GLOBAL ECONOMY & DEVELOPMENT
WORKING PAPER 45 | SEPTEMBER 2011

Global Economy and Development at BROOKINGS

TAKING DEVELOPMENT ACTIVITIES TO SCALE IN FRAGILE AND LOW CAPACITY ENVIRONMENTS

Laurence Chandy
Johannes F. Linn

Global Economy
and Development
at BROOKINGS

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

Johannes F. Linn is a nonresident senior fellow in the Global Economy and Development Program at the Brookings Institution and senior resident scholar at the Emerging Markets Forum.

Acknowledgments:

The authors are grateful for the excellent research assistance of Marek Hlavac, for feedback on earlier drafts from Greg Ellis, Jim Tulloch, Homi Kharas, Pat Duggan, Jennifer Donohoe, John Davidson, Virginia Sprague and participants at various seminars held in Canberra, and for the research funding provided by AusAID's Office of Development Effectiveness.

CONTENTS

List of Boxes

List of Figures

List of Tables

TAKING DEVELOPMENT ACTIVITIES TO SCALE IN FRAGILE AND LOW CAPACITY ENVIRONMENTS

Laurence Chandy
Johannes F. Linn

EXECUTIVE SUMMARY

Fragile states present one of the greatest challenges to global development and poverty reduction. Despite much new learning that has emerged from within the development community in recent years, understanding of how to address fragility remains modest. There is growing recognition that donor engagement in fragile states must look beyond the confines of the traditional aid effectiveness agenda if it is to achieve its intended objectives, which include statebuilding, meeting the needs of citizens, and managing risk more effectively. Current approaches are constrained by relying heavily on small-scale interventions, are weakened by poor coordination and volatility, and struggle to promote an appropriate role for the recipient state.

Scaling up (i.e., the expansion, replication, adaption and sustaining of successful policies and programs in space and over time to reach a greater number of people) is highly relevant to fragile settings, both as an objective and as a strategic approach to development. As an objective, it reinforces the logic that the scale of the challenges in fragile states demands interventions that are commensurate in purpose and equal to the task. As a strategy, it encourages donors to identify and leverage successes, and to integrate institutional development more explicitly into projects and programs. In addition, scaling up can assist donors in addressing the priority areas of improved project design and implementation, sustainability and effective risk management.

Successful scaling up in fragile states almost certainly occurs less often than is possible and does not always involve a systematic approach. Donors should therefore look to more systematically pursue scaling up in fragile states and evaluate their performance with specific reference to this objective. This can be done by incorporating relevant elements of a scaling up framework into operational policies, from strategy development through to program design and monitoring.

Contrary to expectations, there are compelling examples of successful scaling up in fragile states. While the conditions prevailing in fragile states create serious obstacles in terms of "drivers" (the forces that push the scaling up process forward) and "spaces" (the opportunities that need to be created, or potential obstacles that need to be removed for interventions to grow), and in terms of the operational modalities of donors, these can be overcome through the careful

design and delivery of programs with a clear focus on creating scaling up pathways, and through close partnership and sustained engagement of governments, communities and foreign partners.

Case study evidence suggests that the pathways taken to reach scale in fragile states demand different approaches by donors. Donors need to adopt greater selectivity in determining which areas or sectors for scaling up are justified—a strategy that has also assisted some donors in managing risk. More investment and time are required in upfront analysis and building the evidence for successful scaling up pathways. In some cases, donors require longer time horizons to achieve scale, although demand from government or beneficiaries has sometimes forced donors to move immediately to scale, allowing little or no time for piloting. Regardless of the pace of scaling up, donors that were most successful were engaged early and then remained engaged, often far beyond the replication phase of scaling up, to increase the likelihood of interventions being sustained. Other common characteristics of successful scaling up were simple project design and a focus on the institutional aspects of the scaling up pathway.

Case studies also point to the crucial role of drivers in moving the scaling up process forward in fragile states. Proven ideas and practical models have often been picked up in fragile states, contrary to the expectation that actors may be less responsive to recognizing and acting on the utility of promising results. Leaders undoubtedly have a role to play in supporting scaling up, although there are clear dangers that must be avoided, including avoiding the perception that donors are picking (political) winners by nominating leaders, and tying the survival of projects too closely to the fortunes of a leader's political career. Incentives were found to be one of the most important drivers in fragile states, and there is a good case to be made for donors introducing new inducements, greater transparency or similar reforms to strengthen the role incentives play. Finally, and in contrast to the standard scaling up framework, community demand was found to be an important driver in many fragile states, both in demanding the expansion of small-scale projects and by facilitating the community's own resources to support the scaling up process.

The greatest challenge to scaling up in fragile states is the limited spaces these environments provide. This is especially the case in respect to those spaces which concern aspects of governance: political, institutional and policy spaces. When working in fragile states, donors must recognize that spaces for scaling up are almost always more constrained, but look for ways to expand upon them. Some of the most successful examples of scaling up used creative approaches to build space quickly or used existing capacity to the fullest possible extent. Also relevant are the lessons of robust analysis, greater realism and cost control. The case studies confirm the importance of two additional spaces in fragile states. For example, security space often imposed horizontal obstacles to scaling up which could not realistically be overcome while ownership space served as a good indication of the perceived legitimacy of the scaling up process and the likelihood that interventions would be sustained longer term.

Case studies also affirm the importance of emphasizing robust project design and implementation, and the close linkages between the scaling up agenda and the role of risk management and sustainability in fragile states. While sustainability presented a significant problem for many of the projects and programs reviewed, a more focused approach around scaling up may assist donors in addressing sustainability concerns. This would entail adopting a longer-term perspective beyond the immediate confines of any individual project, looking for available drivers and supportive spaces, and focusing on effective implementation and consistent monitoring and evaluation (M&E). Any intervention introduced on a small scale that scores well in sustainability serves as a possible candidate for scaling up.

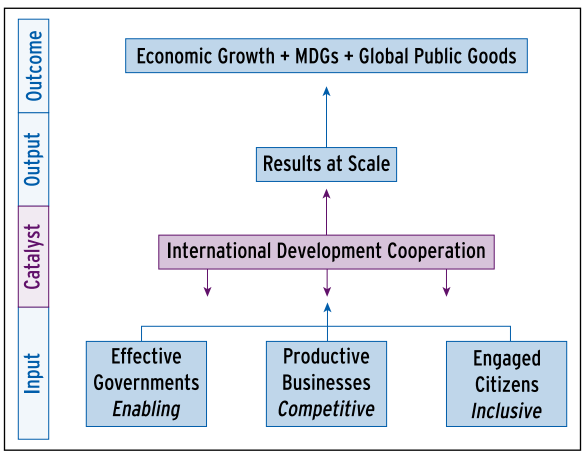
Similarly, many of the methods used by donors for managing risk—an emphasis on analysis, scenario planning, realism and making use of specialized aid instruments—are equally relevant for supporting scaling up in fragile countries. A persuasive argument can be made that the adoption of a more explicit scaling up approach by donors can form part of a risk management strategy in fragile states. Scaling up can enable donors to more ambitiously tackle development risks without allowing institutional and project risks to grow unchecked. Ultimately, a donor approach that combines good risk management and scaling up requires strong leadership and well-aligned incentives.

INTRODUCTION

The case for scaling up in fragile states[1]

If the intention of development assistance is to bring about results at scale then it stands to reason that aid interventions themselves should be designed and implemented with an eye to how they can eventually be scaled up. Hence, both the intended outputs and outcomes of aid have the notion of scale at their core (Figure 1). This same logic applies in fragile states as it is does for stable and better governed countries, as the objectives of transformational change and maximizing impact remain the same for aid in any environment.

Figure 1: The role of aid and scaling up in development



Source: Kharas et al., 2011

In practice, today's global aid system is overrun with small, fragmented interventions— few of which are replicated, expanded, or even sustained after donor support has been withdrawn. Given their limited scope, the results of these interventions are not commensurate in scale with the problem they are tasked with. In those cases where development activities have occurred at a large scale, they often have not relied on tested models or effective implementation strategies. A greater focus by donors on scaling up— translating successful pilot projects into full-scale interventions— would therefore appear warranted (Linn, 2011a).

However, while the case for more and better scaling up in fragile states is straightforward, there is remarkably little understanding as to how—and how often—scaling up occurs in fragile environments and what needs to be done to achieve it.

At least on the surface, scaling up would appear to present a much more onerous challenge in fragile states than elsewhere. First, fragile states are prone to rapidly changing conditions in the social and political environment, which encourage development actors to adopt shorter time horizons for programming.[2] In contrast, scaling up is, by definition, a long-term endeavor which requires a credible planning process into which programming decisions can be incorporated. Second, the weakness of institutions in fragile states renders more difficult the sustainability of successful interventions, especially large-scale ones that require greater levels of capacity. In particular, weak public institutions discredit the default approach to sustaining aid interventions in developing countries: passing control over to state ownership. Third, interventions in fragile states are associated with a higher rate of failure, which affects donors' perception of risk. Donors with low risk tolerance may be wary of reaching scale since this entails the possibility of large scale failure, which would likely be conspicuous and thus draw criticism. These three problems are not entirely unique to fragile states and have their counterparts in stable countries. However, they are undoubtedly more acute in fragile settings.

More generally, some of the more basic assumptions regarding how aid and scaling up work are less applicable to fragile states.[3] Most of what we know about how scaling up occurs has been learned from countries in which governments are relatively strong and it is not obvious to what extent those lessons translate over to fragile environments (Greeley, 2007). The more organic process of scaling up—whereby successful approaches in a particular sector, agency or area create spillover effects when key actors elsewhere recognize the utility of those approaches and replicate them—occurs less readily in fragile states, if at all.

While the challenges to scaling up in fragile states are evident, the danger is that attempts to moderate these can consume donors and cause them to lose sight of the very goals they are trying to achieve at a macro level: the transformation of fragile states into more stable and prosperous environments and the improved well being of their people.

Purpose and scope

The purpose of this paper is to establish a baseline of knowledge on the process of scaling up in fragile states. This is done by identifying a number of case studies of scaling up in fragile states and examining these against a scaling up framework tailored to fragile settings. The analysis is complemented by an assessment of the conditions that commonly prevail in fragile states and existing donor approaches to aiding these countries.

