NUMBER 5 | DECEMBER 2016









BROOKE SHEARER SERIES

Aid effectiveness in fragile states

How bad is it and how can it improve?

LAURENCE CHANDY, BRINA SEIDEL, AND CHRISTINE ZHANG

B | Global Economy and Development at BROOKINGS

BROOKE SHEARER SERIES

This series is dedicated to the memory of Brooke Shearer (1950-2009), a loyal friend of the Brookings Institution and a respected journalist, government official and non-governmental leader. This series focuses on global poverty and development issues related to Brooke Shearer's work, including women's empowerment, reconstruction in Afghanistan, HIV/AIDS education and health in developing countries. Global Economy and Development at Brookings is honored to carry this working paper series in her name.

Laurence Chandy is a fellow in the Global Economy and Development program at the Brookings Institution.

Brina Seidel is a research assistant in the Global Economy and Development program at the Brookings Institution.

Christine Zhang is a former research analyst in the Global Economy and Development program at the Brookings Institution. She is currently a 2016 Knight-Mozilla Open News Fellow based at the *Los Angeles Times* Data Desk.

Acknowledgements: The authors are grateful to Julie Biau and Natasha Ledlie for leading earlier, exploratory research on this topic, and to John McArthur and Homi Kharas for comments on an earlier draft.

Author's Note: The Brookings Institution is a nonprofit organization devoted to independent research and policy solutions. Its mission is to conduct high-quality, independent research and, based on that research, to provide innovative, practical recommendations for policymakers and the public. The conclusions and recommendations of any Brookings publication are solely those of its author(s), and do not reflect the views of the Institution, its management, or its other scholars.

Brookings gratefully acknowledges the program support provided by the Australian Department of Foreign Affairs and Trade, the Bill & Melinda Gates Foundation, the Government of Denmark, and the William and Flora Hewlett Foundation.

Brookings recognizes that the value it provides is in its absolute commitment to quality, independence, and impact. Activities supported by its donors reflect this commitment.

CONTENTS

LIST OF TABLES AND FIGURES

Aid effectiveness in fragile states

How bad is it and how can it improve?

LAURENCE CHANDY, BRINA SEIDEL, AND CHRISTINE ZHANG

ABSTRACT

Fragile states around the world are at risk of being left behind, making them a high priority for increased foreign assistance. But is the donor community capable of delivering that aid effectively? To explore this question, we develop a set of aid effectiveness indicators that are relevant to aid-giving in fragile states and that allow a fair comparison of aid practices in fragile and stable countries. We then make that comparison using different definitions of fragility. Under certain definitions, we show that aid practices in fragile states are inferior to those in stable countries. We then turn the spotlight onto individual donors, revealing substantial variety in their performance in fragile settings. The patterns we find among donors point to workable, though difficult, ways to raise the standard for aid in the countries that need effective aid the most. Chief among these is a call for the worst performing bilateral donors to withdraw themselves from fragile states and support a larger role for multilateral agencies in their place.

I. INTRODUCTION

"I think everyone in the room knows that this is a moment of extraordinary progress. Over the last 30 years, extreme poverty has been cut in half… When USAID was founded more than 50 years ago, donor governments were responsible for nearly three-quarters of all finance flowing to the developing world. Now we account for less than 10 percent."

U.S. Agency for International Development Administrator Gayle Smith, March 2016 [1]

These remarks made by the head of the world's largest bilateral aid agency capture two of the defining trends in global development of the past several years. Together, these trends imply a diminishing role and need for foreign aid.

In fragile states, however, neither of these trends is apparent. The stunning reduction in the number of people worldwide living under the extreme poverty line—from 1.8 billion in 1990 to 800 million in 2013—can be entirely attributed to stable countries. The total number of extremely poor people living in fragile states rose over the same period.[2] Similarly, international capital flows into fragile states look remarkably like the era that we are meant to have left behind. In 2012, the median fragile state still relied on aid for 50 percent of its foreign capital, the other half being made up of non-concessional loans, investment, and remittances. In other developing economies, aid represented only 10 percent.[3]

This is the crux of the argument behind calls to reallocate aid away from stable countries and toward fragile states: Unlike stable countries, fragile states continue to face urgent development needs and are at risk of being left behind. The U.K. government and the World Bank, the leading bilateral and multilateral donors in the global development community, have recently set targets to increase the share of their aid devoted to fragile states.

However, there remains a profound ambivalence about aid's efficacy in these places. If development outcomes are driven by countries themselves rather than by outsiders, how much can aid realistically achieve in situations where partner governments are unable

---

[1] USAID 2016

[2] Calculations based on a static classification of fragility, applying the World Bank's 2015 classification retrospectively over the past 25 years. The same result is found applying the OECD's 2015 classification although the rise in the headcount is more modest. Some fragile states have seen their poverty headcounts fall over this period, but this is canceled out by others whose headcounts have risen. See discussion on the limits of applying an ex-post classification retrospectively in Roesch 2014.

[3] Calculations based on Development Initiatives 2016, using the World Bank's 2015 fragile states classification. Applying the OECD's 2015 fragile states classification, aid represented 43 percent of international capital for the median fragile state.

or unwilling to govern effectively? Even advocates of aid to fragile states are unlikely to argue that aid can dramatically improve how governments govern. Furthermore, aid-giving in fragile states seems especially vulnerable to causing unintended harm. One of the greatest criticisms of aid, forcefully made by Nobel Laureate Angus Deaton, is that it undermines the social contract that holds countries together by inserting donors into the relationship between states and citizens. In fragile states, a relationship that is already an Achilles heel risks being made weaker.[4]

These are valid concerns that cannot easily be argued away. Yet the evidence for aid's impotence in fragile states is notably flimsy and has not stood the test of time. First, in the late 1990s and early 2000s a series of influential papers used cross-country growth regressions to show that the impact of aid was conditional on the quality of the receiving country's policies and institutions—a damning result for fragile states where the quality of policies and institutions are lacking.[5] But, in the following years, a much larger series of papers poked holes in the methods of this earlier research and discredited its results.[6] Second, at the same time that these earlier findings were being taken seriously, the World Bank reported that the share of its projects in fragile states receiving unsatisfactory evaluations was double the share in stable countries. But when projects over the subsequent decade were evaluated, that difference vanished. The recent portfolio performance of the U.K. Department for International Development and the Global Fund in fragile states has been similarly favorable.[7] As for aid's capacity to unintentionally cause harm, some well-documented cases exist, as well as an emerging literature on aid's incitement of conflict in certain conditions.[8] But, to the extent that it can be measured, there is little evidence of any systematic negative effect of aid on recipient institutions.[9]

One upshot of these conflicting narratives is a focus on the operational aspects of aid giving. So long as donors are engaged in fragile states, they face a responsibility to ensure they are delivering aid as effectively as possible—a responsibility that is arguably more critical in fragile states, where aid's usefulness is especially contested, than in other settings. Improving the operations of donors in fragile states offers a practical response to concerns regarding aid's potential to do harm. It is these practices that are the focus of this paper.

---

4 By contrast, others have stressed the potentially transformative effect of aid in fragile states where fragility is conceived of as a low-level equilibrium. See Andrimihaja et al 2011 and Ravallion 2013.

5 Burnside and Dollar 2000; Collier and Dollar 2002

6 Roodman 2007 among many others

7 Chandy 2011; Hellman 2013

8 Strandow et al 2016; Nunn and Qian 2014

9 Jones and Tarp 2016

III. TWO CHALLENGES

There is a large literature devoted to assessing aid operations and their adherence to principles of good practice. Many papers were published in the years immediately following the 2005 High Level Forum on Aid Effectiveness in Paris.[10] This event marked a high point for consensus on the definition of good practice and the urgency of reform, and motivated new cross-country surveys of aid practices, enabling detailed industry-wide analysis for the first time.

This literature has tended to focus on the differences between donors while treating recipients as homogeneous. By contrast, a small number of papers have looked at how aid practices differ between fragile and stable countries.[11] They unanimously find that aid practices tend to be poorer in fragile states.

Interesting though these findings are, they suffer from two limitations.

Measuring good aid practice: First, earlier studies have borrowed measures of aid practice from the broader aid effectiveness literature that inherently discriminate against fragile states. Some measures penalize aid giving to fragile states by rewarding aid that goes to well-governed countries, or those whose governments display high quality management of public finance and procurement. Others penalize aid practices that are not appropriate in fragile states by rewarding the use of direct budgetary support and low overheads.[12] Using these measures, the finding that aid practices tend to be poorer in fragile states than stable countries is spurious.

Measuring fragility: Second, each earlier study employs a single classification of fragile states. This ignores the variety of different classifications of fragile states that abound and the conflicting definitions of fragility that underpin them.[13] It is reasonable to expect that different kinds of fragility and the degree of fragility might affect aid practices in different ways. Thus it is unclear whether the findings from earlier studies can be generalized across fragile environments.

---

[10] Easterly and Pfutze, 2008; Birdsall and Kharas 2010; ActionAid 2011; Knack et al. 2011

[11] OECD 2010; Chandy and Linn 2011; Ishihara 2012; OECD 2012a; OECD 2012b

[12] In a study of World Bank portfolio performance in fragile states, Ralston 2014 finds that high overheads are correlated with more satisfactory outcomes.

[13] See discussions in ACLED 2014 and Putzil and Di John 2012. It is notable that both the U.K. government and the World Bank are in the process of establishing new fragile state classifications.

IV. OUR APPROACH

In this paper we offer a fresh attempt to assess aid practices in fragile states that adopts a more careful approach to the measurement of aid practices and fragility.

Measuring good aid practice

We began by looking for a clear articulation of what constitutes effective aid delivery in fragile conditions. Two sources were especially relevant. The first was the now-retired Organization for Economic Cooperation and Development's (OECD) Principles for Good International Engagement in Fragile States. These were originally drafted in 2005 and intended as a complement to the Paris declaration to adapt its principles to fragile settings. The second was the g7+ New Deal on Fragile States, which offers a more up-to-date and country-led articulation of aid effectiveness principles as they apply in fragile environments.[14]

These sources are products of the traditional aid effectiveness agenda and focus on policies and practices set primarily at the aid agency level. In recent years a new literature has emerged advocating for adaptive, problem-driven approaches to aid practices that is targeted more at program staff.[15] While this can be interpreted as a rival to the traditional aid effectiveness agenda, we instead consider the two to be complementary and reflective of the same underlying principles.[16]

We derived three essential themes of good donor stewardship from these sources: 1) showing *respect* for countries' programming, systems, and staff; 2) working in *cooperation* with other donors; and 3) being a source of *stability* to recipient countries. These themes are all practical applications of the Hippocratic principle of "do no harm." They are also norms that remain relevant in stable countries, so that a comparison of practices in fragile and stable countries is still valid.

Our next step was to identify measures of aid practice that evince the three themes as faithfully as possible. For this we conducted a review of all known aid datasets and donor surveys with extensive international coverage, as well as previous analyses of good aid practice.

This resulted in a list of 10 indicators after singling out those for which donors can reasonably be held accountable for performance, irrespective of the perfor-

[14] The implementation of both the OECD's Principles for Good International Engagement in Fragile States and the g7+ New Deal on Fragile States have been subject to evaluations. However, these have been qualitative in nature and lack rigor. Moreover, by treating aid effectiveness in fragile states as sui generis, they resist efforts to compare donor practices in stable and fragile countries.

[15] See for instance Andrews and Pritchett 2013 and Wild et al. 2015.

[16] Honig 2016 interprets adaptive, problem-driven approaches in terms of traditional aid effectiveness indicators taken from the Paris Declaration monitoring surveys and thus shows how the two agendas are complementary. See later discussion under VII. Policy Implications.

mance of the recipient country. The indicators capture aid practices that we judge as being equally feasible in fragile and stable countries.[17]

The 10 indicators are as follows:

Respect for countries' programming, systems, and staff

1. *Reported on budget*. Recipient governments' ability to optimize their own spending decisions depends on being fully informed of donor spending plans. This requires no more than a commitment to basic communication and courtesy by donors.

2. *Use of public financial management systems, controlling for system quality*. Donors are encouraged to channel their aid through recipient governments' public financial management systems in order to strengthen the systems themselves and to raise accountability for their performance. Yet in many countries, especially in fragile states, these systems are poor quality and vulnerable to corruption as reported by independent assessments. We provide an objective measure of donors' use of systems by controlling for these assessment scores.

