes altogether. Rather than averaging sub-index scores, they average the individual indicators. Index_18 uses the 18 indicators included in our sub-indexes, while Index_20 adds two more—the Transparency and Overhead indicators provided by Easterly and Pfutze (2008)—that we do not include in our sub-indexes or in the Overall index constructed from them. Relative to the latter index, Index_18 and Index_20 give greater weight to aid quality dimensions for which we have more indicators, namely Specialization and Alignment. While these changes in methodology and content affect the rankings, the general contours remain the same: the top five donors were all previously in the top 8, while all of the bottom five were previously in the bottom eight. Among bilaterals, only two smaller donors (Hungary and Luxembourg) have ranks that move more than six spots as a result of the reweightings. The EBRD is the only multilateral that moves much: shifting to Index_20 from Overall drops it by 10 spots, because it ranks poorly on specialization (which has an effective weight of 7/20 in Index_20) and on the transparency index in Easterly and Pfutze (2009).

The fourth column ranks donors on what we call the Allocation Index. As discussed above, the Allocation Index is not an alternative version of the overall index (which could be used for “naming and shaming”), but an index aimed at a different goal: helping to guide allocations of aid resources. In calculating it, we reverse several adjustments made in the first three indexes for factors that were beyond the control of the aid agencies themselves (most notably, geographic mandate). We also code untied aid at 100 percent for the multilateral donors, because the fact that none of their aid is tied is a relevant criterion for allocating resources between a bilateral agency, the EC, and a multilateral. Despite the different in purpose, the Allocation Index rankings do not differ much from the Overall index rankings. The only donors that move more than 5 spots in the rankings are two regional banks, the IDB and the EBRD. This result is not surprising, given that both banks serve largely middle-income regions, which reduces their poverty selectivity when we do not adjust for geographical mandate. Among the sub-indexes, the selectivity sub-index is the most affected, but even it has a rank correlation of 0.84 with the original sub-index; for the other three sub-indexes, the correlations range from 0.94 to 1.0. In summary, moving to an Allocation Index approach does not dramatically affect the rankings, but it does advance several bilaterals past two of the regional development banks in the rankings.²³

²³ Some readers may be uncomfortable with comparing bilaterals and multilaterals (both development banks and vertical funds) on the same ranking, given the inherent differences in their structures and mandates. From a rankings perspective, it is likely that each group has advantages in different areas (such as untied aid for multilaterals and specialization for bilaterals and vertical funds), and it is not clear which would have the greater net advantage. But a reader who prefers to compare only within donor groups can interpret the rankings in that way, by comparing bilaterals against only other bilaterals and multilaterals against only other multilaterals, with or without vertical funds included.

The indicators can be adjusted or aggregated in innumerable ways, some of which may be equally as valid or intuitive as those we present in Table 4. As one example, indicators or sub-indexes could be weighted based on component or factor loadings from principal components or factor analysis. Our explorations with this approach show that it would tend to reduce the weight of our specialization sub-index and its component indicators; it would therefore slightly improve the overall rankings of donors that score poorly on those indicators, relative to donors who do well on them.²⁴ However, the broad outlines of the scores and rankings would not change much under a principal-components or factor-analysis approach.

Comparisons with other indexes of aid quality and quantity

Table 4 also compares our overall index rankings with two previous aid quality indexes, from Easterly and Pfutze (2008) and Roodman (2006 and 2009). As part of the Center for Global Development's annual "Commitment to Development Index" (CDI), Roodman constructs measures of quality-adjusted aid volumes, using plausible but strong assumptions concerning how aid selectivity, aid tying, and project proliferation reduce the effective value of aid.²⁵ Although the final product of this exercise used in the CDI is (quality-adjusted) aid quantities, Roodman produces an implicit aid quality index, equal to his quality-adjusted aid volumes as a fraction of his "net aid" or pre-quality-adjusted aid volumes. The aid data used in the 2009 CDI are from the year 2007, so are comparable to ours.

Easterly and Pfutze (2008) rank donors on an index of "best practices in aid." Their sub-indexes on selectivity (based on poverty, corruption and democracy) and specialization (geographic and sectoral) are fairly similar to two of ours, although their specialization index is constructed to favor smaller donors. A third sub-index on "ineffective aid channels" penalizes donors for providing tied aid, food aid, and technical assistance. The rationale for categorizing food aid and technical assistance as ineffective is mostly that they are tied to purchases from the donor country, so it would be more accurate to call it simply an aid tying index. Their final two sub-indexes are based on newly collected data on overhead costs and on transparency, as measured by providing readily accessible information on operating costs and by full reporting of aid data to the OECD-DAC. While transparency and overhead costs are not unimportant issues, especially for aid research, they have not been central to the aid effectiveness debates and are largely absent from the Paris Declaration. Moreover, Easterly and Pfutze acknowledge that there are serious difficulties in measuring these indicators accurately.

While our various overall indexes produce very similar rankings, they differ considerably from those of Easterly-Pfutze and Roodman. Our main "Overall" index is correlated with Roodman's at .60, and with Easterly-Pfutze's at only .25.²⁶ (See Figure 2.) Of the donors ranked in the top 5 by our Overall index, only the World Bank appears in the top five on both of their rankings. The Asian

²⁴ We do not rely on principal components (or factor) analysis to identify the number or content of sub-indexes, because any donor with missing data on any of the 18 indicators (i.e. nearly half of them) is dropped by that type of analysis. In a principal components analysis on this limited sub-sample, the selectivity, alignment, and harmonization indicators all load most heavily on the first component, and the specialization indicators load most heavily on the second component.