The empirical work in this paper is based on a desk review of program and project evaluations, each evaluation reflecting a particular instance of scaling up in a fragile setting. The research thus takes the form of a meta-evaluation in which a wide range of experiences are captured across a range of different countries and sectors.

While it is hoped that the findings of this paper will be of interest to researchers and practitioners who are interested in the topic, the paper is not intended as an operational guide for those seeking to design and implement development interventions in fragile settings. This would require further research and testing, combined with a rigorous assessment of the broader operational setting within which interventions are being conceived.

Defining fragility

"Fragility" is a loaded term which means different things to different people. In this paper, it is used to describe the condition whereby state structures fail to perform the basic functions required of them, due either to a lack of political will, capacity, or a combination of the two, causing significant challenges for development.[4]

Fragility is typically observed in terms of deficient state authority, legitimacy and/or provision of services (Stewart and Brown, 2010). On this basis, we classify countries as fragile in any one year which appear either in the "Alert" group of the Fund for Peace's Failed States Index, or which score below 3.2 on the World Bank's Country Policy and Institutional Assessment (CPIA).[5]

We deliberately do not attempt to classify fragile states into sub-categories as others do (World Bank, 2005; Moreno Torres and Anderson 2004; OECD DAC, 2007). Such attempts can give the impression that there exists a stylized linear path into and out of fragility, when there is growing recognition that if such a path exists, it is the exception as opposed to the rule (World Bank, 2011; Leader and Colenso, 2005).[6] In addition, the use of sub-categories oversimplifies the different factors that combine to make countries fragile and how these factors relate to each other. We believe that the problem of fragility is highly context specific and these contexts do not lend themselves to clean sub-classifications. Our only exception is to highlight those instances where there are security concerns associated with conflict in fragile environments and to comment on how these specifically affect our analysis.

AID TO FRAGILE STATES: SETTING THE STAGE

The increasing prominence of fragility for aid

Today, there exist around 50 fragile states, together containing more than a billion people. These countries are geographically dispersed although more than half can be found in Sub-Saharan Africa. Together, they present arguably the most urgent and intractable challenge to global development, both now and in the foreseeable future. To date, no fragile country has achieved a single Millennium Development Goal (World Bank, 2011). As global poverty levels fall, driven by progress in more stable developing countries, the share of the world's poor living in fragile states is expected to rise rapidly, from under 30 percent in 2005 to more than half by 2015.

A decade ago, donors approached fragile states with a degree of ambivalence. This position was informed by research suggesting that aid to countries with poor policies and institutions was less effective and thus represented an inefficient allocation of resources (Burnside and Dollar, 1997; Collier and Dollar, 1999). Implicit in this argument was the well-founded supposition that if policy formation is driven primarily by the domestic political economy, using aid to "buy" better policies in recipient countries makes little sense (Devarajan et al., 2001). This justified a shift toward greater policy selectivity in aid allocations and thus away from fragile states. Analysis of aid flows over the 1990s found that aid to fragile states was low, even after accounting for country performance (Levin and Dollar, 2005).

In contrast, recent years have seen a marked, but measured, increase in donor commitment to working with fragile states.[7] This partly reflects an admission by donors that the scale of development needs these countries face is too great to ignore. However, it is also indicative of a growing awareness of the negative external effects associated with fragility, conflict and state failure. Donor governments are increasingly cognizant of the cost effectiveness of preventing state collapse or conflict, and the need to balance the risks of action and inaction.

Aid effectiveness in fragile states

Greater resources for fragile states are unlikely to translate into better results without a clear understanding of how aid should be delivered effectively in such environments.

A starting point is to examine how the global aid effectiveness agenda, as defined in the Paris Declaration on Aid Effectiveness and the Accra Agenda for Action, can be meaningfully applied in fragile states. While equally relevant in stable and fragile countries, the agenda is not "straightforwardly applicable" in the latter (OPM and IDL, 2008). As originally conceived, the Paris principles rely heavily on the role of the recipient state to drive their implementation, which is unlikely to be adequately fulfilled in fragile settings. Alternative approaches to implementing Paris principles that accommodate the limitations of the state and other obstacles present in fragile environments have since been developed.

Common to these approaches are:

- A greater stress on the importance of ***analytical work*** to underpin donor policies, in recognition that fragile settings provide a more complex environment for donors to work and that poorly conceived interventions can lead to unintended consequences. Particular emphasis is given to political analysis (of the state, but also at sub-national level and of individual institutions) to understand the root cause of development failure and fragility.

- A recognition of the important roles played by ***communities and local government***, both as a stand-in for the roles typically performed by the central state in donor-recipient relations, but also as important stakeholders in service delivery whose proximity to end-users/beneficiaries provides a foundation for strengthening accountability.[8]

- ***Flexibility*** by donors, to adjust to quickly changing conditions; to act promptly when presented with successful initiatives which can be build upon, or with failing efforts that need to be altered or brought to a close; to adjust traditional approaches used in stable countries that are unfit for use in fragile settings; to work more closely with other donors; and to minimize the burden placed on recipient partners.

- The use of specially tailored ***tools and frameworks*** for approaching development planning and delivery in fragile states. These include Transitional Results Frameworks for planning, Joint Needs Assessments for needs, Consolidated Appeals Process for resources, Multi-Donor Trust Funds for delivery, and Double Compacts for accountability, among many others. These tools are similar to those employed in stable countries, but are differentiated to account for the particular challenges of delivering aid in fragile settings.

These solutions reflect important new learning on how to deliver effective aid in fragile states that have been developed in recent years. In 2007, the Organization for Economic Cooperation and Development's Development Assistance Committee (OECD DAC) endorsed the Principles for Good International Engagement in Fragile States and Situations whose application is being tracked in select countries. Following the third High Level Forum on Aid Effectiveness in Accra, the International Dialogue on Peacebuilding and Statebuilding was initiated, which has served to deepen discussions between fragile states and their donor partners, and which will report to this year's forum in Busan. Fragile states have formed a new and productive alliance through the g7+, while donors have developed a global network, the International Network on Conflict and Fragility (INCAF), to support new learning.

For the most part, these approaches are simply the corollary of thinking through how the aid effectiveness agenda can be applied in fragile states. However, in three important respects, donor engagement in fragile states must look beyond the parameters of the aid effectiveness agenda if it is to achieve its intended objectives. Specifically, donors must supplement their focus on effectiveness with an overarching concern for statebuilding, a commitment to respond to the needs of citizens, and a different approach to managing risk.

Concern for statebuilding

Statebuilding (as opposed to poverty reduction or economic development) is defined as the central objective of aid to fragile states by the OECD DAC (OECD DAC, 2007). Statebuilding involves fostering the relationship between state and society, and increasing the state's fulfillment of its rightful role. In other words, donors are expected to tackle the causes of fragility itself, rather than just addressing its symptoms.

Concern for the viability of the state means acknowledging that development failings in fragile states are only one part of a bigger puzzle that typically includes security and political challenges. While fragile states are often characterized by poverty and low levels of capacity, it is a lack of stability, confidence and legitimacy in political institutions which is at the root of their development challenge. Stability, confidence and legitimacy can be fostered by securing political settlements between elites which provides the space for institutions to function, by deepening buy-in around those settlements and by establishing greater accountability around institutions' core functions.

Where development is only one of many problems, it is likely only one part of the solution. Thus, development actors are required to work closely with their diplomatic, defense and humanitarian counterparts to achieve policy coherence and to manage the complementarities and tensions that link their interventions. This is especially important on goals where there is clear overlap, such as supporting peace agreements, enhancing capacity for security and preventing the escalation of conflicts.

Coordination with non-development actors represents a challenge to donors who struggle to coordinate effectively among themselves. Whole of government approaches provide a framework for such coordination on a bilateral basis. Some of the planning tools used to engage in fragile states deliberately seek to integrate development, diplomacy and defense objectives and activities. Poor coordination is partly to blame for the unsatisfactory handling of transitions between different phases of support to fragile states. This problem is exacerbated by the false assumption that the trajectory of fragile states traces a linear path from poor to good performance and from war to peace, with traditional tools and approaches designed for the beginning and end of each transition, rather than the middle ground which the majority of fragile states occupy (Leader and Colenso, 2005).

The corollary to a focus on statebuilding is more development programs focused on improving governance. Governance interventions often fail to be adequately grounded in reality, nowhere more so than in fragile states. There are three important lessons here. The first is to adopt realistic targets and timeframes within which governance improvements are expected to occur. A useful benchmark for setting these is what other countries have been able to achieve in similar circumstances (Clemens, 2004; Pritchett and de Weijer, 2011). The second lesson is to place emphasis on the notion of good enough governance (Grindle, 2004). This implies focusing on the core capacities of the state, steering clear of politically or socially sensitive issues that can consciously be avoided, and making full use of existing capacity. In addition, it stresses that institutions earn legitimacy by demonstrating their ability to deliver, rather than merely changing their form (Porter et al., 2011) and therefore accountability mechanisms should be built around changes in actual performance and outcomes, rather than the implementation of reforms. The third lesson is to shift attention toward sub-national levels of government, and the roles and responsibilities that distinguish different levels of government, since overall state effectiveness is only as good as the weakest link from the national to local level.

Responding to the needs of citizens

The presence of humanitarian and unmet basic needs in a country reflects the impact of conflict and/or failure of the state to adequately provide for its citizens. While in most circumstances, donors can provide support to the state or local level authorities to perform their functions more adequately, in the most dire circumstances, donors are faced with little choice but to stand in for these authorities and to provide direct assistance to people—or to encourage NGOs to do the same. While such actions are surely justified (and typically demanded by those whose needs are unmet), their legitimacy is undermined without the articulation—preferably from the outset—of a clear path toward handing responsibility back over to the state. Such a transfer cannot occur, however, until institutions are ready to lead. Donors have to balance the short-term pressure to meet citizens' needs with the long-term objective of enabling the state to perform its rightful role.[9] The longer donors occupy the role traditionally occupied by the state, the further the social contract between state and society is eroded.