3. *Use of project implementation units*. PIUs are freestanding, parallel structures outside of government that replicate public systems. They are typically operated by personnel poached from government with knowledge of how those systems work. PIUs represent one of the most pernicious forms of aid giving, especially in settings where government capacity is already weak.

4. *Use of programmatic aid, excluding budget support*. Programmatic aid describes activities that are determined by recipient governments or nongovernmental organizations, combine donor and recipient resources in a single budget framework, and make some attempt to use recipient systems. Direct budget support is the purest form of programmatic aid but is considered unfeasible in most fragile states, whereas other forms of programmatic aid remain viable.

Cooperation with other donors

5. *Joint activities*. Country missions (i.e., visits by experts or dignitaries) and analytical work are two areas where donors can cooperate to reduce the burden on recipients and create shared knowledge. This approach is feasible everywhere.

6. *Aid through others*. Silent partnerships and delegated cooperation arrangements offer bilateral donors a way to support development in situations where they may lack relevant expertise, seek to reinforce successful engagements by others, and/or minimize the burden on recipients. These approaches are well-suited for providing aid in difficult environments.

[17] One way to understand our choice of indicators is to relate them to the determinants of aid effectiveness. Aid effectiveness is described by Howes 2011 as a function of donor characteristics, recipient characteristics and the interaction of the two. Our interest here is to focus on donor characteristics, and the interaction component as it relates to fragile states. Our approach bears some similarity to Knack et al. 2011, which is unique among the studies of aid agency practices in attempting to control for factors that are outside of agencies' control.

7. *Multilateral aid*. Multilateral aid offers the purest form of cooperation by bilateral donors. To the extent that it substitutes for bilateral aid, it reduces the burden on recipients. While the overall scale of multilateral aid depends on allocation decisions by donor countries, its use in fragile versus stable countries depends on the allocation choices of multilateral agencies themselves.

8. *Multi-bi aid*. So-called "multi-bi" aid, or non-core multilateral aid, offers a form of cooperation between donors in which individual donors retain some control. Aid is pooled in a trust fund and administered by a multilateral agency separately from its core funding. The trust fund is devoted to a specific purpose determined by the sponsoring donors—typically funding collective action problems that might otherwise be underfunded. This makes them well-suited for use in fragile states.

***Stability* to recipient countries**

9. *Predictable funding*. Recipient governments are better able to optimize their own multi-year investment plans when they can reliably anticipate donor spending. Previous studies of aid predictability have often focused on donors' provision of expenditure forecasts while overlooking their accuracy. This indicator assesses the accuracy of donors' spending forecasts on a one-year time horizon.

10. *Volatility*. Aid volatility has been found to negatively affect recipients' economic growth and to reduce levels of domestic investment. In fragile states, volatile aid risks amplifying countries' internal instability. Previous studies of aid volatility have typically focused on the volatility of aggregate aid flows into a country. This overlooks the problem of volatile flows from individual donors, which poses a particular challenge in fragile states where aid is highly projectized and so less easily reallocated to plug shortfalls from other donors. Here we assess the volatility of flows at the level of individual donor-recipient pairs.

Of the 10 indicators, four are borrowed wholesale from those used elsewhere in the aid effectiveness literature (indicators 1, 3, 5, and 7). The remainder are tailored to the study of fragile states in one of three ways: by reworking existing indicators from the literature to capture the most salient aspects for fragile states (indicators 9 and 10); by reworking existing indicators from the literature in a way that controls for the country environment (indicators 2 and 4); or by monitoring the use of aid delivery modalities that are well-suited for fragile states (indicators 6 and 8).

The 10 indicators have some weaknesses and certainly do not capture all relevant aspects of good aid practice. (See Appendix A for a thorough account of each indicator's construction, coverage, and weakness). However, we consider them a significant improvement on those indicators used in prior analyses comparing donor performance in fragile and stable countries.

Two additional points should be stressed.

First, the indicators, and the sources from which they are derived, assume an approach to aid giving in which donors place their engagement with the state at the center. Working with the state poses intrinsic challenges in fragile states but is largely unavoidable so long as donors view state-building and a strengthen-

ing of the relationship between citizens and the state as the ultimate route by which countries can escape fragility. Our assessment of aid practices is of less relevance to donors who opt to work on issues that do not concern the state, and for analyzing humanitarian assistance.

Second, while our choice of indicators represents a faithful interpretation of the literature on good aid practices in fragile states, the evidence linking good practice with development outcomes remains limited.[18] That makes recent research linking aid effectiveness indicators in fragile states with aid performance as measured by project evaluations especially important, and gives greater credence to the approach adopted by this paper.[19]

Measuring fragility

A solution to the measurement of fragility is more straightforward.

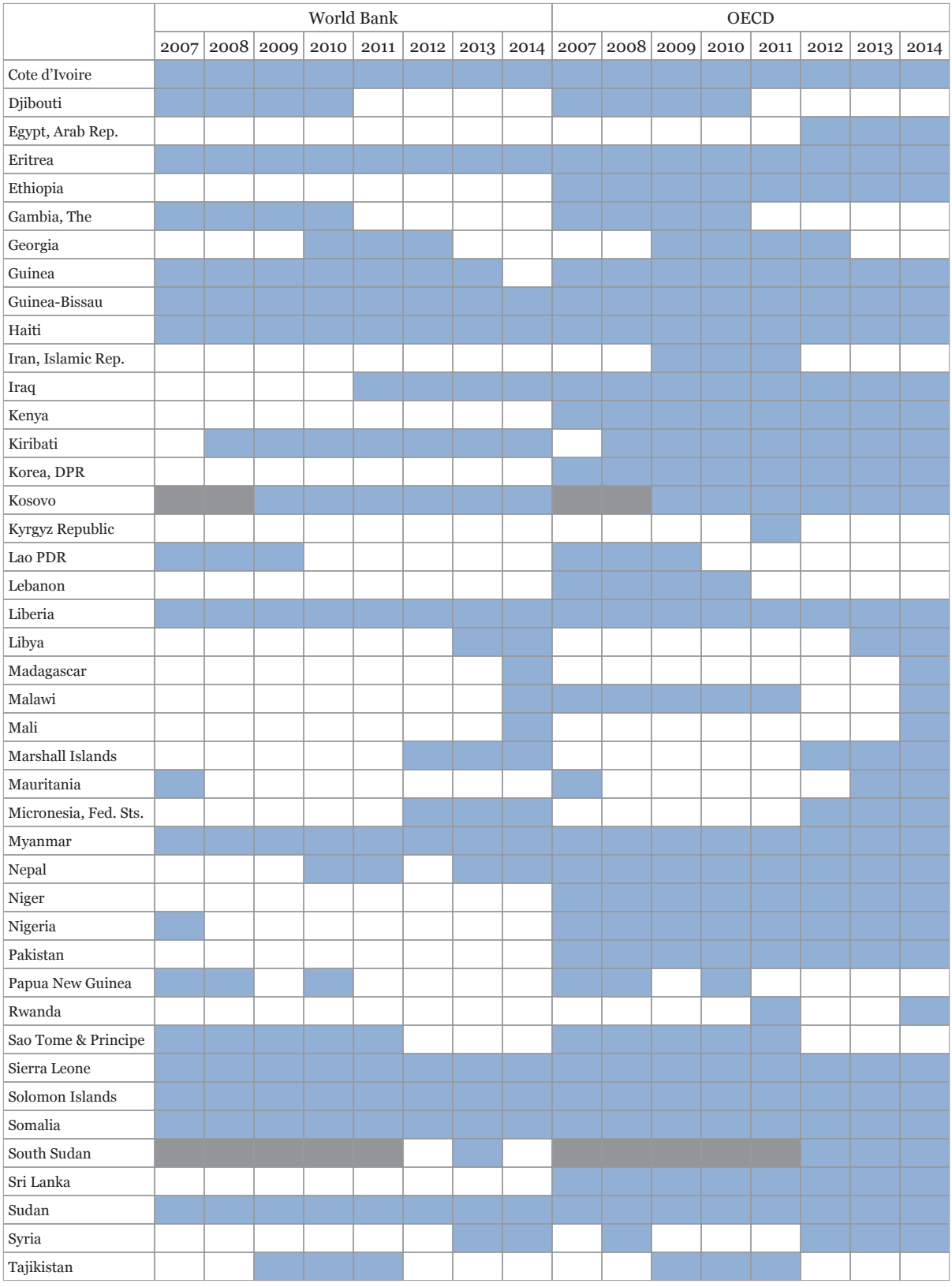
We use both the World Bank and OECD classifications of fragility when assessing aid practices. These two classifications are widely used, well-established, and annually updated. Using both classifications allows us to assess the generalizability of our findings across different kinds of fragile environments. In addition, we examine the design of the two fragility classifications to assess how different kinds of fragility, and the degree of each country's fragility, may affect aid practices. Table 1 provides a list of the countries contained in each of the two fragility classifications.

Table 1: World Bank and OECD fragility classifications

	World Bank								OECD							
	2007	2008	2009	2010	2011	2012	2013	2014	2007	2008	2009	2010	2011	2012	2013	2014
Afghanistan	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Angola	■	■	■	■	■	■	■		■	■	■	■	■	■	■	
Bangladesh									■	■	■	■	■	■	■	■
Bosnia & Herzegovina				■	■	■	■	■				■	■	■	■	■
Burkina Faso											■	■			■	
Burundi	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Cambodia	■	■	■						■	■	■					
Cameroon			■	■						■	■	■	■	■	■	■
CAR	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Chad	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Comoros	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Congo, Dem. Rep.	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Congo, Rep.	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

[18] Klein and Harford 2005; Easterly 2007. One area where there is more evidence is aid volatility, which has been linked with lower investment, exports, and GDP in the recipient country. See Hudson and Mosley 2008; Arellano et al. 2005; Celasun and Walliser 2005.

[19] Honig 2016 provides a complementary argument that the aid effectiveness agenda is of particular importance in fragile states, to the extent that it promotes greater agency autonomy and decentralization. The inherent unpredictability of fragile environments implies less of a role for routinized solutions and more adaptive, problem-driven approaches in their place.


World Bank
OECD
2007 2008 2009 2010 2011 2012 2013 2014 2007 2008 2009 2010 2011 2012 2013 2014
Cote d'Ivoire
Djibouti
Egypt, Arab Rep.
Eritrea
Ethiopia
Gambia, The
Georgia
Guinea
Guinea-Bissau
Haiti
Iran, Islamic Rep.
Iraq
Kenya
Kiribati
Korea, DPR
Kosovo
Kyrgyz Republic
Lao PDR
Lebanon
Liberia
Libya
Madagascar
Malawi
Mali
Marshall Islands
Mauritania
Micronesia, Fed. Sts.
Myanmar
Nepal
Niger
Nigeria
Pakistan
Papua New Guinea
Rwanda
Sao Tome & Principe
Sierra Leone
Solomon Islands
Somalia
South Sudan
Sri Lanka
Sudan
Syria
Tajikistan

	World Bank								OECD							
	2007	2008	2009	2010	2011	2012	2013	2014	2007	2008	2009	2010	2011	2012	2013	2014
Timor-Leste	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Togo	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Tonga	■	■	■	■					■	■	■	■				
Tuvalu							■	■							■	■
Uganda									■	■	■	■	■	■	■	■
Uzbekistan	■	■	■						■	■	■	■				
Vanuatu	■	■							■	■						
West Bank & Gaza	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Yemen, Rep.			■	■	■	■	■	■	■	■	■	■	■	■	■	■
Zimbabwe	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

Note: The OECD classification of fragile states reported in this table differs marginally from published versions of the classification. We have reconstructed the index to correct for the misalignment of years in its construction in some years. See Footnote 26 for details.

V. COMPARING THE QUALITY OF AID PRACTICES IN FRAGILE AND STABLE STATES

Our starting point is an assessment of how aid practices differ in fragile and stable countries across our 10 indicators.