²⁵ In addition to its aid component, the CGD index also ranks donor countries on trade, migration and other policies affecting developing countries that are not applicable to multilateral agencies. Nevertheless, Roodman constructs aid quality measures for the multilaterals, and credits their results to bilateral donors in proportion with the contributions received from each bilateral donor country.

²⁶ Remarkably, our overall index is correlated at .56 with Mosley's (1985) aid quality index for 14 bilateral donors using data from 1979 on poverty selectivity, untied aid, grant element, and share of aid going to agriculture and social infrastructure. Correlations of Mosley's index with our selectivity sub-index (.67) and our poverty selectivity indicator (.76) are particularly high.

Development Bank nearly meets that criterion, but the other three rank poorly on Easterly-Pfütze. At the bottom of the rankings, the CGD rankings line up reasonably well with ours, but there are substantial discrepancies with Easterly-Pfütze: among large donors, the US, Japan, and France all rank at least 15 slots better in their index than in ours. Those three donors all rank fairly high on the Easterly-Pfütze transparency and overhead cost indexes, which we omit from our core index.

The modest correlation of our index with Roodman's, and weak correlation of ours with Easterly-Pfütze, does not by itself suggest that ours is a more valid index of aid quality. It is notable, however, that Roodman and Easterly-Pfütze are even more weakly correlated (at .19) with each other. One explanation for this low correlation is that Easterly and Pfütze are measuring aspects of donor performance that are go beyond the usual aid-quality dimensions; the large weight that they give to their new transparency and overhead cost indexes is likely to lead to weak correlations between their index and other reasonable rankings.

In comparison with the alternatives, our index better meets the criteria of "content validity" (Carmines and Zeller 1979). Namely, the content of our 18 indicators reflects a much more comprehensive and representative sample of "best" aid practices contained in the Paris Declaration and in the research literature on aid effectiveness.²⁷ The adjustments we make to control for factors unrelated to performance (for example, donor size in measuring geographic concentration of aid) also add to the validity of our indicators and indexes produced from them.

Finally, our index and sub-indexes have a much greater degree of internal reliability than those of Easterly-Pfütze. Higher inter-correlations among indicators and among sub-indexes do not prove that they are all good measures of the same underlying and broader concept of aid quality, but low inter-correlations would suggest more doubt. The mean inter-item correlation of our four sub-indexes is a modest .37, but the corresponding value for the five sub-indexes in Easterly-Pfütze is only .14.²⁸ The alpha coefficient of index reliability is .70 for our index and .44 for Easterly-Pfütze. The lowest among the mean inter-item correlations among our four sub-indexes is .26, while two of the five sub-indexes in Easterly-Pfütze have negative mean inter-item correlations.

How is aid quality correlated with quantity of aid delivered, relative to the donor country's GNI (both measured in 2007)? As the bottom right panel of Figure 2 shows, they have a strong positive correlation among the bilaterals, at 0.66. In other words, countries that show a strong commitment to aid quantitatively also deliver that aid in ways that is consistent with the quality standards agreed to by the international community.

Conclusion

In this paper, we have presented a new index of the quality of aid provided by bilateral and multilateral donor agencies. The index is constructed from a more comprehensive set of indicators that cover the four major areas around which the greatest consensus has developed in the aid community: selectivity, alignment, harmonization, and specialization. It combines indicators that have been used in previous indexes with rich new data collected through the recent monitoring survey for the Paris Declaration of the OECD-DAC, as well as the new AidData projects database. Raw indicators were adjusted for the influence of exogenous factors where appropriate, but in fairly simple and intuitive ways that substantially enhance validity at a modest cost in transparency. We also construct sub-indexes for each of the four dimensions of quality noted above, and show

²⁷ Roodman's ability to make use of many more aid quality indicators is constrained by the need in his approach to identify a plausible estimate of the degree to which a "bad" aid practice reduces the effective value of a dollar of aid.

²⁸ Roodman's method does not lend itself to computing measures of reliability.

empirically that the indicators within each sub-index are reasonably well correlated, suggesting that they could be measuring the same underlying phenomenon.

The aid quality of different donors can now be compared along each of these indicators and sub-indexes, as well as on the overall aid quality index. Indeed, as we have argued above, it will often make more sense to use the less aggregated indicators. Combining indicators inevitably raises the question of how to select the appropriate weights, and as we have shown above, the rankings are somewhat sensitive to different weights. The individual indicators and sub-indexes are likely to be more actionable than the overall index because of their greater specificity.²⁹ Individual indicators are also necessary to track progress (in absolute terms) over time, as aggregates built from standardized indicators are designed for comparing relative performance across donors.

At the same time, we believe that an overall index does have its place. Rankings along that single index can bring more pressure to bear on low-ranked donors to move towards “best practices” in delivering aid. In the absence of an overall index, such donors could simply stress the small number of indicators on which they score reasonably well. An overall index prevents this problem.

Although we have tried to be comprehensive in assembling indicators of aid quality as defined implicitly by the Paris Declaration, many aspects of donor agency performance are currently not well measured so were excluded from our index. An example is delegation of aid to another donor; delegation can have the effect of harmonizing and specializing the delegating donor’s activities, but the practice has not yet been measured systematically. A related concern is that our measurement of specialization takes into account its quantity but not its quality: specialization should really take place in areas in which donors have comparative advantage, but we have no data to assess whether this is the case. Other hard-to-quantify areas of donor performance go beyond the content of the Paris Declaration, including the knowledge dimension of aid: the aid agencies’ contributions to our understanding of development problems and solutions. Transparency in data reporting and operating costs (which Easterly and Pftuze attempt to measure) represents only a small part of donors’ information-provision role. For example, some donors conduct and publish systematic evaluations of their projects. Ideally, we would have included measures of the quantity, quality, and transparency of each donor’s impact evaluations, its research, and its country analysis. A related dimension of aid quality, even harder to quantify, is the donor’s willingness to experiment with promising new approaches: innovation is conducive to learning, and some donors may be more innovative and others more risk-averse. Including measures of these dimensions might well change our rankings, which offers another reason to be refrain from over-interpreting small differences in rankings.