Just as donors seek to reinforce the role of the government in service delivery, they similarly look to the state to mediate with its citizens to understand society's needs and priorities. Donors' demand for consultation with civil society to ensure broad-based ownership of aid strategies and projects is therefore normally facilitated by the host government. This, however, relies on an adequate rendering of democratic governance in the country, whereby governments can credibly claim to represent their citizens' views and are capable of soliciting and aggregating diverse opinions.

In fragile states, democratic governance may be dysfunctional or absent, and particular groups within society disempowered. In these circumstances, donors are often obliged to go directly to citizens rather than via the state. In fact, donors often look to consult more extensively than usual in fragile states in recognition that citizens' needs and priorities may be harder to discern. Donors are also encouraged to look out for pockets of exclusion created by societal cleavages and institutional failings.

Donors' focus on communities and civil society in fragile states does not simply reflect the need to substitute for a reliable state partner, but may also be explained by the durability of informal institutions and societal ties, which, while not always immune from the effects of fragility, are often sustained through periods of instability or conflict and thus may serve as building blocks upon which larger state structures can be formed.[10] A focus on communities and civil society can also serve to strengthen societal resilience when conflicts are ongoing and state failings are unable to be effectively addressed. Finally, community participation in civil affairs and development activities can enhance the protection and sustainability of local assets in conflict-affected environments.

Different approach to managing risk

Risks are inherent in donor activities, regardless of where they take place. However, in fragile states the types of risk facing donors are more varied and the level of risk more elevated than those encountered in stable countries. Since the challenges which aid is tasked with tackling in fragile states are significant, donors cannot simply limit themselves to the least risky interventions. A commitment to fragile states obliges donors to identify ways of expanding the space in which they can work and the programming opportunities available to them, while being honest in their assessment of the challenges posed by the environment.[11] In other words, donors must find an approach to risk which is not just *prohibiting* but also *enabling* (OECD DAC, 2011a).

The OECD DAC (OECD DAC, 2011a) distinguishes three sets of risks relevant to any donor agency: developmental risks, which concern the probability of a deterioration in the material condition of the country or sub-region in which the agency is operating; project risks, which concern the probability of a donor intervention failing to meet its stated objectives; and institutional risks, which concern the probability that the standing and integrity of the donor agency may be compromised in some way, specifically among its stakeholders but also among the development community. Donors (and their staff) may value these risks differently; it is often assumed that donors are especially concerned with institutional risks, including fiduciary risk in the case that they are unable to adequately account for expenditure, and reputation risks such as those that emerge when donors choose to align with particular elites whose own reputation is liable to criticism.

An important insight is that these different risks are interdependent. For instance, setting very modest objectives for donor interventions in a country program will succeed in minimizing project risk, but will reduce the likelihood that donor engagement will succeed in reducing the development risks facing the country. Similarly, a donor's choice of aid instruments can help to limit institutional risks but may simultaneously undermine the capacity and visibility of the state in delivering services, which reduces its legitimacy in the eyes of the country's citizens. Given the relationship between these different types of risks, donors are encouraged to develop strategies that look at these risks alongside each other, rather than treating each separately.

If value for money in aid is defined as a function of transaction costs, the expected outcomes of an intervention and project risk (DFID, 2010a), then delivering aid effectively in fragile states almost certainly implies some risk taking, while at the same time learning to better assess and manage risks. Indeed, risk management and analysis are part of a donor's obligation to use resources efficiently and to demonstrate results.[12]

There are various possible components to a donor's strategy for risk management and analysis. Donors can adopt a portfolio approach whereby country programs contain a balance of more and less risky interventions. Donors can invest in better analytical approaches so that risks can be identified and accurately gauged. This may include tracking short-term and leading indicators which are separate from the standard set of indicators monitored in traditional development plans and performance frameworks. Improved analytics can be combined with scenario planning, so that donor staff are forced to reflect on and think through the consequences of a range of possible outcomes that may eventuate, and to anticipate their own response. These sorts of exercises can be undertaken at a whole of government level to enrich the quality of the exercise and to highlight the role different agencies play in responding to changing conditions in a country. Of course, such activities incur transaction costs of their own and so are only worthwhile to the extent that such costs are offset by the benefits of improved risk management. These tradeoffs must be carefully balanced to ensure that risk management activities deliver overall gains in value for money and contribute to more efficient outcomes.

A donor's assessment of risk is reflected in its choice of aid instruments or modalities. These choices determine what aid can realistically hope to achieve, and in extreme cases, can alter the nature of the problem that aid is tasked with addressing. For example, in some of the most uncooperative environments, donors have been forced to operate under the banner of humanitarian work in order to gain the consent of, and to deliberately bypass, a hostile state, despite the needs in the country and the activities undertaken being strictly developmental.[13] A reliance on humanitarian aid leads to assistance being highly projectized, commodity–based, and short-termist in regard to staffing and funding. If employed over a longer time span (for which humanitarian aid is not suited) such an approach causes distortions which can undermine the basis for development in the future (DFID, 2005).

New instruments and modalities provide donors with the means to manage risks more effectively. For instance, the fiduciary risks associated with budget support can be moderated by use of partial earmarks, budget oversight and tracking, and requirements that the recipient government contract out certain functions, such as procurement, financial management or audit services. These approaches have enabled donors to provide budget support in complex environments such as Afghanistan, Burundi, Central African Republic and Guinea Bissau (World Bank, 2011). Some forms of aligned aid have in-built accountability mechanisms that help moderate risk, such as community-driven development programs (where communities monitor expenditures and the quality of investments, often supported by NGOs) and "dual key" mechanisms, whereby government processes are governed by boards comprised jointly by national and international actors.

Aid trends in fragile states[14]

While the notion of effective aid to fragile states is increasingly well defined, the reality of how aid is typically delivered in these environments is often very different. Here, we examine recent trends in aid to fragile states, focusing our attention on four key characteristics of aid that are closely linked with the scaling up agenda: the size of donor activities; coordination and cooperation among donors; the stability of aid flows; and approaches that recognize and reinforce the role of the recipient government.

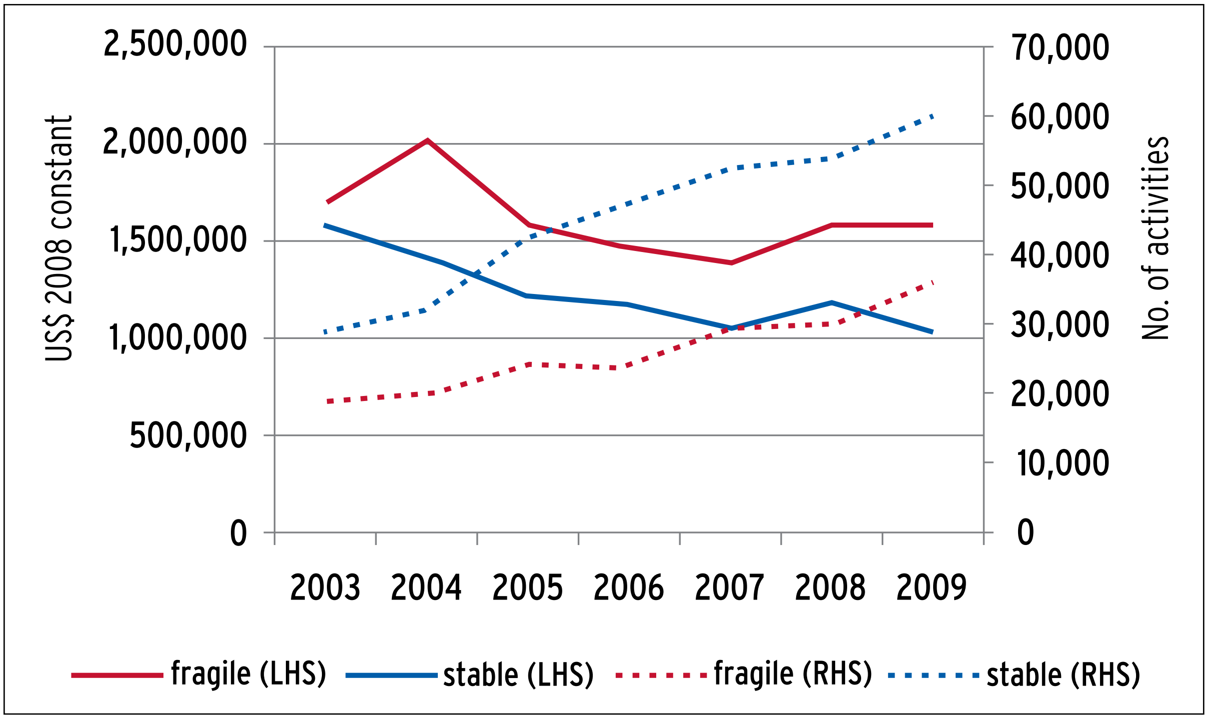
Activity size

An inclination toward bigger activities reflects both the scope of donors' ambition and their capacity to administer complex, transformative interventions. The resources available to fragile states are significant (at the aggregate level) and rising, having almost doubled since 2003 from $21.7 billion to $42.8 billion in 2009.However, these resources are fragmented into tens of thousands of individual activities, with a mean size of only $1.5 million (Figure 2). Since the mean activity size has remained relatively constant in recent years (in contrast to aid to stable countries for which the mean activity size has fallen to approximately $1 million) the problem of fragmentation in fragile states has only worsened with the rise in total flows fully reflected in a rising number of individual activities. The average fragile state today receives more than 750 separate donor interventions per year.

Part of the explanation for this trend is that donor aid programs have become overrun with small activities. Half of all aid activities in fragile states today are valued at under $80,000, following a precipitous fall in the median activity size over just the last few years. While small activities can play an important role, not least by serving as pilots to test new interventions which may eventually be brought to scale, they also raise transaction costs for both donors and recipients. Moreover, such a large number of small interventions means that only a tiny percentage could ever feasibly be brought to scale.

It would be wrong to assume, however, that the large number of small activities in fragile states crowds out larger-scale interventions in financial terms. Despite their frequency, small activities are of such a low budgetary value that they account for a very small share of total aid flows. Where there is a crowding out effect, however, is in terms of the finite technical and administrative resources expended on the development and management of donor activities. This effect is felt within donor agencies but its effect is greatest and most damaging on the recipient side, not least in fragile states where small-scale activities distract government staff and burden planning systems which typically operate under severe capacity constraints.