However, it is impossible to make such an assessment without first determining which states are fragile. We begin by using the World Bank's classification, arguably the best-known and most commonly used fragility classification in the development community. This classification is made up of two overlapping groups of countries: low-income economies with meager state capabilities as reflected in low Country Policy and Institutional Assessment (CPIA) scores, and conflict hotspots where peacekeeping missions are active.[20] Persistently low CPIA scores capture a direct component of fragility in terms of the functioning of the state but also provide a remarkably strong predictor of civil war.[21]

Our analysis is based on a series of regressions in which each of our 10 indicators is separately regressed on a fragile classification dummy variable and a group of relevant controls.[22] Each observation represents the aid received by a given country in a given year from the donor community. This means there are multiple observations from the same country in a single regression, each reflecting aid from a different year.[23] The regressions are weighted using the value of aid represented by each observation so that the results capture

[20] The inclusion of conflict hotspots and the threshold for CPIA scores have not been consistently applied over the history of the World Bank classification. For details see World Bank 2016. Nevertheless, we employ the World Bank's classifications as published, given the difficulty of reconstructing the classification retroactively.

[21] World Bank 2011

[22] The controls include per capita income, population, year dummies, and the share of aid received from each of the world's 27 largest donors. Donor variables are excluded from the regression for multilateral aid to avoid over-specification.

[23] Our use of multiple observations from the same country, but from different years, is potentially problematic. If the observations are duplicative they will artificially drive up the significance of our results. Because many countries are persistently fragile or stable, there is not enough within-country variation in the fragility dummy to permit the use of country fixed effects. For those states that are not persistently fragile or stable, our approach allows us to analyze the relationship between aid practices and fragility as their status changes over time. Our confidence in this approach is strengthened by the fact that our results were unaffected when we tried running our analysis using regional dummy variables as a crude substitute for fixed effects. Our use of multiple observations from each country is further justified by the fact that each country's values on each indicator vary significantly across years.

the effect of fragility on a typical dollar of aid.[24] Our analysis uses all available data for each indicator from the year 2007, when the World Bank fragility classification was sufficiently well-established, to 2014, which is the last year for which detailed estimates of aid flows are available. The number of years for which there are observations and the degree of country coverage vary significantly across our 10 indicators. (See Appendix A for details). This poses an important, though unavoidable, limitation on the robustness of our findings.

Table 2 summarizes the key results. (Full regression results are provided in Appendix B.) Fragility is a significant correlate of worse aid practices for six of the 10 effectiveness indicators, and a significant correlate of better practices for two indicators.[25] The inferior performance of donors in fragile states relative to stable countries is especially apparent on indicators concerned with donor respect for countries' programming, systems, and staff. On the positive side, there is evidence that donors are employing cooperative arrangements in fragile states that are especially suited to these environments—namely, multi-bi aid and delivering aid through others. For several indicators, fragility does not just account for a significant difference in aid practices but the differences are large in absolute terms.

On balance, these results reinforce the findings of earlier studies that donor practices in fragile states are inferior to those in stable countries, at least on the basis of the World Bank's fragility classification. This judgment can now be made with greater certainty

Table 2: Comparing aid practices in fragile and stable states using World Bank fragility classification

Respect	► Aid is reported on budget 12 percentage points *less* in fragile states. ► Donors use recipient public financial management systems 8 percentage points *less* in fragile states. ► There is no significant difference between donors' use of PIUs in fragile and stable states. ► Programmatic aid is 10 percentage points *lower* in fragile states.
Cooperation	► Donors perform 10 percentage points *fewer* of their activities jointly with other donors in fragile states. ► Donors provide 6 *more* percentage points of aid through other donors in fragile states. ► Fragile states receive 9 *fewer* percentage points of their aid from multilateral donors. ► Fragile states receive 7 *more* percentage points of their aid in multi-bi aid.
Stability	► There is no significant difference in the accuracy of donors' aid projections in fragile and stable states. ► Aid volatility is 7 percentage points *higher* in fragile states.

Note: Higher scores indicate better performance on all indicators other than use of PIUs, predictable funding, and volatility.

[24] The value of aid is measured in gross country programmable aid. This is a narrow measure of aid limited to flows available for development projects and programs, which is therefore free of the volatility associated with debt relief and humanitarian assistance.

[25] Significance here and throughout the paper refers to the 10 percent significance level. Of the eight indicators where fragility is a significant determinant of aid practices, seven are also significant at the 5 percent level. See Appendix B for details.

given our choice of effectiveness measures that do not penalize against fragile states but rather capture practices that are equally feasible in both environments.

However, there are at least two alternative explanations for these results that deserve a hearing.

First, it is possible that the results are driven by old inferior donor practices employed in the earlier years of the period under analysis. Since that time, it may be argued, donors have become more aware of the risks of doing harm in fragile environments, and adapted their practices accordingly.

To test for this, we tried interacting our fragility dummy with the year dummies to specifically capture changes in donor performance in fragile states. We also ran the same regressions using a continuous year variable in place of the year dummies in case this might better capture trends over time. We found no evidence of improved donor behavior in either model. This should come as no great surprise given the very modest progress reported against targets in the monitoring frameworks of the Paris Declaration and the Global Partnership for Effective Development Cooperation. (Full regression results are provided in Appendix C.)

A second explanation is that the relationship we observe between donor performance on effectiveness measures and the country setting reflects the different content of aid projects and programs in fragile and stable environments. In other words, the patterns in aid allocations and modalities should not be interpreted as autonomous donor choices in response to the country environment, but rather they are a function of the sectors and activities that are prioritized in each setting.

To test whether this interpretation is valid, we re-ran our regressions with a series of sector share dummies using OECD Development Assistance Committee (DAC) sector headings, and interacted these with our fragility dummy. We find scant evidence of a relationship between our aid effectiveness measures and either the sectors in which aid is employed in general, or the sectors in which aid is employed in fragile states alone. (Full regression results are provided in Appendix D.)

VI. TESTING FRAGILITY CLASSIFICATIONS

Our initial regression results indicate that donors discriminate against fragile states as classified by the World Bank. In other words, the World Bank's definition of fragility is capturing some important recipient characteristics that affect donor behavior. However, what is true for the World Bank classification may not hold for other classifications that capture other types of fragility.

We therefore test again for differences in aid practices between fragile and stable states, substituting the OECD DAC's classification for the World Bank's classification. The OECD DAC's list of fragile states is broader, combining all countries on the World Bank list with additional countries facing a range of risks reflected in an elevated score on the Fund for Peace's Fragile States Index.[26]

Table 3 summarizes the key results. Fragility is a significant correlate of worse aid practices for only two of the 10 effectiveness indicators and a significant correlate of better practices for two indicators. The four significant results are consistent with those found under the World Bank fragility classification. However, for the remaining six indicators, no significant relationship is found. (Full regression results are provided in Appendix E.)

The relationship between aid practices and fragility is thus considerably less apparent under the OECD DAC fragility definition. This is true both in terms of the smaller number of indicators for which fragility accounts for a significant difference in aid practices and the scale of these differences in absolute terms. Critically, our earlier finding using the World Bank's classification that donor practices in fragile states are, on average, inferior to those in stable countries, does not carry across to the OECD DAC's broader classification.

Table 3: Comparing aid practices in fragile and stable states using OECD DAC fragility classification

Respect	► Aid is reported on budget 8 percentage points *less* in fragile states. ► There is no significant difference in donors' use of recipient public financial management systems in fragile and stable states. ► There is no significant difference between donors' use of PIUs in fragile and stable states. ► There is no significant difference between the use of programmatic aid in fragile and stable states.
Cooperation	► There is no significant difference in donors' use of joint activities in fragile and stable states. ► Donors provide 6 *more* percentage points of aid through other donors in fragile states. ► Fragile states receive 3 *fewer* percentage points of their aid from multilateral donors. ► Fragile states receive 6 *more* percentage points of their aid in multi-bi aid.
Stability	► There is no significant difference in the accuracy of donors' aid projections in fragile and stable states. ► There is no significant difference in aid volatility between fragile and stable states.

Note: Higher scores indicate better performance on all indicators other than use of PIUs, predictable funding, and volatility.

[26] We reconstruct the OECD classification ourselves to correct for the misalignment of years in the incorporation of the World Bank and Fund for Peace categories in published versions of the OECD fragile states list. This is necessary to allow a fair comparison of donor practices across the OECD and World Bank classifications.

This serves as an important reminder that fragility is a broad term whose definition is contested. What is true for one set of fragile states cannot be casually applied to all others. Competing fragile state classifications should not be viewed as imperfect proxies of the same underlying characteristics; they are more likely measures of different characteristics entirely.

Our next step was to try to understand what explains the divergence between our results from the two fragility classifications. We did this by comparing the composition of the two country groups to see if we can identify differences in their underlying characteristics. Our approach was to compare countries in the World Bank classification with those uniquely in the OECD DAC classification since it is the addition of the latter countries that drives the differences between our two sets of results.

We began by calculating the average scores of the two country groups on the Fragile States Index. We did this for each of the years 2007 to 2014, using the same aid weights employed in our regressions.[27] The average score for the countries in the World Bank classification is higher, on average, than for countries uniquely in the OECD DAC classification, where higher scores indicate greater levels of fragility. This difference is statistically significant, yet still relatively small.[28] This suggests that it could be the degree of fragility that accounts for the differences in our results, with fragility leading, on average, to inferior aid practices only where fragility is greatest. Yet, when we test this hypothesis by regressing our aid effectiveness indicators on the Fragile States Index scores using only observations from countries on the OECD DAC fragile states list, we fail to establish such a relationship.

To dig a little deeper, we look at the average scores of the two country groups on the 12 components of the Fragile States Index, again using the same aid weights—see Table 4. The countries in the World Bank classification are found to be significantly more fragile than those uniquely in the OECD DAC classification on six of the components, and significantly less fragile on one component.

These components point to different roots and manifestations of fragility that might feasibly influence aid practices differently and therefore account for our divergent regression results. Yet in the absence of a clear theory of how these components relate to each other, this surely risks over-interpreting the differences between the components of the Fragile States Index. This view is reinforced by the inherent contestability of the Fragile States Index as a measure of fragility, and the imperfections in the data used in its construction, both of which have been widely documented.[29]

---

[27] The country groups therefore contain multiple entries of the same country when that country features repeatedly in a classification over several years. These entries will each have a unique Fragile States Index score and aid weight corresponding to a given year.

[28] We obtain a similar result when we replace the Fragile States Index scores with those from the International Country Risk Guide (ICRG)'s Political Risk index. This is based on a smaller sample of countries given the smaller country coverage of the ICRG data.

[29] ACLED 2014

Table 4: Average Fragile States Index scores by fragility classification

FSI Component	World Bank Fragile States	OECD Only Fragile States	Difference
Demographic Pressures	8.67	8.49	0.18
Refugees and IDPs	8.33	7.81	0.52***
Group Grievance	8.49	8.79	-0.30**
Human Flight	7.33	7.34	-0.02
Uneven Development	8.24	8.25	-0.00
Poverty and Economic Decline	7.97	7.41	0.56***
Legitimacy of the State	8.91	8.45	0.46***
Public Services	8.53	7.85	0.68***
Human Rights	8.32	8.29	0.03
Security Apparatus	8.73	8.34	0.38**
Factionalized Elites	8.84	8.98	-0.14
External Intervention	8.90	7.82	1.08***
Total	101.27	97.84	3.43***

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

If our interest is in raising aid effectiveness in fragile states, it makes sense to define fragility in terms of recipient characteristics that actually affect donor behavior. For this reason, in our remaining analysis we focus only on the World Bank's fragility classification in an attempt to understand this relationship more clearly.

VII. COMPARING THE QUALITY OF AID PRACTICES ACROSS DONORS

The preceding analysis showed that fragility, as defined by the World Bank, explains only a portion of aid practices, and our use of donor control variables pointed to differences between donors. In this section, we turn the spotlight on to these differences. This can help to identify which donors are responsible for bad practice in fragile states and most vulnerable to accusations of doing harm. The identification of better and worse performing donors in fragile states might reasonably help to raise aid industry performance by prompting donors to either improve their practices or reallocate their aid in ways that improve the division of labor among donors. To reiterate, we rely exclusively on the World Bank classification of fragile states in this part of our analysis given its success in capturing recipient characteristics that cause donors to perform differently.

To give our analysis a sharper focus, we narrow our attention on 22 donors that each account for at least 1 percent of global aid to fragile states in any year between 2007 and 2014.[30] For each of our aid effectiveness indicators, we examine the performance of these donors in both stable and fragile countries by accumulating unique scores for each donor-recipient pair in each year.[31] These scores are then standardized so that the aid-weighted mean is equal to zero. The standardized scores are then aggregated up across recipients to generate a separate donor score in fragile and stable countries for each indicator, using weights based on the value of aid each observation represents.[32] Finally, we combine each donor's scores across the indicators with a simple average to generate a separate score for its overall effectiveness in fragile and stable countries that can be compared with the other 21 donors.[33]

Table 5 contains the headline results. (Scores for individual indicators both in standardized and absolute terms are provided in Appendix F.) The first column provides donor scores in fragile states, the second column provides donor scores in stable countries, and the third column provides the difference between donor scores in the two environments, which we call the "fragile premium."