Our goal with this paper has been to help improve the measurement of aid quality and donor rankings, with the ultimate objective of improving donor performance. Our intent is not to supplant the other ranking and quality-measurement efforts, which are apparently intending to continue. The CGD/Roodman index serves as an input to the regularly updated CGD Commitment to Development Index, while the Easterly-Pftuze rankings are expected to be updated as well. In addition, a promising separate Brookings/CGD “Quality of Official Development Assistance” measurement effort has recently gotten underway.³⁰ While we do plan to update our

²⁹ White and Woestman (1994) compare DAC bilateral donors on several aspects of aid quality (aid untying, poverty selectivity, and the concessionary element of aid), but do not combine them in a single index because it would obscure the “sources of discrepancy in donor performance.”

³⁰ That project will come out as Birdshall, Kharas, and Mahgoub (Forthcoming). The paper has not been published yet, but based on the conference presentation that is available (from the March 2010 AidData conference at Oxford University), it appears to share a similar motivation to ours. Like our paper, that project emphasizes the importance of assessing donor performance along distinct and well-defined dimensions, and it also uses the Paris Monitoring Indicators

indexes and rankings when the next round of the Paris Monitoring data becomes available, over the longer term we recognize the value of having ranking exercises carried out by independent think tanks and researchers rather than by research staff at a donor agency. We hope that with this paper, we have added some data sources, methodological considerations, and cautions about interpretation that will influence other ranking exercises and—perhaps most importantly—will also inform consumers of the rankings.

as a major source of data. Two differences are that unlike ours, their approach does not appear to make adjustments for factors outside the donor agency's control, nor does it attempt to combine the scores along different dimensions into an overall score and ranking.

Table 1
Explanatory power of donor and recipient dummies for SMPD indicators (2007)

	Donor dummies only	Recipient dummies only	Donor and recipient dummies
Alignment indicators			
Use of PFM systems	.19	.29	.43
Use of procurement systems	.24	.22	.41
Predictability	.21	.14	.34
Coordination of tech. coop.	.15	.18	.32
PIUs*	.31	.09	.39
Harmonization indicators			
Program-based approaches	.14	.23	.36
Coordinated missions	.27	.14	.40
Coordinated country reports	.21	.14	.34

Note: *log of (number of PIUs +1)

Source: Authors' calculations, based on the OECD-DAC Survey on Monitoring the Paris Declaration (2008)

Table 2
Correlations among indicators and sub-indexes (2007)

	Average inter-item correlation	No. of component items	Scale reliability coefficient
Sub-indexes			
Selectivity	0.53	2	0.70
Alignment	0.27	6	0.68
Harmonization	0.30	3	0.56
Specialization	0.26	7	0.71
Overall index (average of 4 standardized sub-indexes)	0.37	4	0.70
Overall index (“unstructured” average of all 18 indicators)	0.15	18	0.76
Overall index (“unstructured” with Easterly-Pfütze <i>Transparency</i> and <i>Overhead</i> indicators)	0.13	20	0.75

Source: Authors' calculations

Table 3
Donor scores and rankings (by sub-index and overall, for 2007)

Donor name	Sub-indices								Overall index		Avg rank difference of sub-indices	No. of indicators with data
	Selectivity		Alignment		Harmonization		Specialization		Score	Rank		
	Score	Rank	Score	Rank	Score	Rank	Score	Rank				
African Dev.Bank	0.505	11	-0.359	28	-0.193	26	1.095	5	0.310	13	14.0	15
Asian Dev.Bank	1.837	2	0.800	2	0.305	14	2.382	1	1.784	1	6.5	15
Australia	0.122	13	-0.042	22	0.708	5	0.253	8	0.348	12	9.3	18
Austria	-0.307	21	-0.654	33	-0.484	31	-0.317	28	-0.641	34	6.5	18
Belgium	0.104	14	0.161	17	-0.191	25	-0.416	29	-0.130	21	8.8	18
Canada	-0.357	23	-0.241	26	-0.416	29	-0.187	21	-0.421	29	4.5	18
Czech Republic	-0.727	28	-0.754	36	-1.150	36	-1.016	38	-1.283	37	5.0	14
Denmark	1.158	5	0.711	4	1.015	2	0.064	14	0.985	3	6.2	18
EBRD	0.960	6	-0.077	23	0.129	18	-0.555	34	0.083	16	14.8	15
EC	-0.612	26	-0.466	31	0.384	13	0.212	10	-0.160	23	12.7	17
Finland	0.035	17	0.447	9	0.574	9	-0.419	30	0.233	15	11.8	18
France	-0.908	33	0.180	15	-0.200	27	0.116	12	-0.226	24	12.5	18
GAVI Alliance	.	.	0.752	3	0.862	3	0.145	11	0.861	5	5.3	9
Germany	-0.452	24	0.407	11	0.205	17	-0.198	22	0.026	17	7.3	18
Global Fund	-0.652	27	0.213	14	-0.094	22	0.019	16	-0.137	22	7.5	15
Greece	-0.828	31	-0.535	32	-0.859	34	-0.602	35	-0.975	35	2.3	18
Hungary	-0.753	30	-1.623	38	-0.604	32	1.492	2	-0.558	31	18.3	17
IDB	0.007	18	0.044	20	0.263	16	0.873	6	0.412	9	7.3	15
IFAD	0.886	7	0.676	5	0.669	6	0.232	9	0.831	7	2.2	15
IMF	1.161	4	0.099	18	0.089	19	1.114	4	0.787	8	9.8	12
Ireland	0.099	15	1.453	1	1.145	1	-0.238	23	0.919	4	13.3	17
Italy	-0.486	25	-0.403	29	-0.159	23	-0.115	20	-0.406	28	4.8	18
Japan	-0.349	22	0.475	8	-0.894	35	-0.278	24	-0.333	27	13.8	18
Korea	-1.128	36	-0.035	21	-1.409	37	-0.768	36	-1.114	36	8.0	15
Luxembourg	0.675	9	-0.427	30	0.552	10	-0.517	33	0.024	18	15.3	18
Netherlands	1.332	3	0.655	6	0.605	8	-0.022	18	0.838	6	7.8	18
New Zealand	0.573	10	0.070	19	-0.459	30	-0.311	27	-0.090	20	11.3	18
Norway	0.380	12	0.387	12	0.280	15	-0.298	25	0.247	14	7.0	18
Poland	-1.098	35	0.161	16	-0.843	33	1.441	3	-0.046	19	18.8	14
Portugal	-1.018	34	-1.562	37	-2.208	38	-0.044	19	-1.718	38	10.0	17
Spain	0.086	16	-0.289	27	-0.066	21	-0.308	26	-0.233	25	6.3	18
Sweden	0.691	8	0.356	13	0.061	20	0.088	13	0.383	11	6.0	18
Switzerland	-0.063	19	-0.122	25	-0.187	24	-0.495	32	-0.317	26	6.7	18
Turkey	-1.447	37	-0.101	24	0.388	12	-0.887	37	-0.640	33	14.7	13
United Kingdom	-0.143	20	0.421	10	0.765	4	0.023	15	0.396	10	8.8	18
United Nations	-0.833	32	-0.700	34	0.621	7	-0.484	31	-0.479	30	13.7	15
United States	-0.751	29	-0.731	35	-0.331	28	-0.004	17	-0.634	32	9.2	18
World Bank	2.301	1	0.600	7	0.522	11	0.623	7	1.291	2	5.0	15