Figure 2: Mean activity size and number of activities



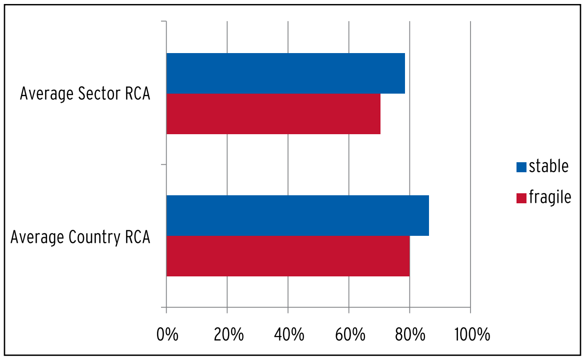
Source: Author's calculations based on OECD DAC 2011b

Coordination and cooperation

High levels of fragmentation suggest that donors are failing to coordinate effectively with one another. On the latest count, a typical fragile state will receive aid from 65 different donor agencies that report to the OECD DAC.[15] This is seven more donors than operate in the average stable developing country. The presence of large numbers of donors in any one fragile state is indicative of an aid system in which aid agencies are proliferating, donors are spreading themselves thinly and operating in environments where they lack specific expertise. The average fragile state receives a smaller proportion of its aid from donors who are meaningfully invested in the country than the average stable developing country (Figure 3).[16] Similarly, a larger share of aid is spent in sectors where donors have little demonstrable experience when compared to aid received by stable countries. This implies that donors are failing to abide by any meaningful division of labor. Donors would be more inclined to scale up and more capable of facilitating the process if they focused their attention on particular countries and sectors which they judged as priorities and in which they had specialized knowledge.

One means of increasing levels of expertise is by having donors work together. This can also support scaling up by allowing donors to pool resources and risk. Levels of cooperation among donors remain very low, however, even on simple tasks such as conducting analysis and missions. Donor cooperation appears marginally higher in fragile states than in stable countries, perhaps reflecting the partial use of common frameworks and pooled funds.

Figure 3: Share of aid received that reflects donors' revealed comparative advantage[17]



Source: Authors' calculations based on OECD DAC 2011b

Stable flows

Stable aid flows enable longer-term development planning and specifically can secure funding for aid that spans the duration of individual projects. In contrast, volatile aid flows tend to dictate programming decisions rather than the other way around. The effect of aid volatility on scaling up may be particular limiting in fragile states if it is the case that results take longer to materialize in these environments (McGillivray, 2005). Aid volatility may not only inhibit donors' capacity to scale up, but may deter governments from attempting large-scale interventions, since volatile aid can distort the composition of public investment (Desai and Kharas, 2010).

A number of studies (Pallage and Robe, 2001; Bulir and Hamann, 2006; Kharas, 2008) have shown aid flows to be highly volatile. While some research has found that the volatility of aid is the same in stable and fragile states, other research suggests that aid to fragile states may be twice as volatile as aid to stable countries (Kharas, 2008; Levin and Dollar, 2005).[18] Regardless of any difference in the degree of volatility, it is evident that the effects of aid volatility are more significant in poorer countries (since aid flows represent a larger share of the economy) and those countries whose domestic institutions are less capable of compensating for uneven donor activity. Both are characteristics common to fragile states.

Sustainable approaches that reinforce the role of the recipient government

A golden rule of aid is to strive for results that can be sustained. The default strategy for sustaining results is typically to work through government, but as discussed earlier, this is not always possible in fragile states while responding to the needs of citizens. At the same time, if the long-term objective is to build up the state, then donors must at least aspire to working with recipient institutions over time.

One relevant measure is to look at the share of aid that is "country programmable" and thus directly supports development projects and programs in partner countries. Country programmable aid serves a similar purpose to recipient governments' own investment spending and so has a greater capacity to reinforce recipients' own development agenda. The Country programmable aid share is approximately three-fifths in fragile states, compared to four-fifths in stable countries.

Three additional indicators examine the extent to which aid employs government systems and institutions. Project Implementation Units (PIUs) represent freestanding structures established by donors that stand outside the government apparatus. On average, another PIU is created in fragile states for each $20 million of aid received. Only 64 percent of aid to fragile states is recorded in recipient budget documents—the key institution for organizing public investments and overall development planning. Only half of aid to fragile states uses government public financial management and procurement systems, while the other half relies on donor systems instead. On each of these indicators, the use of government systems and institutions is marginally higher in fragile states than in stable countries—a result that is contrary to expectations and evidence that governance challenges need not impede donor cooperation with recipient governments.[19] Nevertheless, the overall scores remain unsatisfactory and indicative of an aid system that places insufficient onus on sustainability.

Two final indicators concern the use of program-based aid, whereby interventions within the same sector are integrated around a country-driven strategy. Program-based aid is believed to be more conducive to scaling up than traditional project-based approaches (Linn, 2011a), yet only 40 percent of aid to fragile states is program-based. Part of the explanation for this low figure is that budget support, which is the most common form of program-based aid, is often deemed impractical in fragile states, given that recipient budget systems lack credibility. It is interesting that other forms of program-based aid, such as pooled funds and sector-wide approaches (SWAps), are being employed more often in fragile states than in stable countries, presumably to compensate for the difficulty of budget support.[20] This provides a means of working around the challenge of deficient budget systems without defaulting to project-based approaches.

A SCALING UP FRAMEWORK FOR FRAGILE STATES

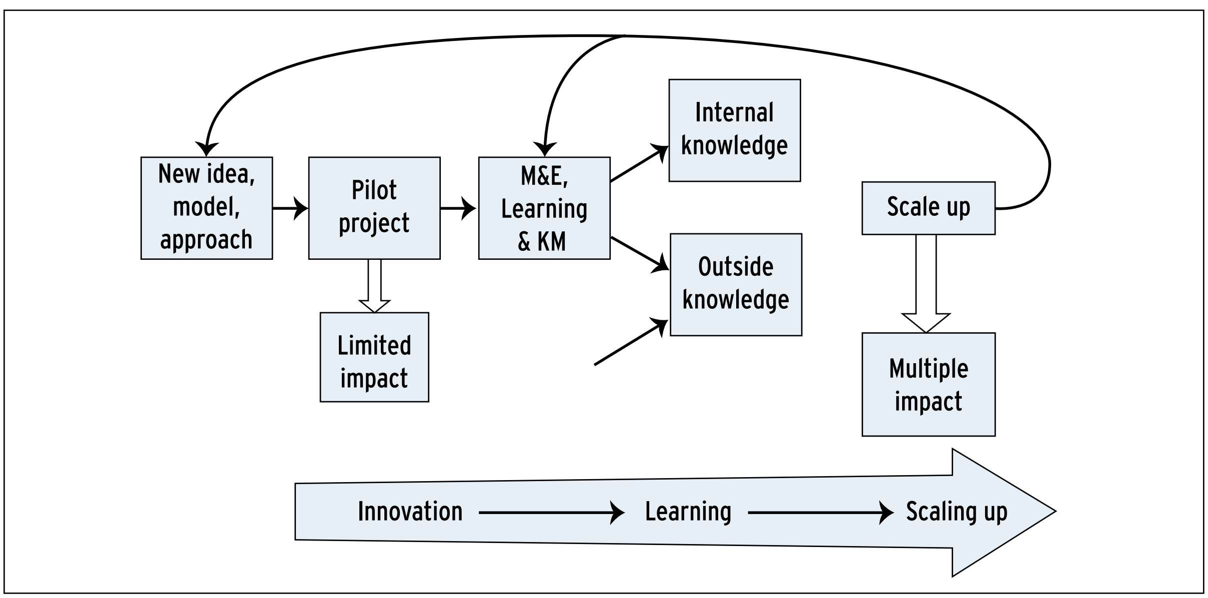
The previous section identified the salient aspects of aid to fragile states. In this section, we use this knowledge to develop a scaling up framework that is relevant for fragile settings. The purpose of the framework is to improve understanding on how the scaling up process occurs in fragile states, and to assist development analysts and practitioners (especially those representing donor agencies) in incorporating a scaling up perspective into their work.

Rather than starting from scratch, our framework is an adaptation of the framework first developed by Hartmann and Linn (Hartmann and Linn, 2008) which we tailor to the special circumstances of fragile states as described in the previous section. Our approach is to describe briefly core elements of the Hartmann and Linn framework and for each element to explain how its application must be altered to make it better suited to fragile contexts.[21]

Scaling up: introducing the concept

Hartmann and Linn define scaling up as *"expanding, replicating, adapting and sustaining successful policies, programs or projects in geographic space and over time to reach a greater number of people."* Scaling up is part of a broader process of innovation, learning and scaling up (Figure 4). A new idea, model or approach is typically embodied in a pilot project with limited impact. By learning from this experience with monitoring and evaluation, organization-internal knowledge is created and organization-external knowledge is disseminated. This knowledge in turn can be used to scale up the model through expansion, replication and adaptation with multiple impacts. The experience from scaling up feeds back into new ideas and learning. Outside knowledge can also feed scaling-up efforts, if an organization picks up on the pilot experience and learning of another organization.

Figure 4: The links of innovation, learning and scaling up



Source: Linn, 2011a

Hartmann and Linn make a number of observations with regard to the innovation-learning-scaling up process, each of which is equally applicable to fragile states:

- *Innovation, learning and scaling up should be treated as separate, albeit linked processes.* Each of the three concepts refers to an important separate stage in the development of an intervention at scale, and each requires its own appropriate process, skills, resources and attention. Innovation and scaling up are often complementary, but there are also times when they compete in terms of resources, managerial attention, and/or political pay-off.

- *Development actors need to focus not only on innovation, but also on learning and scaling up.* The focus on innovation is endemic in the aid industry and the development business, usually to the detriment of an adequate focus on learning and especially on scaling up.

- *The innovation-learning-scaling up process is not linear, but an iterative and interactive cycle.* As indicated in Figure 4, there are many feedback loops from learning and scaling up back to innovation. Indeed, monitoring and evaluation often generates new ideas for better design and implementation and the scaling up process will often require adaptation and innovation in the way the original model or idea is brought to scale. Nonetheless, it is useful to think of the three main components of the process as distinct and separable phases, each of which has its own important role to play.