The first takeaway, which is consistent with our earlier analysis, is that donor performance is typically, though not always, subpar in fragile states. This is in-

[30] We also exclude those donors for whom we have insufficient data to track performance on at least four of our effectiveness indicators—namely, the Arab Fund, the International Monetary Fund, Turkey, and United Arab Emirates.

[31] We drop the multilateral aid indicator as this cannot be measured at the level of individual donor-recipient pairs. When, for any given indicator, we have insufficient data to account for at least 20 percent of a donor's gross country programmable aid in either stable or fragile settings, we exclude that donor's pair-level data from our analysis of that indicator.

[32] The value of aid is again measured in gross country programmable aid.

[33] This approach, which is borrowed from other papers in the aid effectiveness literature—see for instance Birdsall and Kharas 2010—has the virtue of allowing us to include donors that do not record scores in every one of the nine indicators.

Table 5: Standardized donor scores for overall effectiveness

Donor	Fragile Score	Stable Score	Fragile Premium
Denmark	0.47	0.26	0.21***
Canada	0.34	0.25	0.09
Sweden	0.33	0.36	-0.02
Norway	0.32	0.28	0.04
Netherlands	0.26	0.31	-0.05
Italy	0.17	-0.09	0.26***
United Kingdom	0.10	0.29	-0.19***
Japan	0.09	0.01	0.08**
Australia	0.07	0.03	0.05
World Bank	-0.02	0.11	-0.13**
Global Fund	-0.06	0.26	-0.31***
Asian Development Bank	-0.07	0.07	-0.15*
Germany	-0.09	0.02	-0.12***
EU Institutions	-0.11	0.12	-0.22***
United Nations	-0.13	-0.11	-0.01
Switzerland	-0.21	-0.06	-0.15**
Spain	-0.23	-0.13	-0.10
GAVI Alliance	-0.33	-0.19	-0.13***
Belgium	-0.46	-0.09	-0.37***
African Development Bank	-0.47	-0.16	-0.31***
United States	-0.49	-0.22	-0.27***
France	-0.56	-0.03	-0.53***
Total	-0.13	0.04	-0.17***

**** p<0.01, ** p<0.05, * p<0.1*

dicated by the fact that a slim majority of donors have a negative fragile states score, falling short of the industry average across all environments (which would be indicated by a score of zero). Seven of the eight worst donor performances recorded in any setting are in fragile states, as is indicated by the largest negative numbers recorded in the first two columns. These donors are most vulnerable to the accusation of doing harm through their aid practices in fragile states. The story is more positive for the nine donors at the top of the table. Their performance in fragile states exceeds the industry average in all environments. In fact, it is striking how strong some of these performances are: they account for four of the top six donor performances recorded in any setting.

The second related takeaway is that performance at the individual donor level broadly mirrors that at a global level: individual donors tend to perform better in stable countries than fragile states. This is indicated by the fact that the fragile premiums for 16 of the 22 donors are negative.

While donor practices are on average less effective in fragile states than in stable environments, performance in fragile states still varies significantly between donors. That variance might be explained by one or both of two factors: differences between donor performance in general that are independent of the environment, or differences between donors in how they differentiate between stable and fragile conditions. In other words, a bad (or good) donor in fragile states may be one who is bad (or good) everywhere, reflected in a high correlation in the scores across the first two columns, or one who discriminates negatively (or positively) toward fragile states, reflected in a high correlation in the scores between the first and third column.

Our results show that both factors are relevant. The correlation between donor scores in fragile and stable settings is 0.79. The eight worst performing donors in fragile states also underperform in stable countries, where they fall short of the industry average across all environments. Similarly, all the donors who exceed the overall industry average in fragile states exceed that standard in stable countries as well—with the notable exception of Italy. At the same time, the correlation between donor scores in fragile states and their

Figure 1: Relative rankings of donor performance in fragile and stable states

fragile premium is 0.80. The five worst performers in fragile states alter their practices significantly for the worse in fragile settings, whereas none of the six donors with the highest scores in fragile states perform significantly worse in fragile states than in stable ones.

These two correlations explain the relationship between donor rankings in fragile and stable settings illustrated in Figure 1. The figure is composed of both horizontal lines, where donor rankings across stable and fragile countries are similar, and diagonal lines where rankings diverge. The predominance of the horizontal lines is due to the fact that the best performers in fragile states are good performers in all environments and minimize the difference in their practices across the two environments. At the same time, poor performers in fragile states tend to perform poorly everywhere and alter their practices for the worse in fragile settings.

Next, we consider the performance of bilateral donors versus multilateral donors in fragile states. Table 6 provides the average scores for each donor type, weighted by aid dollars. The average score of bilateral donors in fragile states falls far below the global average for the aid industry in all environments, whereas the average score for multilateral donors falls only marginally short of that standard. That is a remarkable result given that bilateral donors account for each of the nine positive donor scores recorded in fragile states. Those performances are offset, however, by the poor performance of the U.S. It alone is responsible for 40 percent of all bilateral aid to fragile states from our sample of donors between 2007 and 2014. (See Appendix A for donor shares of global aid flows.)

By contrast, the average performance of bilateral and multilateral donors in stable settings are very similar. (Scores for bilateral and multilateral donors on each indicator are available in Appendix G.)

The superior average performance of multilateral donors in fragile states is all the more striking considering the earlier result that multilateral aid is used less in fragile states than stable countries. This points to a failure in the division of labor in the aid industry.

Table 6: Average standardized donor scores for multilateral and bilateral donors

	Multilateral Score	Bilateral Score	Difference
Stable States	0.08	0.05	0.03**
Fragile States	-0.08	-0.32	0.24***

**** p<0.01, ** p<0.05, * p<0.1*

VIII. POLICY IMPLICATIONS

The aid industry's approach to fragile states has undergone significant change since the start of the century. Whereas 15 years ago donor policy toward fragile states could be characterized as mildly allergic, recent years have seen earnest attempts to deepen support and find more effective ways of working in these countries, motivated by some combination of evidence, theology, and politics.

Three aspects of this change are noteworthy. First, donors have come to see fragile states as a shared responsibility of the global community—a global public bad that will be insufficiently resourced unless all donors bear part of the burden. Second, donors have sought to structure their engagement with fragile states in terms of a mutually accountable relationship, whereby both donor and recipient jointly commit to actions that can improve development outcomes. This recognizes that donor action alone is insufficient to deliver countries from fragility but hinges on actions taken by countries themselves. Third, program staff have been encouraged to adopt a more adaptive, problem-driven approach to project design and implementation, which forces attention on understanding how progress can be supported in a given context.

The analysis in this paper suggests that these well-intentioned policies may have had unintended consequences. Conceiving of fragile states as a shared burden encourages donors to give aid to countries where they may have little or no expertise, which increases the likelihood of doing harm. Conditioning changes in donor practice on recipient country reforms is a recipe for inaction in fragile states, where fragility extends to the ability of government to fulfill their commitments.[34] Finally, a focus on adaptive, problem-driven approaches has been interpreted as an alternative to the traditional aid effectiveness agenda and shifted attention away from agency-level reforms, providing cover for those efforts to stall.

The latest set of proposals for advancing aid effectiveness in fragile states seems to double down on these ideas.[35] We propose two alternative policies that can serve as a course correction.

The first is to define some minimum standards for donor practices in fragile states. This would serve as a practical manifestation of a Hippocratic Oath for donors. To be credible, these standards would have to be observable and verifiable. The set of 10 indicators used in this paper provide a possible starting point for thinking about how such standards could be defined.

The truth is that upholding such standards is going to be a challenge. Attempts to improve donor performance through the agreement of goals and the monitoring of indicators have a woefully poor track record.

---

[34] This may help to explain the disappointing track record in the implementation of the New Deal on Fragile States—specifically its failure to deliver meaningful reform in donor practices.

[35] For instance, Greenhill 2016 explicitly pits adaptive problem-driven approaches as an alternative to the traditional aid effectiveness agenda and argues that the former is more appropriate in fragile states. Meanwhile, Norris 2016 proposes a new modality for U.S. engagement in borderline fragile states built explicitly around a mutual accountability framework, which, rather than buttressing countries against their internal instability, could easily reinforce these effects.

Earlier this month, the Global Partnership on Effective Development Cooperation met in Nairobi to declare that none of the industry goals set five years ago in Busan have been met. The Busan meeting, it should be recalled, was the culmination of the previous attempt to raise effectiveness among donors, which proved a similar flop. The only difference in Nairobi, compared to Busan, was the absence of any sense of outrage. Expectations had been calibrated. A new effort to raise donor performance specifically in fragile states will have to think of ways to avoid this fate.

The second policy, which could form part of the minimum standards, is to encourage a stronger division of labor among donors in fragile states. We specifically would encourage greater reliance on the better performing multilateral agencies by the worst performing bilateral donors. This too is not a straightforward ask—especially in the U.S. since its poor performance looms over the other donors given its outsized share of global aid to fragile states. Nevertheless, this may prove to be the simplest way to lift aid industry performance in these settings. The World Bank's upcoming International Development Association replenishment, which promises to deliver a significant scaling up of resources and intent in fragile states, provides a promising opportunity to jumpstart this change.

APPENDIX A: DESCRIPTION OF INDICATORS, DATA COVERAGE, AND SUMMARY STATISTICS

Definitions of Indicators

Reported on budget

Our measure of the extent to which aid is reported on recipients' budgets compares two values from the Paris and Global Partnership surveys: the amount of ODA disbursed to the government sector as reported by donors, and the same amount as recorded on recipients' budget sheets. We divided aid recorded by aid disbursed and calculated the absolute deviation from 1 for each donor-recipient pair in each year. Our use of the absolute deviation captures both under- and over-reporting because both forms of miscommunication between donors and recipients lead to less effective aid.

However, it should be noted that the extent to which aid appears on recipient budgets may be determined in part by the recipient as well as by the donor. If recipients do not have the will or the capacity to record and publish accurate accounts of aid received, there may be discrepancies between aid recorded and aid disbursed through no fault of the donor. In this respect, the indicator fails to fully control for recipient country characteristics.

Use of public financial management (PFM) systems

The Paris and Global Partnership surveys asked donors to report what percent of their ODA to the government sector uses 1) national budget execution procedures, 2) national financial reporting procedures, and 3) national auditing procedures. We averaged these three percentages to measure the share of aid that goes through public financial management systems.

The use of public financial management systems is clearly less feasible in countries where systems are of poor quality, and our indicator takes this into account. The Paris and Global Partnership surveys used subcomponent 13 of the World Bank's Country Policy and Institutional Assessment (CPIA) scores to measure the quality of recipient systems. To measure donor use of PFM systems conditioned on system quality, we first regressed our measure of system quality on the share of aid channeled through PFM systems for each donor-recipient pair in each year. We then calculated the residuals from this regression. This captures the extent of donor use of PFM systems that is unexplained by system quality.

Use of project implementation units (PIUs)

The Paris surveys asked donors how many parallel PIUs were used in delivering aid to the government sector. Because the question specifically asked about parallel PIUs, our indicator avoids punishing donors for creating project implementation units in places where no comparable government system exists. In order to control for aid volume, our indicator measures the number of PIUs per $100 million of official development assistance (ODA) to the government sector for each donor-recipient pair in each year.

Use of programmatic aid, excluding budget support

The Paris surveys asked donors for the amount of ODA that was provided in support of initiatives adopting program-based approaches. Donors report

two sub-components of that amount: direct budget support and other programmatic aid. We measured only other programmatic aid to avoid discriminating against fragile states, where direct budget support is in most cases considered unfeasible. We calculated this volume as a share of total ODA for each donor-recipient pair. Following Paris Indicator 9, we divided by total ODA instead of ODA to the government sector because programmatic approaches are possible when working with the nongovernment sector.[36]

In stable countries, it is possible that donors use direct budget support as a substitute for other forms of programmatic aid. If this is the case, our indicator would be biased toward finding greater amounts of programmatic aid in fragile situations.