Note: Sub-index scores are normalized in standard deviation units with a mean of 0.

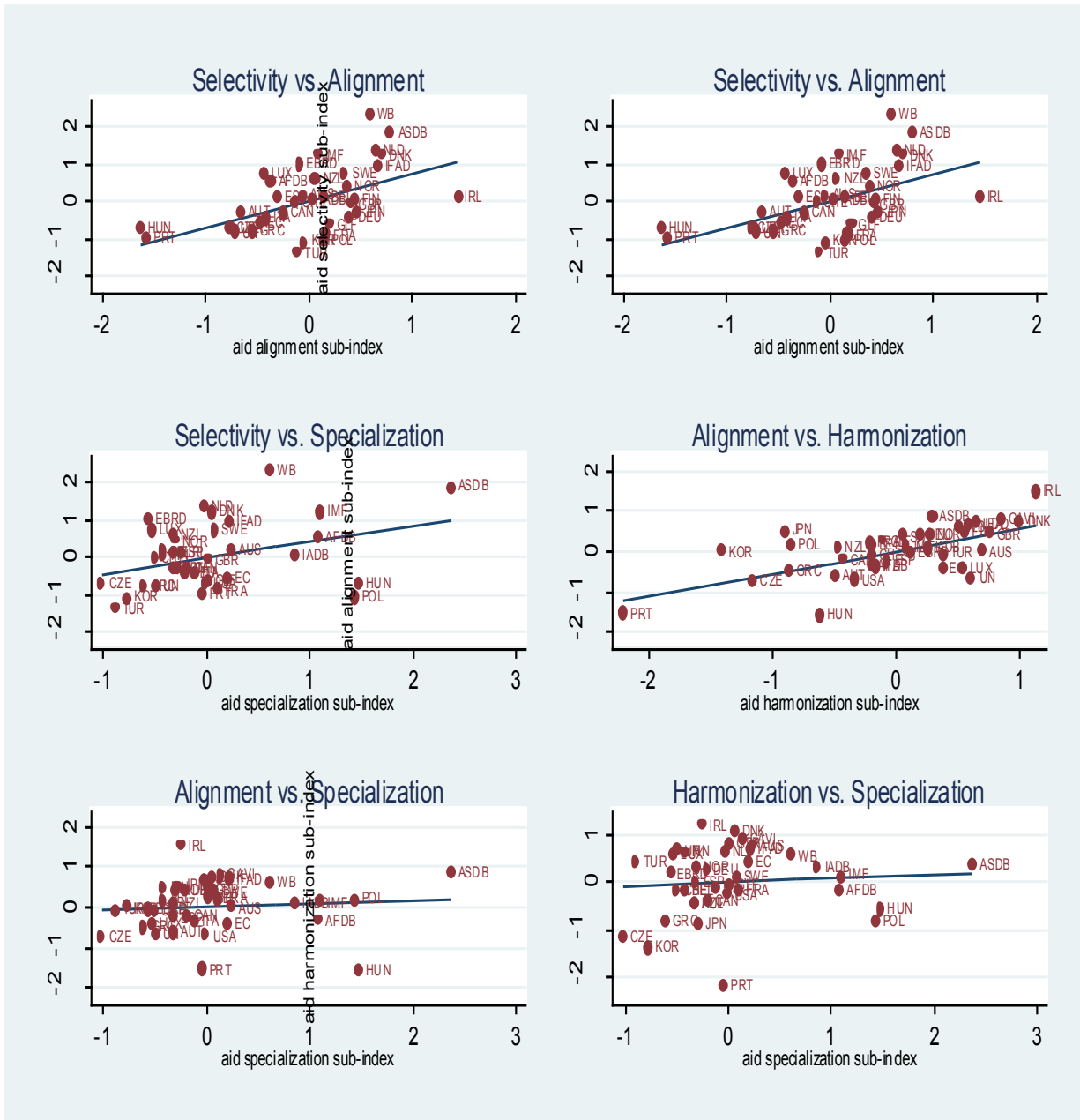
Source: Authors' calculations

Table 4
Donor ranks with different weightings (2007), compared with CGD and EP ranks