- *Not every innovation can or should be scaled up. Not every scaling up needs to involve an innovation.* Many innovations may not be suitable for scaling up. It is precisely the experimental nature of the innovation process that needs to be recognized as important it its own right and the risk of pilots not succeeding must be accepted as an integral part of the innovation and learning process. Failed pilots may offer as many and perhaps more lessons than successful ones. But of course, failed or unsustainable pilots should not be scaled up. By the same token, not every scaling up effort needs to involve an innovation.

- *Scaling up involves two types of possible errors: "type 1 error"—no or too little scaling up; and "type 2 error"—wrong scaling up.* Much attention in the scaling up literature is on the prevailing lack of attention to scaling up—a "type 1 error". This is most typical with smaller aid organizations, which cannot hope to scale up interventions with their own limited resources, and hence tend to focus on small and limited projects with limited impact. But there are also cases where aid organizations support scaling up or operate at a large scale, but do so sub-optimally or in a wrong way—a "type 2 error". The latter type of error is found more frequently in the large development banks, which can and often do aim to go to scale in their country strategies and programs.

Pathways for scaling up

The Hartmann and Linn framework places a particular focus on the notion of scaling up pathways. A "pathway" is the sequence of steps that need to be taken in the innovation-learning-scaling up cycle to assure that a successful pilot or practice is taken from its experimental stage through subsequent stages to the scale ultimately judged to be appropriate for the intervention pursued. In general, there are many possible pathways for scaling up a successful intervention. For each case an aid organization needs to explore potential pathways early on and take proactive steps to plan and prepare for scaling up—in terms of dimensions, desired ultimate scale, drivers and spaces (see later), the agency's operational modalities, intermediate results, and monitoring and evaluation. In practical terms, this means developing a strategic approach to the aid intervention by developing a country, sector or subsector strategy, in which the scaling up pathway is defined and the role of the project or intervention in helping the country move along the pathway is clearly identified.

Selecting the "dimensions": Scaling up pathways can follow different dimensions: They may simply expand services to more clients in a given geographical space. They can also involve "horizontal" replication, from one geographic area to another; "functional" expansion, by adding additional areas of engagement; and "vertical" up-scaling, i.e., moving from a local or provincial engagement to a nation-wide engagement, often involving policy dialogue and technical assistance to help achieve the policy and institutional conditions needed for successful national level scaling up.

Defining the desired scale: It is important to define up-front the ultimate scale to which an intervention should or could be taken, given the needs of the target population and the nature of the intervention, and to consider realistically the time horizon over which the scaling process needs to extend in order to achieve the desired ultimate scale. Hartmann and Linn found that successful scaling up of programs to national scale can take five to 10 years, or longer.

Defining intermediate results: Along the scaling up pathway, it is important that the program delivers intermediate results. This is necessary to allow for the testing and, where needed, adaptation of the approach. But it also helps with ensuring the buy-in of community, government and other stakeholders (Schaffer and Ashkenas, 2005).

Selecting operational modalities for scaling up: Donor agencies have various options for applying their operational modalities in supporting the pathways for scaling up:

- They can use their own resources for scaling up (top-up, repeater projects, programmatic approaches), work in partnership with other agencies (co-financing, SWAps) or hand-off to other donors, the government or non-governmental partners.
- They can finance investments, provide technical assistance or engage in policy dialogue.
- They can scale up an intervention within a country or across countries.

Putting in place monitoring and evaluation: M&E are key ingredients of a successful scaling up strategy in various important respects: first, during the implementation of the pilot or experimental stage, the intervention needs to be monitored to learn what are the drivers and spaces (opportunities and constraints) that may affect an eventual scaling up process and the impact of the pilot in terms of the lives of the rural poor need to be evaluated (preferably against a control group); second, during the scaling up process monitoring will provide important feedback on any unforeseen aspects of the scaling up pathway and permit the adaptation of the pathway as needed. Intermittent evaluation of the impact of the scaled up program during implementation and after completion is needed to ensure that the expected results actually materialize.

There is an important feedback loop from scaling up to M&E. Evidence shows that most donor and recipient agencies do not make a serious effort in implementing effective M&E of development projects. One explanation for this is that when the focus of M&E is only on the project, as opposed to the scaling up pathway of which the project is one part, M&E does not contribute much to the success of the project by itself. However, once project managers buy into the importance of the scaling up pathway, and recognize the contribution that learning from project experience can make to the successful pursuit of such a pathway, the benefits of

M&E become more apparent and project managers will hence have an incentive to develop effective M&E processes (Linn et al, 2010).

Risks of inadequate attention to scaling up pathways: A number of problems can result from not paying due attention to scaling up pathways, including the following:

- opportunities for scaling up may be missed ("Type 1 error") or scaling up may be done badly ("Type 2 error");
- not paying attention to costs may create "boutique" approaches that only work at a small scale;
- setting up special purpose entities (e.g., project implementation units) rather than working through ministries may limit institutional options later;
- working with limited financing mechanisms, not identifying policy constraints and working with small implementing partners (such as NGOs) may limit the potential for scaling up later; and
- lack of effective, timely M&E may lead to poor decisions in scaling up.

Adaptation to fragile states

In fragile settings, the sequence of steps taken to reach scale is likely to vary in a number of important respects from those taken in stable environments. We identify eight principles reflecting these differences, each of which are described briefly below:

Employ greater selectivity determining in which areas or sectors scaling up is justified. In stable settings, there are few if any ground rules for determining in which areas or sectors scaling up may be pursued. Donors can be active in any area where there exists significant development needs, as identified by national development strategies and reflected in country assistance programs. Wherever donors identify a small-scale intervention that has demonstrated success, the possibility of scaling up is in principle worth pursuing. Fragile settings, at least in the most extreme cases, are different. While development needs can be numerous, donors must focus on areas that are key to securing stability and improving the viability of state. The 2011 *World Development Report* describes these succinctly as "security, justice and jobs". One might add to these other core functions of the state, such

Box 1: The Monrovia Roadmap

At the June meeting of the International Dialogue on Peacebuilding and Statebuilding in Monrovia, the g7+ group of fragile states and their development partners defined five international peacebuilding and statebuilding objectives. According to the group, progress against these objectives is "necessary to reduce and prevent conflict and human suffering and to reach the MDGs in situations of fragility and conflict". It is therefore intended that the objectives be used to guide priorities at the national level and to direction international assistance. The five objectives are as follows.

- Legitimate politics: fostering inclusive political settlements and conflict resolution
- Security: establishing and strengthening people's security
- Justice: addressing injustices and increasing people's access to justice
- Economic foundations: generating employment and improving livelihoods
- Revenues and services: managing revenues and building capacity for accountable and fair social service delivery

Source: International Dialogue on Peacebuilding and Statebuilding, 2011 as public financial management. While the difficulty of working in fragile states may tempt donors to scale up "anything that works" or to work in any area where they can find an enthusiastic partner, a better strategy may be to focus attention on areas that are most important to overcoming a country's fragility (Box 1). Selectivity is also important because of the limited institutional capacity in-country and the high costs of operating for donors.

Conduct more upfront analysis and place more onus on evidence. It is widely accepted that donor interventions in fragile states are more prone to failure. Part of the explanation for this is that fragile environments can be more complex: the policy process is less easy to follow, government failures are multiple and pervasive, and donors lack a coherent "theory of change". Donors cannot expect to successfully scale up interventions without an adequate grasp of the environment in which they are operating and confidence that the interventions they are sponsoring genuinely work. While there is an inevitable experimental element to any scaling up process, donors need to invest in upfront analysis and put greater emphasis on the role of evidence to manage these challenges.

Use scenario planning in anticipation of changing conditions and outcomes: While Hartmann and Linn encourage donors to maintain a vision of an intervention's ultimate scale, they also stress that there is no scaling up blueprint and the need for flexibility. In fragile states, donors should be prepared not only for varied outcomes from the scaling up process, but changing country conditions, such as a sudden deterioration in governance which directly affect sponsored interventions. Scenario planning can be a useful tool for thinking through different outcomes and how donors will respond accordingly.

Accept longer time horizons, be prepared to take slower and smaller steps and adapt the pathway as needed at each step in the light of experience. Since conditions in fragile states can change quickly, donors may choose to adopt shorter programming cycles, typically of two years. Yet interventions themselves are likely to take longer to develop, to scale up and to sustain. Collecting analysis and evidence, as described above, take time. Each phase in a scaling up pathway may take longer to complete and the pathway may need to be broken down into more phases so that the expansion process can be adequately managed and adapted as necessary. This is consistent with lessons from the field that stress the need for greater realism in fragile states. Where there is a strong humanitarian case to roll out services quickly, donors may choose to scale up over a much shorter timeframe than can readily be sustained by existing institutions. In these instances, the timeframe for scaling up and achieving sustainability may markedly diverge. It is nevertheless important that consideration of both timeframes be incorporated into program design and closely monitored throughout implementation.

Compensate for higher cost by ensuring designs are kept simple. A higher rate of failure, more upfront costs and a slower scaling up pathway all translate into higher costs for delivering aid. One way in which donors can moderate costs is by ensuring project designs are simple. This also increases the likelihood that interventions can be successfully replicated and sustained by partners.

Focus on institutional aspects of the scaling up pathway and look for institutional innovations. Working in fragile states obliges donors to work alongside weak or unreliable institutional partners. Along the scaling up pathway, institutional partners can struggle to exert effective ownership, to benefit from the process by learning, retaining knowledge and conducting M&E, and to sustain interventions initiated with donor support. These weaknesses need to be monitored and adjusted for. Where donors identify effective strategies for improving institutions or for working alongside less reliable institutional partners, these strategies themselves can be replicated in other areas of donor engagement.