Joint activities

Donors reported on two types of joint activities in the Paris surveys: joint missions and joint analytical work. We calculated joint missions as a share of total missions and joint analytical work as a share of total analytical work, and took the simple average of these two percentages for each donor-recipient pair.

Aid through others

The Paris and Global Partnership surveys ask donors to report the value of any ODA to the government sector that they disburse through other donors. Aid given through other donors can include delegated cooperation arrangements, "silent" partnerships, or contributions channeled through multilateral donors at the country level. We calculated the ratio of this value to all ODA to the government sector provided directly by each donor, to reflect a donor's reliance on this aid modality relative to the size of its direct efforts for each donor-recipient pair.

Multilateral aid

We calculated the share of each recipient's country programmable aid (CPA) in each year that comes from multilateral donors.

This indicator cannot be measured at the donor-recipient level—the share of multilateral aid from a multilateral donor is always 100 percent and the share of multilateral aid from a bilateral donor is always 0 percent. The indicator is therefore excluded from our analysis of individual donors' effectiveness.

Multi-bi aid

Our measure of multi-bi aid, also known as non-core multilateral aid, is drawn from the OECD DAC's records of CPA channeled through multi-donor trust funds. We calculated multi-bi CPA as a share of total CPA for each donor-recipient pair in each year, excluding all regional funds.

Multi-bi aid has been rightly criticized for its distortionary effect on multilateral agencies.[37] However, it represents an unambiguous improvement over bilateral aid from the perspective of recipients by enabling donors to cooperate and harmonize efforts.

[36] OECD 2010

[37] OECD 2011

The multi-bi aid modality is employed only by bilateral donors. We therefore excluded this indicator from our measurement of donor performance for multilateral donors.

Predictable funding

We assessed the accuracy of donors' spending predictions by comparing predicted and actual ODA to the government sector as reported in the Paris and Global Partnership surveys. Donors reported their predicted and actual spending for the same year, allowing us to measure predictability on a one-year time horizon.

To determine this measure, we calculated the absolute difference between actual ODA and predicted ODA and divide this by predicted ODA to obtain a deviation ratio—strictly speaking, a measure of unpredictability. We used absolute deviations to capture the fact that both under- and over-delivering on promised aid undermine recipient governments' ability to optimize their own budget allocations. We also capped the ratio at 100 percent to ensure symmetry between under- and over-delivering; while donors cannot provide fewer than zero dollars, there is no upper limit on how much they can provide. Over-delivering by 400 percent is not four times worse than under-delivering by 100 percent and failing to give any aid at all. The cap ensures that we avoided misinterpreting such situations.

Unlike many measures of predictability, our indicator is calculated at the donor-recipient pair level. This is predicated on the fact that aid from different donors is not perfectly fungible. This is especially the case in fragile states where aid is highly projectized, so aid from one donor cannot necessarily be used to fill in the gaps left by another. Using a measure of the predictability of each recipient's aggregate aid receipts, aid would be considered highly predictable in a country where all donors' forecast expenditure collectively add up to $1 billion and $1 billion is provided, even though individual donors provide drastically more or less than they promise.

It is possible that the amount of aid delivered could fall short of the predicted amount due to the limitations of the recipient government's implementation capacity. Predictability is largely but not entirely a function of donor practices, and this should be kept in mind when interpreting donor performance on this indicator.

Volatility

Our indicator measures volatility in CPA which, unlike ODA, excludes sectors such as debt relief and humanitarian aid where volatility is expected and often welcome. To measure volatility, we regressed CPA for each donor-recipient pair on time using four years' worth of data—for example, 2010 volatility scores are based on data going back to 2007—and calculated the residuals. These residuals represent the de-trended flows of CPA, or the difference between actual CPA and the amount of CPA we would have expected based on trends over time.

We then calculated the coefficient of variation—the standard deviation as a share of the mean—of the residuals within that four-year period to obtain a volatility score. Finally, we adjusted the coefficient of variation to account for the importance of aid in the recipient country. Volatility matters most where aid is large relative to the share of the recipient economy;

for instance, highly volatile aid could destabilize Malawi but will barely register in Brazil. We therefore took the value of CPA for each donor-recipient pair, divided by the recipient's share of GDP, and took the square root to moderate the impact of extreme values. We then multiplied each volatility score by this value divided by the aid-weighted mean of all such values. This serves to increase the volatility score for countries in which aid is a greater-than-average share of GDP, and vice versa.

The OECD DAC's CPA data is less complete prior to 2007 and we would run the risk of interpreting gaps in reporting as evidence of volatile aid flows if we were to calculate volatility scores using data from earlier years. Because our scores for each year rely on data from the three preceding years as well, we can only calculate volatility scores for 2010 through 2013. Donor-recipient pairs with CPA flows in fewer than three of the four years are also excluded.

It should be noted that, in rare cases, aid volatility may be a sign that donors are being appropriately responsive to changing situations in recipient countries. For example, a large unexpected increase in aid would be justified if donors chose to show strong support for a new government that takes on positive reforms. For further discussion of aid volatility as a response to changing circumstances in recipient countries, see Desai and Kharas (2010).

As with predictability, it is important to measure volatility at the donor-recipient pair level instead of at the aggregate level because aid is highly projectized in fragile states.

Coverage

Table A.1a: Number of recipients with data on each indicator

	2007	2008	2009	2010	2011	2012	2013	2014
Reported on budget	48	--	--	70	--	--	45	--
Use of PFM systems	41	--	--	56	--	--	33	--
Use of PIUs	54	--	--	77	--	--	--	--
Use of programmatic aid	54	--	--	77	--	--	--	--
Joint activities	54	--	--	77	--	--	--	--
Aid through others	--	--	--	77	--	--	40	--
Multilateral aid	145	145	146	145	147	146	147	145
Multi-bi aid	146	146	147	147	148	148	148	146
Predictable funding	54	--	--	77	--	--	39	--
Volatility	--	--	--	134	134	134	135	134

Table A.1b: Share of global aid represented by donor-recipient pairs with data

	2007	2008	2009	2010	2011	2012	2013	2014
Reported on budget	45.5%	--	--	57.6%	--	--	31.7%	--
Use of PFM systems	40.2%	--	--	49.7%	--	--	29.5%	--
Use of PIUs	53.3%	--	--	66.5%	--	--	--	--
Use of programmatic aid	52.3%	--	--	66.5%	--	--	--	--
Joint activities	41.7%	--	--	51.9%	--	--	--	--
Aid through others	--	--	--	66.5%	--	--	28.0%	--
Multilateral aid	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Multi-bi aid	64.0%	65.1%	61.8%	62.1%	63.5%	62.4%	62.9%	61.1%
Predictable funding	43.0%	--	--	50.8%	--	--	23.1%	--
Volatility	--	--	--	96.3%	96.5%	96.3%	93.7%	96.6%

Summary Statistics

These summary statistics describe all recipient-year observations, which are constructed by taking the aid-weighted mean of donor-recipient pair level data for a given year.

Table A.2: Indicator Summary Statistics

	Recipient-Year Observations	Mean	Standard Deviation	Max	p5	p95	Max
Reported on budget	163	47.4	21.0	0.0	10.6	77.8	97.0
Use of PFMs (residual)	130	2.3	21.3	-40.6	-25.1	40.9	59.1
Use of PIUs	131	7.4	8.6	0.0	0.3	24.4	51.5
Use of programmatic aid	131	18.4	13.6	0.0	0.9	41.4	65.7
Joint activities	131	40.1	16.1	5.0	13.3	67.4	100.0
Aid through others	117	8.5	10.0	0.0	0.0	34.5	50.0
Multilateral aid	1,176	36.7	21.8	0.0	3.9	73.2	98.9
Non-core multilateral aid	1,176	7.4	9.2	0.0	0.0	25.1	83.2
Predictable funding	170	32.2	17.6	3.0	9.5	67.4	81.8
Volatility	671	25.7	33.7	0.6	3.1	72.6	370.0

Table A.3: Country Programmable Aid from Each Donor, 2007-2014

Donor	% of Total CPA	% of CPA to Fragile States	% of CPA to Stable States
African Development Bank	1.9%	1.6%	2.0%
Arab Fund (AFESD)	0.9%	0.9%	0.9%
Asian Development Bank	2.6%	2.1%	2.7%
Australia	2.4%	3.5%	2.1%
Belgium	0.5%	0.8%	0.4%
Canada	1.3%	1.5%	1.2%
Denmark	1.2%	0.9%	1.3%
EU Institutions	8.9%	10.9%	8.3%
France	4.9%	3.0%	5.5%
GAVI Alliance	0.9%	1.1%	0.8%
Germany	5.2%	4.2%	5.5%
Global Fund	3.2%	3.1%	3.2%
IMF (Concessional Trust Funds)	1.5%	2.6%	1.2%
Italy	0.6%	0.6%	0.6%
Japan	10.6%	6.5%	11.8%
Netherlands	1.5%	1.2%	1.6%
Norway	1.5%	2.3%	1.3%
Spain	1.3%	0.8%	1.4%
Sweden	1.3%	1.7%	1.2%
Switzerland	0.8%	0.9%	0.8%
Turkey	0.9%	1.2%	0.8%
United Arab Emirates	2.0%	1.5%	2.2%
United Kingdom	5.1%	5.9%	4.9%
United Nations	1.8%	3.0%	1.4%
United States	17.1%	22.8%	15.4%
World Bank	12.3%	9.1%	13.2%
Other	7.9%	6.3%	8.4%

Fragile states were identified using the World Bank fragility classification.

APPENDIX B: REGRESSION USING WORLD BANK FRAGILITY CLASSIFICATION

Table B.1: Regression using World Bank fragility classification

	Reported on budget	Use of PFM systems	Use of PIUs	Use of program-matic aid	Joint activities	Aid through others	Multi-lateral aid	Multi-bi aid	Predict-able funding	Volatility
Fragile Dummy	-12.0***	-7.7*	2.1	-10.1**	-10.0**	6.2**	-9.3***	7.3***	1.7	6.8***
	(4.4)	(4.5)	(1.7)	(4.2)	(4.1)	(2.9)	(1.2)	(0.6)	(4.6)	(1.7)
log(GDP Per Capita)	-0.6	-0.1	1.8*	-2.1	2.7	-1.5	-9.2***	-3.3***	1.7	-6.8***
	(2.9)	(3.1)	(1.0)	(2.5)	(2.5)	(1.6)	(0.5)	(0.4)	(2.9)	(1.0)
Population (millions)	-0.0	0.1	-0.0	0.0	-0.0	0.0*	-0.0***	-0.0***	0.0	-0.0
	(0.0)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Year = 2008							0.6	0.7		
							(2.0)	(0.8)		
Year = 2009							3.8*	2.1***		
							(2.0)	(0.8)		
Year = 2010	-1.8	-5.2	-5.0***	-3.6	-4.3		2.2	1.9**	6.9*	
	(3.8)	(4.8)	(1.5)	(3.8)	(3.8)		(2.0)	(0.8)	(4.1)	
Year = 2011							1.3	2.0**		1.3
							(2.0)	(0.8)		(1.7)
Year = 2012							2.2	2.3***		-0.2
							(2.0)	(0.8)		(1.7)
Year = 2013	-5.2	-3.8				0.6	2.4	1.9**	4.2	2.9
	(5.1)	(6.2)				(2.1)	(1.9)	(0.8)	(5.3)	(1.7)
Year = 2014							4.1**	1.9**		2.2
							(2.0)	(0.8)		(1.8)
Constant	105.8***	-23.3	2.0	64.2**	36.1	10.7	103.2***	22.2***	14.8	106.1***
	(32.1)	(36.2)	(12.7)	(31.9)	(31.4)	(18.1)	(4.3)	(4.4)	(33.6)	(12.0)
Observations	161	130	131	131	131	115	1,068	1,068	170	661
R-squared	0.6	0.7	0.5	0.4	0.5	0.5	0.2	0.6	0.4	0.5
# of Recipients	81	61	80	80	80	83	135	135	81	134

Standard errors in parentheses. Donor control variables excluded from table. Higher scores indicate better performance on all indicators other than use of PIUs, predictable funding, and volatility.
**** p<0.01, ** p<0.05, * p<0.1*