Donor name	Code	Rank on overall index	Ranks on alternative summary indexes			Rank given by	
			Index_18	Index_20	Allocation index	CGD/ Roodman	Easterly & Pfutze
African Dev.Bank	AFDB	13	14	14	9	2	3
Asian Dev.Bank	ASDB	1	1	1	2	7	4
Australia	AUS	12	10	10	12	29	12
Austria	AUT	34	33	32	33	33	15
Belgium	BEL	21	23	22	21	19	13
Canada	CAN	29	30	28	30	16	19
Czech Republic	CZE	37	37	38	37	26	.
Denmark	DNK	3	7	5	4	6	19
EBRD	EBRD	16	22	26	22	15	26
EC	EC	23	25	25	23	10	25
Finland	FIN	15	15	16	13	14	21
France	FRA	24	21	18	25	18	9
GAVI Alliance	GAVI	5	3	3	8	.	.
Germany	DEU	17	18	20	19	34	24
Global Fund	GLF	22	19	19	17	8	.
Greece	GRC	35	35	35	35	25	27
Hungary	HUN	31	16	17	31	13	.
IDB	IADB	9	12	12	15	1	4
IFAD	IFAD	7	6	7	3	3	31
IMF	IMF	8	5	8	6	.	23
Ireland	IRL	4	4	4	5	5	15
Italy	ITY	28	28	24	28	35	14
Japan	JPN	27	26	21	27	36	8
Korea	KOR	36	36	36	36	31	.
Luxembourg	LUX	18	24	27	16	9	22
Netherlands	NLD	6	8	6	7	12	18
New Zealand	NZL	20	20	23	20	21	30
Norway	NOR	14	17	15	14	20	6
Poland	POL	19	13	13	18	32	.
Portugal	PRT	38	38	37	38	24	9
Spain	ESP	25	27	29	24	27	29
Sweden	SWE	11	9	11	10	11	7
Switzerland	CHE	26	29	30	26	22	9
Turkey	TUR	33	34	33	32	30	.
United Kingdom	GBR	10	11	9	11	17	2
United Nations	UNTOT	30	31	34	29	28	28
United States	USA	32	32	31	34	23	15
World Bank	WB	2	2	2	1	4	1
Notes:							
"overall index" is the core index, reproduced from Table 3							
"index_18" equally weights all 18 indicators included in any of the 4 sub-indexes							
"index_20" equally weights all 18 indicators included in equal_weight_18 index plus Easterly-Pfutze <i>Transparency & Overhead</i> sub-indexes							
"allocation index" is derived from overall index, but does not adjust indicators for factors outside the donor's control							
Roodman index varies from 0-1, computed as quality-adjusted aid divided by net aid. Easterly-Pfutze varies from 0-100 (average percentile rank of 5 sub-indexes). Other indexes are in standard deviation units with a mean of 0.							

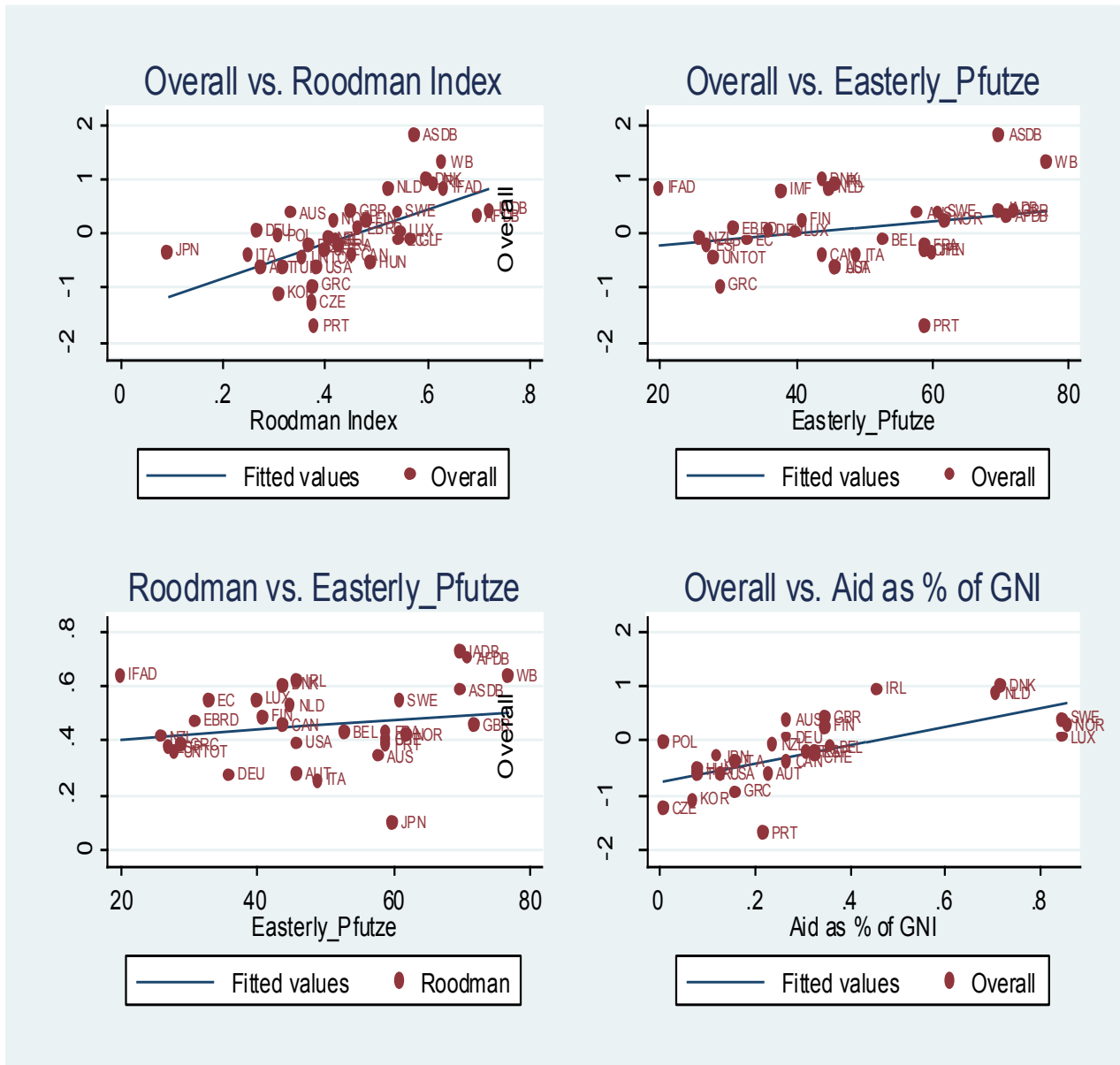
Source: Authors' calculations, based in part on Roodman (2006 and 2009) and Easterly and Pfutze (2008)