Integrate transitions across different phases of support and/or aid instruments into scaling up pathways. Poorly handled transitions between different phases of support (i.e., security, humanitarian, development) can pose a threat to the sustainability of long-term scaling up endeavors, just as certain aid instruments (e.g., PIUs, short-term contracts) can impose a limit on the scope and sustainability of interventions. However, with the right preparation and vision, interventions can be designed in such a way that early stages serve as a stepping stone on a scaling up pathway. In its report on service delivery in fragile states, the OECD DAC recommends that "aid programs should be structured from the start in such a way as to promote a timely transition to sustainable service delivery and accountable governance," and specifically "government ministries should be active participants in planning and program development. Contracts for non- state providers should include transition planning and allow the alignment of state administrative practice, donor systems and humanitarian principles." (OECD DAC, 2008b, pp. 40, 42)

Use horizontal and vertical scaling up to overcome rather than reinforce obstacles. In most fragile states, other than the smallest countries, there exists considerable variety in the stability of different regions and the capacity of regional governance. This type of landscape lends itself to a particular pathway for horizontal scaling up, whereby interventions are tested in more stable, better governed areas before being gradually rolled out in increasingly complex settings. While seemingly logical, this approach can result in the creation of "islands of excellence" while areas that are the most fragile are altogether avoided. Similar constraints can occur in vertical scaling up, when donors focus their attention at a particular level—communities, local government, central government—with whom they can reliably partner. Yet an exclusive approach of this kind can easily falter: national programs that are poorly anchored in local or community demand and implementation modalities are less likely to take root, while an exclusively local-level focus may fail to achieve institutional development in the sense of embedding rule-based systems, organizational coherence and complete accountability chains. A focus on transformation change forces donors to respect obstacles to horizontal and vertical scaling up without avoiding them. This is part of a broader point that donors should not allow scaling up pathways to be defined purely by where scaling up is most feasible, but equally by what is an appropriate vision of an intervention's scale.

Drivers

According to Hartmann and Linn, a key consideration in determining a scaling up pathway is the role of "drivers". These are the forces that push the scaling up process forward. Hartmann and Linn identify four common drivers:

- *Ideas and models*: There has to be an idea or model that works at a small scale. These may emerge from research or practice.
- *Vision and leadership*: A vision is needed to recognize that scaling up of an idea is necessary, desirable and feasible. Visionary leaders or champions often drive the scaling up process forward.

- *External catalysts*: Political and economic crises or pressure from outside actors (donors, EU, NGOs, etc.) may drive the scaling up process forward.

- *Incentives and accountability*: Incentives are key to drive the behavior of actors and institutions toward scaling-up. They include rewards, competitions and pressure through the political process, peer reviews and other evaluations. Monitoring and evaluation against goals, benchmarks and performance metrics are essential ingredients to establish incentives and accountability.

Adaptation to fragile states

The identification and assessment of drivers to propel the scaling up process is arguably especially important in fragile states since scaling up pathways take longer to complete and can quickly stall when faced with obstacles. Strong drivers may be particularly important in overcoming particular challenges, such as when a country is transitioning between different stages of external support. Each of the four common drivers is relevant to fragile states, although their role may be modified in certain ways:

Ideas and models: Recent years have seen a justified push-back against the notion of international best practice and blueprints in aid, in recognition that replication rarely works without tailoring interventions to country circumstances. Nevertheless, tried and tested models have a role to play, especially in fragile settings where donor experience and knowledge remains limited and results difficult to achieve. The 2011 *World Development* Report argues that reliable models exist for delivering better outcomes in the core areas of justice, jobs and security. Other areas where new models have proven successful in one setting and since been replicated in other fragile states include social protection, and disarmament, demobilization and reintegration (DDR) programs.

Vision and leadership: A vision that a better, more peaceful, just and prosperous future is possible, with benefits of reconstruction and recovery delivered widely across different segments of the population, is critical to turn around the cycle of conflict, destruction, lack of capacity and despair in many of the worst affected fragile states. Individual and charismatic leaders with a clear vision and the capacity to organize a response can prove very effective in driving the scaling up of a particular project and overcoming obstacles to progress. Their skill is often to make use of informal institutions (and to understand how formal and informal institutions interact) and to leverage personal networks—in essence knowing "how things get done" in a particular country setting. Valuable though such individuals can be, there are two major risks to such a strategy. First, choosing leaders in an environment where settlements between political elites are unresolved or fragile may be interpreted as "picking winners". This can be divisive and counterproductive for addressing a country's development problems. Second, the links between a single leader and a project can be hard to pry apart once an association has been formed. This means that the individual and the project rise and fall together. If the leader withdraws from his/her role or loses influence, the scaling up of the project is undermined. Equally, if the leader's legitimacy erodes over time, as often occurs in the political bargaining required to achieve piece in fragile states, donors may feel obliged to abandon the project. Given these concerns, Pritchett and de Weijer argue that interventions in fragile states are more likely to succeed when they are driven by multiple leaders at once, rather than a single player. "It often takes many people, situated in many different positions and with impartial agreement and commitment—to drive reform.…Activities have to be flexible enough that many different people are committed to success and feel they are leaders of at least some component of the overall endeavor." (Pritchett and de Weijer, 2011: p39.)

Local NGOs and civil society organizations are of special significance in post-conflict and fragile settings for initiating, spearheading and sustaining successful scaling up processes.

External catalysts: Country ownership is often poorly manifested in fragile states so it is little surprise that interventions tend to rely more heavily on donors to advocate for and help implement the process of scaling up. Nevertheless, there are limits to how effectively donors can drive the scaling up process without a thorough grasp of local politics and especially in the later stages of the scaling up pathway when attention shifts to the question of sustainability. International NGOs and UN agencies, which often have a longer-standing presence and deeper local grounding than other donors in fragile settings, can be especially helpful if they are participating in a scaling up process in partnership with other donors and local stakeholders. In all these cases, the term "catalyst" is appropriate since external actors cannot substitute for national drivers.

Incentives and accountability: Incentives and accountability can certainly play an important role in driving scaling up in fragile states and there may be a case for introducing new inducements, greater transparency or similar reforms to strengthen the role they play. However, donors should first take the time to better understand the existing incentives operating on actors—both government and donors—which can be complex in fragile settings. Likewise, since fragile states typically suffer from weak systems of accountability, donors require a credible theory of change if they are to attempt to alter the environment.

In addition to the four common drivers described above, we identify two additional drivers which may lend momentum to scaling up pathways in fragile settings:

Unmet needs: A commitment by donors to respond directly to the needs of citizens in fragile states builds a strong moral and logical case for developing programs at scale, capable of serving large populations. While scaling up pathways may take longer to navigate in fragile states, the persistence of unmet needs can serve as a regular reminder of the need to expand coverage to the greatest extent possible. In contrast to the humanitarian sector, development actors should be wary of scaling up interventions in which they substitute for the state without a clear sense of how the state can be incorporated back into a leadership role once it is capable.

Community demand: Communities can serve as a powerful actor in driving forward the scaling up process, both by demanding the expansion of small-scale projects and by committing their own resources to support the scaling up process. Irrespective of the dysfunction of national level institutions in fragile states, communities often enjoy efficient forms of democratic governance which enable them to organize and to identify community priorities. Their simultaneous role as project beneficiaries means they are well placed to monitor services and to account for project delivery. Where appropriate, communities can contribute human capital—sometimes referred to as "sweat equity"—to support the scaling up effort and to convey their stake in successful rolling out.

Spaces

Another important concern in determining the feasibility of a scaling up pathway is the role of "spaces". These are the opportunities that can be created, or potential obstacles that need to be removed to open up the space for interventions to grow. Hartmann and Linn identify eight spaces in their framework:

- *Fiscal/financial space:* Fiscal and financial resources need to be mobilized to support the scaled up intervention; and/or the costs of the intervention need to be adapted to fit into the available fiscal/financial space.

- *Natural resource/environmental space:* The impact of the intervention on natural resources and the environment must be considered, harmful effects mitigated, and beneficial impacts promoted.

- *Policy space:* The policy (and legal) framework has to allow for, or be adapted to support, scaling up.

- *Institutional/organizational/staff capacity space:* The institutional and organizational capacity has to be created to carry the scaling-up process forward.

- *Political space:* Important stakeholders, both those in support and those against the intervention, need to be attended to through outreach and suitable safeguards to ensure the political support for a scaled up intervention.

- *Cultural space:* Possible cultural obstacles or support mechanisms need to be identified and the intervention suitably adapted to permit scaling up in a culturally diverse environment.

- *Partnership space:* Partners need to be mobilized to join in the effort of scaling up.

- *Learning space:* Knowledge about what does and does not work in scaling up needs to be harnessed through monitoring and evaluation, knowledge sharing and training.

Adaptation to fragile states

An unavoidable inference in applying the concept of "spaces" to fragile states is that they are almost always more constrained. This is especially the case in respect to those spaces which concern aspects of governance: the political, institutional and policy spaces. Since scaling up in fragile states usually involves engagement with the state to some degree, this presents a serious challenge.

Thankfully, there has been considerable learning in recent years on how to deliver aid more effectively in fragile states, much of which is implicitly concerned with how to overcome small governance spaces. Governance interventions are an important part of most country programs in fragile states and can usefully be integrated with the scaling up of any particular intervention. Of course, spaces are not static and opportunities may arise for donors to scale up as and when spaces open up. At the same time, donors can endeavor to create necessary spaces over time, as part of the design and implementation of scaling up pathways.

The recognition that spaces for scaling up in fragile states are often small serves as a reminder of many of the lessons that emerge from our earlier analysis of scaling up pathways in fragile states. These include the need for robust analysis, scenario planning, greater realism, cost control, and slower and smaller steps.

We identify two additional spaces which may be relevant to analyzing scaling up in fragile states:

Ownership: Ownership by national stakeholders might ordinarily be thought of as a driver in scaling up, reflecting elements already captured in the drivers of leadership and community demand. However, in fragile states it is arguably more relevant as a space. Ownership is needed to legitimize the scaling up process since interventions originally conceived by external actors are unlikely to ever achieve a strong sense of local ownership, thus reducing the likelihood of their longer-term sustainability. The challenge is that ownership is rarely manifested strongly in fragile states. Donors often look for little more than a superficial buy-in from elites, which counts for little in an environment where political settlements are weak. As explained earlier, ownership need not necessarily involve the state, but can be engendered by civil society, communities and/or sub-national government. The key to a successful scaling up pathway is to explicitly consider how to develop ownership over time. Ownership may be manifested in multiple stakeholders by engaging them incrementally and by designing and implementing appropriate accountability and social justice mechanisms.