APPENDIX C: REGRESSIONS TESTING FOR CHANGING EFFECT OF FRAGILITY OVER TIME

Table C.1: Regression interacting fragility with year dummies

	Reported on budget	Use of PFM systems	Use of PIUs	Use of program-matic aid	Joint activities	Aid through others	Multi-lateral aid	Multi-bi aid	Predict-able funding	Volatility
Fragile Dummy	-17.2**	-11.6	1.5	-8.8*	-12.9**	1.6	-5.4*	7.7***	9.5	10.7***
	(7.1)	(7.4)	(2.0)	(5.0)	(4.9)	(4.2)	(3.3)	(1.3)	(7.9)	(3.1)
log(GDP Per Capita)	0.2	-0.2	1.8*	-2.1	2.7	-1.1	-9.2***	-3.3***	1.0	-6.8***
	(3.0)	(3.1)	(1.0)	(2.6)	(2.5)	(1.6)	(0.5)	(0.4)	(2.9)	(1.0)
Population (millions)	-0.0	0.1	-0.0	0.0	-0.0	0.0*	-0.0***	-0.0***	0.0	-0.0
	(0.0)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Year = 2008							0.8	0.4		
							(2.3)	(0.9)		
Year = 2009							4.2*	1.4		
							(2.2)	(0.9)		
Year = 2010	1.0	-2.8	-4.5**	-4.7	-1.8		2.0	1.6*	1.5	
	(4.6)	(5.5)	(1.8)	(4.5)	(4.4)		(2.3)	(0.9)	(5.0)	
Year = 2011							1.7	1.8*		0.9
							(2.3)	(0.9)		(1.9)
Year = 2012							2.6	1.5		1.6
							(2.3)	(1.0)		(2.0)
Year = 2013	-1.2	-0.8				2.3	2.3	1.2	-2.9	2.6
	(6.0)	(6.9)				(2.4)	(2.2)	(0.9)	(6.4)	(2.0)
Year = 2014							3.3	1.3		1.5
							(2.2)	(0.9)		(2.0)
Year 2007 x Fragile	11.5	11.0	1.5	-2.9	6.8		-3.7	-2.5	-17.8*	
	(9.4)	(11.0)	(2.6)	(6.6)	(6.4)		(4.9)	(1.9)	(9.9)	
Year 2008 x Fragile							-4.7	-0.9		
							(4.8)	(1.9)		
Year 2009 x Fragile							-5.6	0.9		
							(4.6)	(1.7)		
Year 2010 x Fragile	4.3	3.4				6.5	-2.7	-1.2	-4.1	-3.8
	(8.3)	(8.4)				(4.4)	(4.5)	(1.7)	(8.6)	(4.0)
Year 2011 x Fragile							-5.5	-1.3		-2.3
							(4.5)	(1.7)		(4.0)
Year 2012 x Fragile							-5.3	0.7		-11.2***
							(4.6)	(1.8)		(4.1)
Year 2013 x Fragile							-3.4	0.2		-2.4
							(4.5)	(1.7)		(4.0)
Constant	95.7***	-26.7	1.1	65.9**	32.3	7.0	103.1***	22.7***	27.8	106.3***
	(33.1)	(36.5)	(12.8)	(32.3)	(31.5)	(18.2)	(4.4)	(4.4)	(34.0)	(12.0)
Observations	161	130	131	131	131	115	1,068	1,068	170	661
R-squared	0.6	0.7	0.6	0.4	0.5	0.5	0.2	0.6	0.4	0.6
# of Recipients	81	61	80	80	80	83	135	135	81	134

Standard errors in parentheses. Donor control variables excluded from table. Higher scores indicate better performance on all indicators other than use of PIUs, predictable funding, and volatility.
**** p<0.01, ** p<0.05, * p<0.1*

Table C.2: Regression with continuous year variable interacted with fragility

	Reported on budget	Use of PFM systems	Use of PIUs	Use of programmatic aid	Joint activities	Aid through others	Multilateral aid	Multi-bi aid	Predictable funding	Volatility
Fragile Dummy	-6.2	-1.8	3.0	-11.7**	-6.2	14.5**	-10.6***	6.5***	-6.4	3.3
	(6.3)	(7.5)	(2.2)	(5.6)	(5.5)	(6.3)	(2.3)	(1.0)	(6.4)	(4.7)
log(GDP Per Capita)	0.0	0.0	1.8*	-2.1	2.7	-1.1	-9.2***	-3.3***	0.6	-6.8***
	(2.9)	(3.1)	(1.0)	(2.6)	(2.5)	(1.6)	(0.5)	(0.4)	(2.9)	(1.0)
Population (millions)	-0.0	0.0	-0.0	0.0	-0.0	0.0*	-0.0***	-0.0***	0.1	-0.0
	(0.0)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Years Since 2007	-0.3	0.1	-1.5**	-1.6	-0.6	0.8	0.3	0.1	-0.6	0.4
	(1.0)	(1.1)	(0.6)	(1.5)	(1.5)	(0.8)	(0.2)	(0.1)	(1.1)	(0.5)
Years Since 2007	-1.9	-2.0	-0.5	1.0	-2.3	-2.2	0.4	0.3	3.3**	0.7
	(1.5)	(1.8)	(0.9)	(2.2)	(2.1)	(1.5)	(0.5)	(0.2)	(1.6)	(0.9)
Constant	98.8***	-32.9	1.1	65.9**	32.3	4.7	104.4***	23.7***	33.1	104.6***
	(32.3)	(35.7)	(12.8)	(32.3)	(31.5)	(18.5)	(4.2)	(4.4)	(33.7)	(12.1)
Observations	161	130	131	131	131	115	1,068	1,068	170	661
R-squared	0.6	0.7	0.6	0.4	0.5	0.5	0.2	0.6	0.4	0.5
# of Recipients	81	61	80	80	80	83	135	135	81	134

Standard errors in parentheses. Donor control variables excluded from table. Higher scores indicate better performance on all indicators other than use of PIUs, predictable funding, and volatility.
**** p<0.01, ** p<0.05, * p<0.1*

APPENDIX D: REGRESSION TESTING ROLE OF AID ALLOCATIONS TO SECTORS IN EXPLAINING AID PRACTICES

Table D.1: Regression testing role of aid allocations to sectors in explaining aid practices

	Reported on budget	Use of PFM systems	Use of PIUs	Use of programmatic aid	Joint activities	Aid through others	Multilateral aid	Multi-bi aid	Predictable funding	Volatility
Fragile Dummy	77.9	5.4	-1.4	38.2	-63.1	-39.7	13.6	-7.5	19.1	-46.6**
	(56.9)	(69.9)	(19.6)	(43.2)	(46.1)	(30.3)	(12.0)	(5.0)	(50.0)	(20.7)
log(GDP Per Capita)	-1.3	1.7	2.2*	-5.2*	2.4	1.0	-8.0***	-2.7***	-0.6	-7.4***
	(3.4)	(3.6)	(1.2)	(2.7)	(2.9)	(1.7)	(0.7)	(0.4)	(3.0)	(1.3)
Population (millions)	-0.0	-0.1	-0.0	0.0	-0.0	0.0	-0.0***	-0.0***	0.0	-0.0
	(0.0)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Education Sector	0.0	1.0	0.3	-0.0	-0.2	-0.8**	0.2*	-0.1	0.3	-0.6***
	(0.5)	(0.9)	(0.2)	(0.5)	(0.5)	(0.3)	(0.1)	(0.1)	(0.5)	(0.2)
Health Sector	-0.6*	-0.0	-0.0	-0.0	0.2	-0.6***	0.1*	-0.3***	1.3***	-0.6***
	(0.4)	(0.6)	(0.1)	(0.3)	(0.3)	(0.2)	(0.1)	(0.0)	(0.3)	(0.1)
Water and Sanitation Sector	0.7	0.1	-0.3	0.2	0.8	-0.0	-0.0	-0.2***	0.4	-0.2
	(0.5)	(0.7)	(0.3)	(0.6)	(0.6)	(0.3)	(0.1)	(0.1)	(0.5)	(0.2)
Gov't and Civil Society Sector	0.1	-0.5	0.1	-0.7**	-0.5*	-0.7***	-0.1	-0.1**	0.6*	-0.6***
	(0.4)	(0.7)	(0.1)	(0.3)	(0.3)	(0.2)	(0.1)	(0.0)	(0.4)	(0.1)
Physical Infrastructure Sector	-0.0	0.6	-0.1	0.3	0.2	-0.6***	0.1	-0.2***	-0.0	-0.5***
	(0.3)	(0.5)	(0.1)	(0.3)	(0.3)	(0.2)	(0.1)	(0.0)	(0.3)	(0.1)
Financial Infrastructure Sector	-0.5	-0.7	-0.1	-0.0	-0.2	-0.2	0.5***	-0.1***	1.0**	-0.6***
	(0.5)	(0.8)	(0.2)	(0.4)	(0.4)	(0.2)	(0.1)	(0.0)	(0.4)	(0.2)
Agriculture Sector	0.8	0.1	0.1	-0.7	-0.8*	-0.4	0.6***	-0.3***	0.3	-0.5**
	(0.5)	(0.7)	(0.2)	(0.4)	(0.5)	(0.3)	(0.2)	(0.1)	(0.5)	(0.2)
Other Economic Production Sector	0.2	-0.9	0.1	-0.5	-0.8*	0.0	-0.0	-0.0	0.4	-0.2
	(0.4)	(0.8)	(0.2)	(0.4)	(0.4)	(0.2)	(0.1)	(0.0)	(0.4)	(0.2)
Education Sector x Fragile	-0.1	-0.3	-0.3	-0.6	1.5	-0.1	-0.7**	0.1	0.8	2.2***
	(1.3)	(1.6)	(0.5)	(1.2)	(1.3)	(0.8)	(0.3)	(0.1)	(1.4)	(0.4)
Health Sector x Fragile	-0.4	0.3	0.3	-0.6	0.0	0.0	0.2	0.2***	-0.2	0.6**
	(0.7)	(0.9)	(0.2)	(0.5)	(0.5)	(0.3)	(0.2)	(0.1)	(0.6)	(0.3)
Water and Sanitation Sector x Fragile	-1.8	-0.0	0.0	0.7	-0.3	1.2	2.0***	-0.1	0.7	-1.0
	(1.7)	(1.8)	(0.7)	(1.5)	(1.6)	(1.0)	(0.5)	(0.2)	(1.7)	(0.7)
Gov't and Civil Society Sector x Fragile	-1.5**	0.3	-0.1	-0.1	1.1*	1.1***	-0.4***	0.3***	-0.6	0.4
	(0.7)	(0.9)	(0.2)	(0.5)	(0.6)	(0.4)	(0.2)	(0.1)	(0.6)	(0.3)
Physical Infrastructure Sector x Fragile	-1.9***	-1.8*	0.1	-1.2**	-0.1	0.5	-0.0	0.1	0.5	0.6**
	(0.7)	(0.9)	(0.2)	(0.5)	(0.5)	(0.4)	(0.2)	(0.1)	(0.6)	(0.3)
Financial Infrastructure Sector x Fragile	-0.7	2.8*	0.1	-1.7*	2.3**	0.8	-0.6*	0.1	-2.0*	0.8
	(1.2)	(1.6)	(0.4)	(1.0)	(1.0)	(0.9)	(0.3)	(0.1)	(1.1)	(0.6)
Agriculture Sector x Fragile	-0.2	0.4	-0.1	-0.0	0.8	0.8	-1.5***	0.2	-0.3	0.7
	(1.1)	(1.4)	(0.5)	(1.0)	(1.1)	(0.7)	(0.3)	(0.1)	(1.1)	(0.5)
Other Economic Production Sector x Fragile	-1.6*	-0.3	0.2	-0.3	1.3	-0.0	-0.0	-0.4***	-1.1	0.9**
	(0.9)	(1.5)	(0.6)	(1.3)	(1.3)	(0.5)	(0.3)	(0.1)	(1.5)	(0.4)
Constant	110.0***	-87.5*	-2.3	60.3*	29.8	39.8*	82.8***	34.8***	5.4	154.7***
	(40.5)	(51.5)	(15.9)	(35.0)	(37.4)	(20.5)	(8.6)	(5.0)	(37.8)	(16.8)
Observations	161	130	131	131	131	115	937	937	170	531
R-squared	0.7	0.8	0.6	0.6	0.6	0.7	0.4	0.7	0.6	0.6
# of Recipients	81	61	80	80	80	83	135	135	81	134