Figure 1
Correlations among scores on the four sub-indexes



Source: Authors' calculations

Figure 2
Correlations between our overall index and
alternative indicators of aid quality and quantity



Source: Authors' calculations, based in part on Roodman (2006 and 2009) and Easterly and Pfutze (2008)

Appendix Table A1

Scores for 18 individual indicators (normalized to Z-scores), by donor

Donor	Selectivity		Alignment						Harmonization			Specialization						
	S1	S2	A1	A2	A3	A4	A5	A6	H1	H2	H3	Sp1	Sp2	Sp3	Sp4	Sp5	Sp6	Sp7
AfDB	-0.46	1.47	.	-0.71	0.53	-0.11	-0.87	-0.64	0.06	-0.46	-0.01	.	.	2.34	-0.47	-0.16	0.09	1.72
AsDB	0.89	2.79	.	1.19	1.40	0.15	0.75	0.52	1.39	-0.42	-0.68	.	.	0.02	0.99	0.82	4.84	3.25
Australia	0.23	0.01	0.68	0.56	-0.17	-0.71	-0.95	0.35	0.19	0.28	0.87	-0.39	0.61	-0.81	1.19	0.76	0.80	-0.04
Austria	0.32	-0.93	0.03	-1.23	-0.86	0.01	-1.32	-0.54	-0.63	-1.27	0.08	-0.71	1.14	-0.61	-0.67	-1.00	0.09	-0.13
Belgium	0.25	-0.04	0.68	0.19	-0.17	1.25	-0.78	-0.21	-0.63	-0.81	0.84	-0.64	0.43	-0.54	-0.48	-0.35	-0.61	-0.50
Canada	-0.24	-0.48	-0.12	-0.60	0.10	0.38	-0.84	-0.37	-0.21	-0.54	-0.69	0.31	-0.76	-0.47	-0.17	-1.00	0.13	0.11
Czech Rep.	-0.43	-1.02	.	0.55	-1.38	-0.56	-2.81	0.43	-1.37	-1.27	-1.38	-1.56	0.15	.	-0.70	-0.97	.	.
Denmark	1.60	0.71	0.68	0.64	0.85	1.27	-0.03	0.86	1.38	0.83	1.41	0.43	-0.71	0.82	-0.34	0.34	-0.59	0.70
EBRD	1.66	0.26	.	0.16	-0.71	-0.51	1.00	-0.31	-1.18	-0.86	1.93	.	.	1.43	-1.36	-0.43	-0.85	-1.94
EC	-0.65	-0.58	-2.80	0.97	0.25	-0.51	-0.51	-0.19	0.20	0.27	0.94	.	-0.11	0.26	-0.21	-0.55	-0.22	1.09
Finland	0.42	-0.35	0.54	0.71	0.29	0.78	0.40	-0.03	0.88	0.65	1.13	-0.03	-0.94	0.04	-0.59	-0.50	-0.62	0.14
France	-0.86	-0.96	0.46	-0.10	0.32	0.79	-0.05	-0.35	-0.43	-0.07	0.04	-0.10	0.35	-0.03	-0.13	-0.47	0.08	0.35
GAVI	.	.	.	0.25	0.81	-1.02	1.15	2.56	0.88	3.49	-1.57	.	.	0.14
Germany	-0.35	-0.55	-0.08	0.79	-0.07	0.61	0.65	0.55	-0.21	0.48	0.87	-0.10	1.09	-1.09	-0.24	-0.28	-0.32	-1.10
Global Fund	-1.13	-0.18	.	0.01	0.26	0.16	1.15	-0.52	1.55	-0.31	-0.75	.	.	1.88	-0.55	-0.05	-1.23	-0.07
Greece	-0.63	-1.03	-2.29	-2.15	0.43	0.60	0.38	-0.18	0.62	-1.27	-1.57	-1.46	-1.64	-0.27	0.05	-1.36	0.28	0.58
Hungary	-0.06	-1.45	.	-2.59	-1.91	-2.52	-0.23	-0.85	-1.95	-0.68	-0.40	3.08	2.22	0.77	3.99	2.03	0.15	0.84
IDB	-0.30	0.31	.	0.45	1.50	0.15	-1.91	0.03	0.03	0.37	-0.04	.	.	2.05	0.81	-0.18	-0.44	-0.05
IFAD	0.31	1.46	.	-0.48	1.33	2.00	-0.20	0.74	-0.53	2.07	1.00	.	.	0.84	-1.00	0.55	1.30	0.11
IMF	1.08	1.24	.	-0.51	1.11	-0.44	1.10	-0.77	0.82	-0.19	-0.71	.	.	.	0.28	3.98	.	.
Ireland	-0.38	0.58	0.83	0.57	2.25	2.10	1.01	1.96	2.25	0.46	1.30	-0.21	0.21	.	-0.07	-0.02	-0.35	-0.53