Security space: Donor activity may be precluded in the most unstable environments where there is open conflict. This may impose horizontal obstacles to scaling up that cannot be readily overcome. Nevertheless, some donors have succeeded in finding ways of continuing their work by borrowing lessons from the humanitarian sector.[22] A successful scaling up pathway will take explicit notice of the evolving security conditions and needs, and adapt interventions accordingly. For instance, security-related aspects of program design may be gradually phased out as and when the probability of conflict recedes.

Three overarching priorities: design and implementation, sustainability and risk

Moving beyond the Hartmann and Linn framework, we identify three themes which are central to any discussion of scaling up in fragile states: design and implementation, sustainability and risk. Each theme has synergies with the scaling up agenda that we explore briefly below. An important inference is that a scaling up approach provides a framework that can help improve design, manage risk and buttress sustainability.

Design and Implementation

As discussed previously, development assistance in fragile states differs from assistance to stable countries in terms of both what it is intended to achieve and how it is delivered. Questions concerning the design and implementation of assistance are central to the fragile state setting (Box 2) and are closely linked to the scaling up agenda with its emphasis of a longer-term, programmatic approach to aid management.

The frameworks and instruments used to deliver aid in fragile states may prove useful for planning scaling up pathways. For instance, joint needs assessments can serve as a useful entry point to identify where scaling up may be required and consolidated appeals processes can be explicitly organized around the required financial space that is required to reach a desired scale. At the same time, assessments carried out rapidly will likely miss aspects of the local context, so designs may need to be altered over time. Similarly, shorter- and longer-term funding vehicles should be linked so that early investments can be built upon over time.

Other tools may need to be partially reengineered if they are to support scaling up objectives. For instance, the "Double Compacts" proposed by Ashraf Ghani and Clare Lockhart offer a promising approach for distinguishing government and donor roles and responsibilities in a way that promotes mutual accountability, without undermining each actor's accountability to other stakeholders, such as citizens, parliaments and auditors (Ghani and Lockhart, 2008). During a scaling up pathway, the balance between government and donor roles and responsibilities is typically expected to change over time, as donors initially drive the process forward but governments are expected to take on greater leadership through latter phases of planning and execution. Double Compacts could potentially be employed to manage this transition by linking compact

> **Box 2: Pritchett and de Weijer on the "Middle Way"**
>
> Pritchett and de Weijer advocate for a more explicit focus by donors on design and implementation issues in fragile states. They identify the limits of donors' standard approaches which they characterize as being of two types. The first are large programs which achieve scale, scope and speed, but at the expense of having virtually no—or worse, a perverse—impact on the capability of the state. The second are small projects, which have the virtue of being flexible, innovative and adaptable, all of which are important in fragile settings and in enabling aid to achieve local success, yet are incapable of bringing about systemic change.
>
> Pritchett and de Weijer wonder whether there might be a "middle way": one that "has the benefits of systemic impact and scale but the virtues of flexibility." They continue: "the starting point of a 'middle way' might be to refine accountability around outcomes, where outcomes include the capability of the system." For this, "one needs a combination of outcome targets that includes measures of both immediate outputs and outcomes but also how equipped the system is to continue to deliver those into the future." (Pritchett and de Weijer, 2010, p. 38)

Thus, changes in accountability would be triggered by the completion of phases in a scaling up pathway. Such an approach would not only allow accountability to serve as a driver in scaling up, but would help to embed scaling up into broader governance arrangements.[23]

Taking another example, donors often employ "Transitional Results Frameworks" in post-conflict settings to organize development planning over relatively short time horizons. These frameworks typically last for two years and are organized into six-month segments, around which specific goals and reporting requirements are organized. Linking a framework to the scaling up pathway of a particular intervention would appear relatively straightforward, but would require explicit consideration of how those pathways that extend beyond the timeframe of the framework would be affected, to ensure that interventions carry over to future planning strategies.

Designing and implementing institution building and capacity building is an integral part of any scaling up framework in a fragile state context. At the outset of an initiative, local capacity may not be sufficient to allow progress and hence may require donor-led implementation involving special project implementation units. However, as part of a scaling up pathway, institution and capacity building can be brought to bear to assure that over time the appropriate combination of community, local and higher-level government, and non-governmental institutional capacity is created and can carry forward the scaling up process in an effective and sustainable manner.

Sustainability

Sustainability is a central problem of delivering aid to fragile states and is the reason why many interventions fail. The problem of sustainability is inextricably linked to the weakness of formal institutions, limited state capacity to generate and manage revenues, and the changeable conditions that define fragile settings. However, it is also reinforced by donor approaches that lack recipient ownership, are too ambitious or gold-plated, and fail to adequately understand how local institutions work. On occasions, donors may attempt to deliver interventions that are recognized ex ante as unsustainable in their urgency to provide for citizens' needs, or in an effort to secure peace. But more often, donors approach interventions with an eye to their long-term sustainability and yet these efforts nevertheless fall short.

The objectives of scale and sustainability are closely linked. Scaling up can be conceived as a two stage process: replication or rollout first, and consolidation and sustaining second.[24] There is little point attempting the first stage unless donors have confidence that the latter stage is achievable. For donors that have struggled to achieve sustainability in most of their interventions in a particular country, the objective of scaling up may therefore appear daunting. A scaling up approach stresses two different points: first, donors should question whether those interventions that proved unsustainable were attempted at the appropriate scale to enable learning in the first place; and second, donors should build on—and learn from—those interventions that have proved sustainable. Thus, any intervention introduced on a small scale that scores well on sustainability serves as a possible candidate for scaling up. Given this close link between sustainability and going to scale, it is not surprising that many of the factors which make scaling up possible are also those that support sustainability: a longer-term perspective beyond the immediate confines of the project, effective drivers and supportive spaces, good implementation and consistent M&E.

Risk

Donors' approaches to managing risk in fragile states are closely connected with strategies for scaling up. Many of the methods used by donors for managing risk—an emphasis on analysis, scenario planning, realism, making use of specialized aid instruments—are equally relevant for developing a scaling up approach in fragile countries. These similarities are perhaps to be expected since both strategies are concerned with expanding the scope of donors' activities in an environment which is otherwise limiting.

In addition to this general connection, a persuasive argument can be made that the adoption of a scaling up approach in a donor's country program (i.e., the inclusion of some interventions with an explicit scaling up pathway) can form part of a risk management strategy in fragile states. To explain this, we use the risk framework described earlier, which distinguished three different types of risks: developmental risks, project risks and institutional risks.

In fragile states, donors typically imagine a tradeoff between developmental risks on the one side, and project and institutional risks on the other. That is, any attempt to use aid more determinedly to tackle the root challenges of fragility forces donors to accept a higher likelihood that project objectives will not be met and to adopt approaches that expose the donor agency to a greater reputational risk.

Scaling up attempts to moderate this tradeoff, enabling donors to more ambitiously tackle development risks without allowing institutional and project risks to grow unchecked. Project and institutional risks are moderated by donors employing greater selectivity in determining which interventions to scale up and basing those decisions on evidence of what works. Interventions that are identified for scaling up are not only those that have delivered against their objectives at a smaller scale, but ones that have scored strongly across the board in evaluations on measures of relevance, efficiency, effectiveness, impact and sustainability are built upon a strong analytical foundation and have strong backing by donors and recipients. This implies that only a selection of interventions within a country program should adopt a scaling up pathway. This is consistent with the Hartmann and Linn framework which stresses that not all interventions should be scaled up. Furthermore, it echoes the idea that country programs should adopt a portfolio approach, combining interventions with higher and lower risk ratings, and those with and without scaling up elements.[25]

Institutional (reputational) risk may well be deemed higher when managing large-scale projects, given the possibility that their failure—however remote—will draw attention by virtue of their size. However, the common perception that smaller projects carry less project risk is very likely mistaken. For a given amount of aid, the risk per project in a portfolio of many small interventions may well add up to a larger total risk than for fewer projects that get systematically scaled up over time, precisely because the scaling up process is built around a stronger evidence base of what works, greater familiarity with the institutions and actors, and the potential for developing their capacities. Moreover, the scope for managing risks proactively is much greater in the case of a carefully designed and adapted scaling up process than in the case of a proliferation of fragmented one-time interventions.

Ultimately, a donor approach that combines good risk management and scaling up requires strong leadership and aligned incentives. The incentives for donor staff in most aid agencies are perceived as being very narrowly defined in terms of risk. The combination of a clear preference for institutional risk avoidance at the senior management level and the often misunderstood nature of project risk at the level of mid-level managers and project leaders can be a key factor biasing institutional incentives against scaling up and in favor of many smaller one-off projects. For donors to move beyond the standard approach to risk management requires strong leadership that understands the benefits of scale. Good leadership will ensure that staff are not stymied by a fear that they will be blamed should development interventions go awry, but will be rewarded for achieving success at scale.

LESSONS: A REVIEW OF SCALING UP CASE STUDIES IN FRAGILE STATES

Methodology

We now turn to an assessment of project and country program case studies to determine what lessons can be derived for scaling up approaches in fragile states. This review takes the form of a meta-evaluation, drawing on existing formal evaluations and reviews of selected donor supported programs. Such an assessment is helpful because it allows us to pit our hypotheses in the preceding section against the reality on the ground. And perhaps most importantly it permits a reality check for what might be the most troublesome question a skeptic might raise, namely: Is scaling up appropriate and feasible in fragile states?

We rely on two types of evaluations: first, evaluations and assessments of specific programs and projects funded by selected multilateral, bilateral and non-governmental donors in fragile states; and second, evaluations of institutional experience at the country-level in fragile states, drawing on evaluations for the World Bank, the Asian Development Bank (ADB) and the Global Fund for HIV/AIDS, Tuberculosis and Malaria (GFATM). In both cases, the focus is on examples that involve scaling up in one form or another. Annex 3 and 4 provide detailed accounts of the different case studies and evaluations referenced in this section.