Standard errors in parentheses. Donor control variables and year dummies excluded from table. All sector variables measure the percent of total ODA delivered in each sector to each recipient in each year, based on CRS codes. Higher scores indicate better performance on all indicators other than use of PIUs, predictable funding, and volatility.
**** p<0.01, ** p<0.05, * p<0.1*

APPENDIX E: REGRESSION USING OECD DAC FRAGILITY CLASSIFICATION

Table E.1: Regression using OECD DAC fragility classification

	Reported on budget	Use of PFM systems	Use of PIUs	Use of programmatic aid	Joint activities	Aid through others	Multilateral aid	Multi-bi aid	Predictable funding	Volatility
Fragile Dummy	-7.9**	-5.4	1.6	-4.8	-3.3	6.4***	-3.3***	5.9***	2.9	1.0
	(3.6)	(3.6)	(1.5)	(3.8)	(3.8)	(2.3)	(1.1)	(0.5)	(4.0)	(1.6)
log(GDP Per Capita)	0.3	-0.2	1.7*	-1.2	4.1	-0.6	-8.7***	-4.1***	2.1	-8.3***
	(2.9)	(3.1)	(1.0)	(2.6)	(2.6)	(1.7)	(0.6)	(0.3)	(2.9)	(1.0)
Population (millions)	-0.0	0.1	-0.0	0.0	-0.0	0.0	-0.0**	-0.0***	0.0	-0.0
	(0.0)	(0.1)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
Year = 2008							0.6	0.9		
							(2.1)	(0.8)		
Year = 2009							3.6*	2.8***		
							(2.0)	(0.8)		
Year = 2010	-3.6	-7.9*	-4.7***	-5.5	-6.1		1.9	2.4***	7.1*	
	(3.8)	(4.7)	(1.5)	(3.8)	(3.8)		(2.0)	(0.8)	(4.0)	
Year = 2011							0.8	3.0***		1.6
							(2.0)	(0.8)		(1.7)
Year = 2012							2.0	3.1***		-0.1
							(2.0)	(0.9)		(1.8)
Year = 2013	-7.8	-7.2				1.0	2.4	2.9***	4.7	3.0*
	(5.2)	(6.1)				(2.0)	(2.0)	(0.8)	(5.2)	(1.8)
Year = 2014							4.1**	2.8***		2.4
							(2.0)	(0.8)		(1.8)
Constant	102.3***	-15.6	1.5	62.7*	32.4	3.4	99.1***	29.0***	12.7	114.4***
	(32.5)	(36.5)	(12.8)	(32.8)	(32.4)	(18.2)	(4.7)	(4.3)	(33.7)	(12.0)
Observations	161	130	131	131	131	115	1,068	1,068	170	661
R-squared	0.6	0.7	0.5	0.4	0.4	0.5	0.2	0.6	0.4	0.5
# of Recipients	81	61	80	80	80	83	135	135	81	134

Standard errors in parentheses. Donor control variables excluded from table. Higher scores indicate better performance on all indicators other than use of PIUs, predictable funding, and volatility.
**** p<0.01, ** p<0.05, * p<0.1*

APPENDIX F: INDIVIDUAL DONOR SCORES, IN FRAGILE AND STABLE SETTINGS

Table F.1: Donor scores across all indicators in fragile and stable states, with standard deviations

Donor	Fragile Score	Stable Score	Fragile Premium
Denmark	0.47 (0.08)	0.26 (0.02)	0.21*** (0.06)
Canada	0.34 (0.09)	0.25 (0.04)	0.09 (0.09)
Sweden	0.33 (0.06)	0.36 (0.03)	-0.02 (0.06)
Norway	0.32 (0.06)	0.28 (0.03)	0.04 (0.07)
Netherlands	0.26 (0.06)	0.31 (0.03)	-0.05 (0.07)
Italy	0.17 (0.07)	-0.09 (0.02)	0.26*** (0.05)
United Kingdom	0.10 (0.05)	0.29 (0.02)	-0.19*** (0.05)
Japan	0.09 (0.06)	0.01 (0.01)	0.08** (0.04)
Australia	0.07 (0.07)	0.03 (0.02)	0.05 (0.06)
World Bank	-0.02 (0.07)	0.11 (0.03)	-0.13** (0.06)
Global Fund	-0.06 (0.07)	0.26 (0.03)	-0.31*** (0.06)
Asian Development Bank	-0.07 (0.06)	0.07 (0.05)	-0.15* (0.08)
Germany	-0.09 (0.03)	0.02 (0.01)	-0.12*** (0.03)
EU Institutions	-0.11 (0.05)	0.12 (0.02)	-0.22*** (0.05)
United Nations	-0.13 (0.03)	-0.11 (0.03)	-0.01 (0.05)
Switzerland	-0.21 (0.05)	-0.06 (0.03)	-0.15** (0.06)
Spain	-0.23 (0.12)	-0.13 (0.02)	-0.10 (0.08)
GAVI Alliance	-0.33 (0.04)	-0.19 (0.03)	-0.13*** (0.05)
Belgium	-0.46 (0.05)	-0.09 (0.02)	-0.37*** (0.05)
African Development Bank	-0.47 (0.07)	-0.16 (0.04)	-0.31*** (0.07)
United States	-0.49 (0.04)	-0.22 (0.02)	-0.27*** (0.04)
France	-0.56 (0.05)	-0.03 (0.01)	-0.53*** (0.04)
Total	-0.13 (0.01)	0.04 (0.01)	-0.17*** (0.01)

Standard deviations are shown in parentheses in the first two columns, and the standard error of the difference is shown in parentheses in the third column. Standard deviations were calculated by determining the standard deviation of all pair-level scores for each donor on each indicator, then pooling the standard deviations for each donor across all indicators.
****p<0.01, ***p<0.05, *p<0.1*

Table F.2a: Raw individual donor scores in fragile states

Donor Name	Reported on budget	Use of PFM systems	Use of PIUs	Use of program-matic aid	Joint activities	Aid through others	Multi-bi aid	Predictable funding	Volatility
African Development Bank	55.2	-13.3	7.5	2.1	12.9	0.8		40.3	32.8
Asian Development Bank	74.5	55.5	4.1	3.6	32.7	0.2			43.9
Australia	53.3		6.6	13.6	29.0	35.6	22.9	23.3	36.5
Belgium	16.3	-24.4	38.6	13.1	18.4	10.2	17.4	23.0	11.1
Canada	41.8	36.1	24.6	17.9	35.8	73.7	45.6	46.5	21.4
Denmark	59.2	59.1	7.1	18.2	54.0	49.2	28.4	52.4	9.3
EU Institutions	45.5	-3.4	2.3	28.7	41.9	11.6		33.5	33.1
France	25.5	11.8	12.0	10.8	23.7	6.8	1.2	60.4	48.7
GAVI Alliance	5.0	-22.8	0.0	30.9		1.3		36.5	17.1
Germany	25.6	0.1	5.2	10.0	53.0	2.6	11.3	22.7	14.0
Global Fund		-11.3	1.0	30.5	21.1	37.0			36.2
Italy	51.1	-3.8	13.0	10.0		38.8	31.3	27.8	15.3
Japan	65.2	-17.5	0.0	19.7	0.6	46.3	25.9		23.1
Netherlands	15.6	-12.9	0.1	11.1	35.3	63.1	55.4	59.6	11.8
Norway	21.9	-7.7	6.2	25.5	17.0	75.0	45.3	36.3	9.6
Spain	16.6	20.6	30.8	25.7	39.4	40.9	25.0	72.5	33.9
Sweden	35.5	-18.0	0.8	14.5	32.7	42.6	48.9		6.1
Switzerland	14.8	-16.4	24.2	16.1	28.9	16.5	22.0	23.1	5.7
United Kingdom	45.2	-4.8	4.1	7.5	37.9	31.6	43.2	44.0	26.5
United Nations	39.2	-18.0	11.0	18.2	51.2	0.7		22.9	7.2
United States	8.2	-26.4	3.1	8.2	15.5	7.4	11.1	26.2	42.2
World Bank	58.5	13.8	2.3	19.7	56.3	1.0		26.2	36.9

When, for any given indicator, we have insufficient data to account for at least 20 percent of a donor's gross CPA in either stable or fragile settings, we exclude that donor's pair-level data from our analysis of that indicator.

Table F.2b: Raw individual donor scores in stable states

Donor Name	Reported on budget	Use of PFM systems	Use of PIUs	Use of program-matic aid	Joint activities	Aid through others	Multi-bi aid	Predictable funding	Volatility
African Development Bank	44.3	16.3	6.6	17.7	32.8	1.6		36.7	20.6
Asian Development Bank	71.9	42.3	0.6	15.6	27.4	0.6			20.2
Australia	27.7		6.3	7.0	51.0	29.6	17.3	46.5	10.6
Belgium	50.3	-10.3	18.5	19.8	40.4	16.3	8.8	28.0	8.2
Canada	51.4	28.4	8.9	23.7	38.7	21.9	23.7	20.1	8.6
Denmark	61.9	23.3	4.5	38.1	66.8	4.9	3.3	19.0	7.2
EU Institutions	55.7	11.3	2.4	16.2	51.1	14.1		23.0	16.4
France	59.0	25.2	4.1	22.5	27.3	5.8	0.2	27.6	13.9
GAVI Alliance	9.6	-16.8	0.0	26.6		0.3		19.3	11.7
Germany	49.7	3.3	1.4	25.2	57.5	0.6	0.8	28.8	7.8
Global Fund		18.2	0.0	77.8	23.7	5.0			18.3
Italy	35.8	3.0	11.6	29.7		8.5	7.4	31.8	9.1
Japan	67.8	27.1	0.0	28.7	14.9	3.4	0.9		10.6
Netherlands	59.7	25.5	1.9	23.6	52.3	19.2	17.3	22.4	8.6
Norway	64.2	34.7	3.8	15.2	48.0	18.0	17.2	20.2	10.7
Spain	37.4	7.9	5.7	8.3	40.4	13.1	7.4	43.3	11.6
Sweden	58.9	22.9	3.0	23.4	52.0	22.6	21.4		6.2
Switzerland	39.6	8.0	26.2	16.1	46.3	13.9	21.0	34.8	5.1
United Kingdom	71.3	34.2	2.4	24.5	50.6	15.7	15.4	30.4	13.4
United Nations	41.3	-14.3	24.9	31.2	55.5	3.9		26.0	3.4
United States	34.3	-13.4	6.2	28.9	33.4	5.5	3.2	41.0	12.8
World Bank	73.6	25.1	0.8	18.3	47.5	0.2		24.1	24.0

When, for any given indicator, we have insufficient data to account for at least 20 percent of a donor's gross CPA in either stable or fragile settings, we exclude that donor's pair-level data from our analysis of that indicator.

Table F.3a: Scaled individual donor scores in fragile states

Donor Name	Reported on budget	Use of PFM systems	Use of PIUs	Use of program-matic aid	Joint activities	Aid through others	Multi-bi aid	Predictable funding	Volatility
African Development Bank	0.07	-0.69	-0.21	-0.67	-0.93	-0.37		-0.35	-0.63
Asian Development Bank	0.62	1.23	0.00	-0.62	-0.23	-0.40			-1.11
Australia	0.02		-0.16	-0.29	-0.36	1.19	0.76	0.21	-0.79
Belgium	-1.02	-1.00	-2.14	-0.30	-0.74	0.05	0.44	0.23	0.31
Canada	-0.30	0.69	-1.27	-0.15	-0.12	2.90	2.04	-0.56	-0.14
Denmark	0.19	1.33	-0.19	-0.13	0.53	1.80	1.07	-0.75	0.38
EU Institutions	-0.20	-0.42	0.12	0.21	0.10	0.11		-0.12	-0.64
France	-0.76	0.01	-0.49	-0.38	-0.55	-0.10	-0.47	-1.02	-1.32
GAVI Alliance	-1.33	-0.96	0.26	0.29		-0.35		-0.22	0.05
Germany	-0.76	-0.32	-0.07	-0.41	0.49	-0.29	0.10	0.24	0.18
Global Fund		-0.64	0.19	0.27	-0.64	1.25			-0.78
Italy	-0.04	-0.43	-0.55	-0.41		1.33	1.23	0.07	0.13
Japan	0.35	-0.81	0.26	-0.09	-1.37	1.67	0.92		-0.21
Netherlands	-1.04	-0.68	0.25	-0.37	-0.13	2.42	2.59	-0.99	0.28
Norway	-0.86	-0.54	-0.13	0.11	-0.79	2.96	2.02	-0.22	0.37
Spain	-1.01	0.25	-1.65	0.11	0.01	1.43	0.87	-1.42	-0.68
Sweden	-0.48	-0.83	0.21	-0.26	-0.23	1.50	2.23		0.52
Switzerland	-1.06	-0.78	-1.25	-0.20	-0.36	0.33	0.71	0.22	0.54
United Kingdom	-0.21	-0.46	0.00	-0.49	-0.04	1.01	1.90	-0.47	-0.36
United Nations	-0.38	-0.82	-0.43	-0.13	0.43	-0.38		0.23	0.47
United States	-1.24	-1.06	0.06	-0.46	-0.84	-0.08	0.09	0.12	-1.04
World Bank	0.16	0.06	0.11	-0.09	0.61	-0.36		0.12	-0.81

When, for any given indicator, we have insufficient data to account for at least 20 percent of a donor's gross CPA in either stable or fragile settings, we exclude that donor's pair-level data from our analysis of that indicator. Normalized scores on three indicators (Use of PIUs, Predictability, and Volatility) were multiplied by -1 so that higher scores indicate better performance on all indicators.