Appendix Table A1 (continued)

Scores for 18 individual indicators (normalized to Z-scores), by donor

Donor	Selectivity		Alignment						Harmonization			Specialization						
	S1	S2	A1	A2	A3	A4	A5	A6	H1	H2	H3	Sp1	Sp2	Sp3	Sp4	Sp5	Sp6	Sp7
Italy	-0.73	-0.25	-0.84	0.11	-0.88	0.31	-0.63	-0.48	0.05	-0.56	-0.12	0.77	0.97	-0.66	-0.26	-0.30	-0.62	-0.57
Japan	-0.07	-0.62	0.10	1.18	-0.33	-0.66	1.25	1.30	-0.16	-1.03	-0.67	-0.08	-0.09	-1.07	0.56	-0.41	-0.02	-1.77
Korea	-1.36	-0.90	.	-0.44	-1.08	-1.17	0.67	1.84	-1.63	-0.10	-1.57	.	.	-0.72	-0.66	-1.04	-0.46	-0.92
Luxembourg	1.13	0.22	0.83	1.01	-1.12	-1.41	-0.30	-1.57	-0.69	0.43	1.23	-0.17	-0.42	-0.38	-0.78	-0.15	-0.65	-0.28
Netherlands	2.07	0.59	0.17	0.33	0.87	1.48	0.81	0.26	1.16	1.23	0.18	-0.15	0.10	-0.01	-0.35	0.32	-0.31	0.22
New Zealand	1.58	-0.43	0.25	0.39	-0.53	-0.42	0.28	0.46	-0.07	0.95	-1.57	-0.28	-1.07	0.01	-0.75	-0.13	0.40	0.54
Norway	0.58	0.18	0.83	-0.11	0.58	0.71	0.81	-0.50	0.43	0.15	0.74	0.30	-0.40	-0.72	-0.50	-0.05	-0.13	-0.68
Poland	-1.21	-0.99	.	2.27	-0.24	-0.56	1.40	-2.06	-0.37	-1.27	-1.57	2.40	2.22	.	2.36	-0.28	.	.
Portugal	-1.15	-0.89	-1.42	-2.24	-2.15	-1.77	0.70	-2.48	-2.78	-1.27	-1.57	-1.02	0.19	-0.55	0.53	1.40	0.15	0.35
Spain	-0.54	0.71	-0.44	-1.24	0.36	0.32	-0.13	-0.60	-0.57	-0.46	-0.20	-0.15	0.81	-0.77	-0.56	-0.05	-0.59	-1.09
Sweden	1.24	0.15	0.64	0.48	0.59	0.22	0.30	-0.10	0.29	0.20	0.23	0.54	-0.72	-0.04	-0.59	-0.13	0.49	1.03
Switzerland	-0.14	0.01	0.83	-0.53	0.02	0.40	-0.52	-0.93	-0.29	-0.42	0.80	0.29	-1.21	-0.59	-0.94	-0.65	-0.27	0.00
Turkey	-2.09	-0.80	.	-1.09	-1.39	0.86	0.66	0.46	-0.15	-0.37	.	-1.20	-1.60	.	-0.06	-0.34	.	.
UK	-0.10	-0.19	0.83	0.11	0.59	0.25	0.26	0.49	0.90	1.51	0.57	0.21	-0.76	-0.54	0.13	-0.14	-0.25	0.92
UN	-1.08	-0.58	.	0.84	-1.02	-1.27	-2.27	0.22	0.18	0.73	0.64	.	.	-1.47	-0.14	-0.94	0.30	-0.83
US	-1.03	-0.47	-0.37	-0.39	-1.32	-1.15	-0.77	-0.38	-0.22	-0.70	-0.23	-0.06	-0.07	-1.24	1.41	-0.09	-0.15	-0.77
World Bank	1.60	3.00	.	0.66	0.88	0.00	0.43	1.03	0.78	0.22	0.49	.	.	1.98	0.26	1.82	-0.43	-0.64

Note: See Appendix list (next page) for interpretation of indicator codes S1 through Sp7.

Appendix

Aid Quality Indicators: Data Sources and Mapping to Sub-indexes

Selectivity

S1: policy (DAC tables and World Bank [CPIA])

S2: poverty (DAC tables and World Development Indicators [GDP per capita, PPP- adjusted])

Alignment

A1: untied tied (CRS)

A2: aid predictability (SMPD)

A3: use of PFM systems (SMPD*)

A4: use of procurement systems (SMPD)

A5: use of PIUs (SMPD)

A6: technical cooperation coordinated with country programs (SMPD)

Harmonization (in country)

H1: use of program-based approaches (SMPD)

H2: coordinated missions as share of all missions (SMPD)

H3: coordinated country analytic reports as share of all reports (SMPD)

Specialization

Sp1: contributions to multilaterals (DAC tables)

Sp2: administrative costs (DAC tables)

Sp3: average project size (PLAID)

Sp4: geographic concentration (DAC tables)

Sp5: number of recipients (DAC tables)

Sp6: sectoral concentration (PLAID & CRS)

Sp7: average number of sectors per recipient (PLAID)