Table F.3b: Scaled individual donor scores in stable states

Donor Name	Reported on budget	Use of PFM systems	Use of PIUs	Use of program-matic aid	Joint activities	Aid through others	Multi-bi aid	Predictable funding	Volatility
African Development Bank	-0.23	0.13	-0.15	-0.15	-0.22	-0.34		-0.23	-0.10
Asian Development Bank	0.54	0.86	0.22	-0.22	-0.42	-0.38			-0.09
Australia	-0.70		-0.13	-0.50	0.42	0.92	0.44	-0.56	0.33
Belgium	-0.06	-0.61	-0.89	-0.08	0.05	0.32	-0.04	0.06	0.43
Canada	-0.03	0.47	-0.29	0.05	-0.01	0.57	0.80	0.32	0.42
Denmark	0.26	0.33	-0.02	0.52	0.99	-0.19	-0.35	0.36	0.48
EU Institutions	0.09	-0.01	0.11	-0.20	0.43	0.23		0.23	0.08
France	0.18	0.38	0.00	0.01	-0.42	-0.15	-0.53	0.07	0.19
GAVI Alliance	-1.21	-0.79	0.26	0.14		-0.39		0.35	0.28
Germany	-0.08	-0.23	0.17	0.10	0.66	-0.38	-0.50	0.03	0.45
Global Fund		0.19	0.25	1.84	-0.55	-0.19			0.00
Italy	-0.47	-0.24	-0.46	0.25		-0.03	-0.12	-0.07	0.39
Japan	0.43	0.43	0.25	0.21	-0.86	-0.25	-0.49		0.33
Netherlands	0.20	0.39	0.14	0.04	0.47	0.45	0.44	0.25	0.41
Norway	0.33	0.65	0.02	-0.24	0.32	0.40	0.44	0.32	0.32
Spain	-0.43	-0.10	-0.10	-0.46	0.05	0.18	-0.12	-0.45	0.29
Sweden	0.18	0.32	0.07	0.04	0.46	0.61	0.67		0.52
Switzerland	-0.36	-0.10	-1.37	-0.20	0.26	0.22	0.64	-0.17	0.57
United Kingdom	0.53	0.63	0.11	0.07	0.41	0.30	0.33	-0.02	0.21
United Nations	-0.32	-0.72	-1.29	0.29	0.59	-0.23		0.12	0.64
United States	-0.51	-0.70	-0.13	0.22	-0.20	-0.16	-0.36	-0.37	0.24
World Bank	0.59	0.38	0.20	-0.13	0.30	-0.40		0.19	-0.25

When, for any given indicator, we have insufficient data to account for at least 20 percent of a donor's gross CPA in either stable or fragile settings, we exclude that donor's pair-level data from our analysis of that indicator. Normalized scores on three indicators (Use of PIUs, Predictability, and Volatility) were multiplied by -1 so that higher scores indicate better performance on all indicators.

APPENDIX G: BILATERAL AND MULTILATERAL DONOR SCORES ACROSS ALL INDICATORS

Table G.1: Bilateral and multilateral donor scores on each indicator

	Stable States			Fragile States		
Indicator	Multilateral Score	Bilateral Score	Difference	Multilateral Score	Bilateral Score	Difference
Reported on budget	0.30 (0.05)	0.01 (0.03)	0.28*** (0.06)	-0.03 (0.07)	-0.73 (0.06)	0.71*** (0.10)
Use of PFM systems	0.23 (0.04)	0.08 (0.04)	0.15** (0.06)	-0.19 (0.07)	-0.69 (0.06)	0.50*** (0.09)
Use of PIUs	0.06 (0.04)	-0.00 (0.03)	0.06 (0.05)	0.03 (0.04)	-0.13 (0.11)	0.15 (0.13)
Use of programmatic aid	0.05 (0.05)	0.10 (0.04)	-0.05 (0.07)	-0.02 (0.07)	-0.36 (0.05)	0.34*** (0.08)
Joint activities	0.19 (0.05)	0.02 (0.05)	0.16** (0.07)	0.21 (0.09)	-0.56 (0.09)	0.76*** (0.13)
Aid through others	-0.24 (0.03)	-0.01 (0.04)	-0.23*** (0.06)	-0.09 (0.07)	0.61 (0.12)	-0.69*** (0.14)
Predictable funding	0.15 (0.05)	-0.10 (0.05)	0.26*** (0.07)	0.01 (0.07)	-0.11 (0.08)	0.12 (0.11)
Volatility	-0.05 (0.02)	0.30 (0.01)	-0.36*** (0.02)	-0.56 (0.05)	-0.57 (0.04)	0.01 (0.06)
All Indicators	0.08 (0.01)	0.05 (0.01)	0.03** (0.01)	-0.08 (0.02)	-0.32 (0.03)	0.24*** (0.04)

Standard deviations are shown in parentheses beneath the multilateral scores and bilateral scores, and the standard error of the difference is shown in parentheses beneath the difference values. Standard deviations on each indicator were calculated by determining the standard deviation of all pair-level scores for all bilateral or multilateral donors, and these standard deviations were pooled to calculate the standard deviation across all indicators.
****p<0.01, ***p<0.05, *p<0.1*

REFERENCES

ActionAid (2011). "Real Aid: Ending Aid Dependency." https://www.actionaid.org.uk/sites/default/files/doc_lib/real_aid_3.pdf

Andrews, M. and L. Pritchett (2013). "Escaping Capability Traps Through Problem Driven Iterative Adaptation (PDIA)." *World Development* 51: 11.

Andrimihaja, N. A., et al. (2011). "Avoiding the Fragility Trap in Africa." *World Bank Working Paper 5884*.

Arellano, C., et al. (2005). "The Dynamic Implications of Foreign Aid and Its Variability" *IMF Working Paper 05/119*, International Monetary Fund.

Armed Conflict Location & Event Data (ACLED) Project (2014). "Real-Time Analysis of African Political Violence." *Conflict Trends 28*.

Birdsall, N. and H. Kharas (2010). "Quality of Official Development Assistance Assessment." Brookings Institution and Center for Global Development.

Burnside, C. and D. Dollar (2000). "Aid, Policies, and Growth." *The American Economic Review* 90(4): 847-868.

Celasun, O. and J. Walliser (2005). "Predictability of Budget Aid: Experiences in Eight African Countries." World Bank Practicioner's Forum on Budget Support.

Chandy, L. (2011). "Ten Years of Fragile States: What Have We Learned?" *Brookings Institution Global Views Policy Paper*, The Brookings Institution.

Chandy, L. and J. F. Linn (2011). "Taking Development Activities to Scale in Fragile and Low Capacity Environments." *Global Economy and Development Working Paper No. 45*, The Brookings Institution.

Collier, P. and D. Dollar (2002). "Aid Allocation and Poverty Reduction." *European Economic Review* 46(8): 1475–1500.

Desai, R. M. and H. Kharas (2012). "The Determinants of Aid Volatility." *Global Economy and Development Working Paper No. 42*, The Brookings Institution.

Development Initiatives (2016). "Unbundling Other Official Flows." http://devinit.org/#!/post/unbundling-other-official-flows

Easterly, W. (2007). "Are Aid Agencies Improving?", Centre for Economic Policy Research.

Easterly, W. and T. Pfutze (2008). "Where Does the Money Go? Best and Worst Practices in Foreign Aid." *Journal of Economic Perspectives* 22(2): 29-57.

Greenhill, R. (2016). "Development Effectiveness for the SDG Era: Five Reasons Why We Need a New Agenda." Overseas Development Institute. https://www.odi.org/comment/10451-development-effectiveness-sdg-era-new-agenda-aid

Hellman, J. (2013). "Surprising Results from Fragile States." *World Bank Future Development Blog*. http://blogs.worldbank.org/futuredevelopment/surprising-results-fragile-states

Honig, D. (2016). "Penny Wise, Point-Foolish: The Costs of Politically Constrained Organizational Autonomy in Foreign Aid Implementation."

Howes, S. (2011). "An Overview of Aid Effectiveness, Determinants and Strategies." *Development Policy Centre Discussion Paper No. 1*, Development Policy Centre.

Hudson, J. and P. Mosley (2008). "The Macroeconomic Impact of Aid Volatility." *Economic Letters* 99(3): 486–489.

Ishihara, Y. (2012). "Identifying Aid Effectiveness Challenges in Fragile and Conflict-Affected States." *Policy Research Working Paper No. 6037*, The World Bank.

Jones, S. and F. Tarp (2016). "Does foreign aid harm political institutions?" *Journal of Development Economics* 118: 266-281.

Klein, M. and T. Harford (2005). *The Market for Aid*, International Finance Corporation.

Knack, S., et al. (2011). "Aid Quality and Donor Rankings." *World Development* 39(11): 1907-1917.

Norris, J. (2016). "A Better Approach to Fragile States: The Long View." Center for American Progress.

Nunn, N. and N. Qian (2014). "US Food Aid and Civil Conflict." *American Economic Review* 104(6): 1630-1666.

Organization for Economic Cooperation and Development (OECD) (2010). *Resource Flows to Fragile and Conflict-Affected States*, OECD Publishing.

Organization for Economic Cooperation and Development (OECD) (2012a). "Aid Effectiveness 2011: Progress in Implementing the Paris Declaration."

Organization for Economic Cooperation and Development (OECD) (2012b). "Ensuring Fragile State Are Not Left Behind: 2011 Report on Financial Resource Flows."

Putzel, J. and J. Di John (2012). "Meeting the Challenges of Crisis States." *Crisis States Research Centre Report*, Crisis States Research Centre at the London School of Economics.

Ralston, L. (2014). "Success in Difficult Environments: A Portfolio Analysis of Fragile and Conflict-affected States." *Policy Research Working Paper No. 7098*, The World Bank.

Ravallion, M. (2013). "Cutting Development Assistance after a Coup May Be Bad Response." *Center for Global Development*. http://www.cgdev.org/blog/cutting-development-assistance-after-coup-may-be-bad-response

Roesch, A. (2014). "Understanding Poverty and Fragility: Moving from Rhetoric to Analysis," USAID.

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

Strandow, D., et al. (2016). "Foreign Aid and the Intensity of Violent Armed Conflict." A*idData Working Paper*.

United States Agency for International Development (2016). “Remarks by Administrator Gayle Smith: U.S. Leadership in International Development.” https://www.usaid.gov/news-information/speeches/mar-9-2016-administrator-gayle-smith-us-leadership-international-development

Wild, L. et. al (2015). “Adapting Development: Improving Services to the Poor.” Overseas Development Institute.

World Bank (2011). “Conflict, Security, and Development.” *World Development Report*.

World Bank (2016). “Information Note: The World Bank Group’s Harmonized List of Fragile Situations.” http://pubdocs.worldbank.org/en/586581437416356109/FCS-List-FY16-Information-Note.pdf

Cover photos courtesy of the World Bank
(in order, top to bottom)
Curt Carnemark
Jamie Martin
Kubat Sydykov
Igor Castro da Silva Braga
Jamie Martin
John Isaac
Kubat Sydykov
Tran Thi Hoa

BROOKINGS
1775 Massachusetts Ave., NW
Washington, D.C. 20036
202.797.6000
www.brookings.edu/global