*OECD-DAC Survey on Monitoring the Paris Declaration

References

- Acharya, A., A. T. Fuzzo de Lima and M. Moore (2006). "Proliferation and Fragmentation: Transactions Costs and the Value of Aid." *Journal of Development Studies* 42 1: 1-21.
- Amprou, J., P. Guillaumont and S. Guillaumont Jeanneney (2007). "Aid Selectivity According to Augmented Criteria." *World Economy* 30 5: 733-63.
- Birdsall, N., H. Kharas and A. Mahgoub (Forthcoming). "A Quality of Development Assistance Index." Center for Global Development and the Wolfensohn Center for Development at the Brookings Institution.
- Bulow, J. and K. Rogoff (2005). "Grants versus Loans for Development Banks." *American Economic Review* 95 2: 393-97.
- Burnside, C. and D. Dollar (2000). "Aid, Policies, and Growth." *American Economic Review* 90(4): 847-68.
- Carmines, E. G. and R. A. Zeller (1979). Reliability and Validity Assessment. *Sage University Paper series of Quantitative Applications in the Social Sciences, series no. 07-001*. Beverly Hills and London, Sage Publications.
- Center for Global Development (2007). Measuring the Quality of Development Assistance: Discussion Draft. Washington, DC: Center for Global Development.
- . (2010). "FAQs: Commitment to Development Index 2009." Retrieved March 7, 2010, from <http://www.cgdev.org/section/initiatives/active/cdi/faq>.
- Djankov, S., J. G. Montalvo and M. Reynal-Querol (2009). "Aid with Multiple Personalities." *Journal of Comparative Economics* 37 2: 217-29.
- Dollar, D. and V. Levin (2006). "The Increasing Selectivity of Foreign Aid, 1984-2003." *World Development* 34 12: 2034-46.
- Easterly, W. (2002). Evaluating Aid Performance of Donors. Washington, DC: Center for Global Development. Background paper for the Commitment to Development Index.
- (2007). "Are Aid Agencies Improving?" *Economic Policy* 52: 633-68, 675-78.
- Easterly, W., R. Levine and D. Roodman (2003). New Data, New Doubts: A Comment on Burnside and Dollar's 'Aid, Policies, and Growth' (2000). New York, New York University.
- (2004). "Aid, Policies, and Growth: Comment." *American Economic Review* 94 3: 774-80.
- Easterly, W. and T. Pfutze (2008). "Where Does the Money Go? Best and Worst Practices in Foreign Aid." *Journal of Economic Perspectives* 22 2: 29-52.
- (2009). "Where Does the Money Go? Best and Worst Practices in Foreign Aid: Response." *Journal of Economic Perspectives* 23 1: 245-46.

- Hansen, H. and F. Tarp (2000). "Aid Effectiveness Disputed." *Journal of International Development* 12(3): 375-98.
- (2001). "Aid and Growth Regressions." *Journal of Development Economics* 64(2): 547-70.
- Jepma, C. J. (1991). *The Tying of Aid*. Paris, OECD Development Centre.
- Kharas, H. (2008). *Measuring the Cost of Aid Volatility*. *Wolfensohn Center for Development Working Paper No. 3*.
- Kilby, C. (2010). "What determines the size of aid projects?": Villanova School of Business Economics Working Paper No. 10.
- Klein, M. and T. Harford (2005). *The Market for Aid*. Washington, D.C., International Finance Corporation.
- Knack, S. and N. Eubank (2009). *Aid and trust in country systems*. The World Bank, Policy Research Working Paper Series, 5005.
- Knack, S. and A. Rahman (2007). "Donor Fragmentation and Bureaucratic Quality in Aid Recipients." *Journal of Development Economics* 83 1: 176-97.
- Martens, B., U. Mummert, P. Murrell and P. Seabright (2002). *The institutional economics of foreign aid*. With a foreword by Elinor Ostrom. Cambridge; New York and Melbourne, Cambridge University Press.
- Mosley, P. (1985). "The Political Economy of Foreign Aid: A Model of the Market for a Public Good." *Economic Development and Cultural Change* 33 2: 373-93.
- Nielson, R. (2010). "Does Aid Follow Need? Humanitarian Motives in Aid Allocation." Presented at the AidData Conference, Oxford, UK (March).
- OECD (2008). *2008 Survey on Monitoring the Paris Declaration: Effective Aid by 2010?* Paris, OECD (www.oecd.org/dac/hlfsurvey <<http://www.oecd.org/dac/hlfsurvey>>).
- (2009a). *Aid Effectiveness: A Progress Report on Implementing the Paris Declaration*. Paris, OECD.
- (2009b). *Implementation of the 2001 DAC Recommendation on Untying Aid to the LDCs: 2009 Review*. Background document to the OECD High-Level Meeting of 27-28 May 2009.
- Ramankutty, R., E. Berglof, W. Easterly and T. Pfütze (2009). "Comments." *Journal of Economic Perspectives* 23(1).
- Roodman, D. (2006 and 2009). "An Index of Donor Performance." Center for Global Development Working Paper Number 67. Data set updated 2009.
- Rwanda Development Partners (2009). "Rwanda Donor Performance Assessment Framework: Preliminary results from the 2009 round of monitoring."

White, H. and L. Woestman (1994). "The Quality of Aid: Measuring Trends in Donor Performance." *Development and Change* 25 3: 527-54.

World Bank (1998). *Assessing Aid: What Works, What Doesn't, and Why*. Washington, DC, World Bank.

--- (2003). *World Development Report 2004: Making services work for poor people*. Washington, DC, Oxford University Press for the World Bank.

--- (2010). "Bank Support to Program-Based Approaches: What Does the 2008 Survey on Monitoring the Paris Declaration Tell us?" Washington: World Bank, OPCS Aid Effectiveness Unit.