Project and country program evaluations by their very nature cannot provide conclusive answers to questions about scaling up, let alone statistically significant tests of hypotheses. Rather, they represent snapshots of program implementation at a given time.[26] The evaluations are shaped by the methodology employed by the implementing agency and the judgment of its evaluators, neither of which may be comparable across agencies or even across programs supported by the same agency. Evaluations generally do not involve controlled experiments that would test interventions against a counterfactual in a rigorous manner. And since most evaluation offices operating in the development community have yet to include scaling up systematically in their evaluation methodologies, it is not possible to draw on standardized evaluation results.[27]

The *project evaluations* that we draw on allow a reasonable assessment of the design of scaling up pathways and drivers as they support successful scaling up experiences. They generally, however, offer less insight into failures to scale up, the challenge of limited spaces and the institutional constraints of donor organizations. By contrast, *country program reviews* permit a more comprehensive assessment of success and failure and provide particularly good insight on spaces. These are less helpful in exploring specific scaling up pathways or the drivers which push successful scaling up forward. Taken together, the range of evaluations used for this paper allows us to draw useful preliminary conclusions that can be exposed to further examination in due course, whether through additional case studies or through the process of implementing specific projects and programs.

Since we are testing the various program and project case studies against the scaling up framework described in this previous section, here we follow a similar structure of examining the pathways, the drivers and the spaces to achieving scale, before reflecting on the important role of program design and implementation, sustainability and risk management.

The inventory of evaluations

From an original collection of more than 100 evaluations of development activities in fragile states, we

Table 1: Distribution of projects and programs by region, country, sector and agency

	Health	Education	Local/Community Development	Agriculture/Rural Development	Other Sectors	Country Programs
Asia						
Afghanistan	USAID/EC/DA/ WB, GFATM				WB	
Cambodia		UNICEF	UNDP	ADB		WB, ADB
Nepal						ADB
Pakistan						ADB
Timor-Leste	WB				MD	WB
Africa						
Ethiopia		SCN			WB/DFID	
Kenya	USAID	DFID/WB				
Mozambique			UNDCF/UNDP			
Uganda						WB
Middle East						
Yemen			DFID			
West Bank-Gaza						WB
Europe / Central Asia						
Bosnia Herzegovina						WB
Tajikistan	GF/UNDP		UNDP			
Latin America						
Peru				IFAD		
Global / Regional						WB, ADB

Legend: ADB = Asian Development Bank; DA = Danish International Development Agency; DFID = Department for International Development; EC = European Commission; GFATM = Global Fund for HIV/AIDS, Tuberculosis and Malaria; IFAD=International Fund for Agriculture; MD = Multidonor; SCN = Save the Children Norway; UNCDF = UN Capital Development Fund; WB = World Bank

identified and reviewed 17 project or program case studies that involved various degrees of successful scaling up. In addition, we reviewed a number of recent country program evaluations and cross-country reviews of operations in fragile states by the World Bank Group's Independent Evaluation Group (IEG) and the ADB's Independent Evaluation Department (IED). Table 1 shows the distribution of evaluations by region, country, sector and agency. With cases covering five regions, 14 countries, and 11 donors (eight multilateral donors, three bilateral donors and one international NGO), the sample of projects and programs is widely dispersed across multiple dimensions.

Pathways

The case studies demonstrate a variety of different scaling up pathways. Some programs started off as pilots (e.g., the USAID funded health program in Kenya, the Seila local development and education programs in Cambodia, and the Afghanistan DOTS-TB program). Others moved to national scale directly (e.g., the

Box 3: "Classical" scaling up pathways for a health program in Kenya

The *Kenya Adolescent Reproductive Health Program (KARHP)*, supported by USAID, aimed to improve knowledge about reproductive health and encourage a responsible and healthy attitude toward sexuality among adolescents; to delay the onset of sexual activity among younger adolescents; and to decrease risky behaviors among sexually active adolescents. It followed a "classical" scaling up pathway: starting with an innovative idea, implementing experimentally through pilots, evaluating the pilots before replicating the intervention geographically in multiple sequential steps, and sustaining these over a long period to reach national scale. The approach involved both horizontal scaling up (geographic replication) and vertical scaling up (advocacy and policy engagement at the national level). No functional scaling up (e.g., expanding the coverage of health areas or target age groups) appeared to have been attempted.

The interventions were implemented through existing governmental structures, networks and systems. Working with government ministries ensured sustainability of the interventions through integration of activities within routine operational plans, protocols and policies. It has also ensured ownership and the prioritization of issues concerning adolescent sexual reproductive health, including through national budgetary allocations. Key to the ministries' adoption of KARHP was their initial involvement in the planning and design of the model. The use of evaluation evidence was critical to both convince the ministries and USAID of the effectiveness of the model and to improve and adapt the original interventions within existing government structures. In sum, institutional, ownership, learning and partnership spaces were effectively created to allow scaling up to take place.

Source: Annex 3

Timor-Leste health program and the Ethiopia social transfer program). Among the former, the Kenya health project followed what might be characterized as a "classical" scaling up pathway by careful and deliberate sequencing from pilot to larger rollout and eventually to national scale (Box 3).

Regardless of the ultimate scope of each project and program, some salient aspects can be identified from the scaling up pathways pursued.

Scale and dimensions of projects and programs

Most of the projects reviewed reached nation-wide scale and involved horizontal and vertical scaling up. A few (e.g., the area development program supported by ADB under its Agricultural and Rural Development program in Cambodia) reached only regional or sub-regional scale. The degree of engagement in nation-wide policy dialogue and institution-building differed across programs. The World Bank and ADB virtually always attempted to scale up vertically because of their focus on policy and institutional reform, but even the NGO supported education project in Ethiopia involved some support for policy formation. Of particular interest is the scaling-up experience of the GFATM since it represents an innovative approach to scaling up that is increasingly emulated by other vertical funds (i.e., single-purpose funding mechanisms with very explicit scaling up objectives) (Box 4).

Selectivity, simplicity and flexibility

Successful scaling up appears to be more readily achievable when donors are selective in the programs they sponsor, and in the objectives and interventions they pursue. In addition, it helps when the process

Box 4: Scaling up experience of the Global Fund

The GFATM has supported the scaling up of programs to fight HIV/AIDS, tuberculosis and malaria in developing countries since 2002. A recent evaluation published in *Global Health Governance* in 2010 reports on the performance of 198 grants to 41 fragile states with a total disbursement of $4.9 billion.[28] Contrary to earlier evaluations that had found no significant differences between GFATM operations in fragile and stable states, including the GFATM's own report (GFATM, 2005), this evaluation concludes that there were significant differences between the two types of country programs.

With 40 percent of its funding going to fragile states, the GFATM is heavily invested in these countries. Overall, GFATM programs in fragile states perform remarkably well in terms of reaching their targets, their disbursement rates and extension rates between different phases of support. Performance has been less strong for M&E, data quality and grant extension beyond the first five years. Across all performance categories, fragile states performed somewhat less well than other states and countries affected by humanitarian crises scored even less well. Some two-thirds of grants in fragile states have to be administered by external (multilateral) organizations, compared to one-third in stable settings, reflecting the relatively weak development of local institutions.

While fragile states, as expected, create greater challenges than other countries for the scaling up strategy of the GFATM, successful scaling up does take place in four-fifths of fragile states. Three areas are identified as deserving special attention, however: data quality and reporting systems; M&E; and country ownership and administration by national agencies.

Source: Annex 4

of implementation is kept as simple as possible. The World Bank's evaluation of engagement in fragile states observed that the World Bank's standard processing procedures often created serious delays. Streamlined operational procedures for processing emergency operations were a major factor in allowing the World Bank to engage in Bosnia-Herzegovina and Timor-Leste much more quickly than under ordinary circumstances, and ultimately to reach scale faster. In Bosnia-Herzegovina, this did not affect the quality of operations and fiduciary aspects negatively, while in Timor-Leste, the evaluation attributed significant quality and fiduciary problems to the fact that short-cut procedures were applied. Streamlined operational procedures were identified as a problem for the ADB in Afghanistan and Timor-Leste in terms of the quality of project preparation and the effectiveness of fiduciary standards.

A large number of the case studies specifically mentioned the importance of flexibility in project preparation and implementation. The GFATM represents an interesting case in this regard, judging from the Tajikistan case study. While initially regarded as exceptionally non-bureaucratic, flexible and speedy in its procedures, over recent years the GFATM has lost some of its advantage over other donors, as increased processing and fiduciary requirements were added in response to the perception of excessive performance, fiduciary and reputational risks.

As these examples show, donors face many tensions and dilemmas in fragile states. More generally, as noted in the World Bank evaluations in Bosnia-Herzegovina, West Bank and Gaza and Timor-Leste, there is frequently a tension between quick delivery of financial support and pursuit of long-term development outcomes. The World Bank's evaluations conclude that too much attention was devoted to crisis response with detrimental impacts on long-term development outcomes. In Timor-Leste, the evaluation attributed this to the intense pressure on the World Bank by local and international politicians (as well as by other donors) to show quick impact in the face of humanitarian emergencies and destruction of essential infrastructure. Tensions and dilemmas were noted for the ADB in terms of the tradeoff between speed of response and quality of preparation.

Analytical work and M&E

Many of the project and country program evaluations stressed the importance of preparatory analytical work to ensure effective understanding of the country context, including political and cultural dimensions. A good example of this is the Mozambique Decentralized Planning and District Finance Project supported by United Nations Development Program (UNDP). The evaluation of pilots and the measurement of progress along scaling up pathways were part of a systematic approach to M&E in many of the successfully scaled up projects. For example, the careful evaluation of pilots was noted in the USAID-assisted Kenya health program, and the measurement of progress of the GFATM-supported projects in Tajikistan and Afghanistan are notable.

The country program evaluations tend to paint a more negative picture on M&E than the project reviews. The cross-country evaluation of World Bank activities describes M&E in fragile states as "negligible". This was particularly notable in the World Bank projects of Timor-Leste, where the bank's own evaluation identified inadequate M&E and insufficient learning from one project to the next in the sequence of projects as a major problem. The ADB's Cambodia and Nepal evaluations characterized M&E as weak and insufficient, respectively.

Continuity of engagement

Sustained engagement was clearly a major contributing factor in cases of successful scaling up. All the projects reviewed for this study involved a longer-term horizon in donor engagement for a sequence of steps in the scaling up pathway, often involving multiple projects or financing tranches over five to 10 years. In the case of the Peru Highland Area Development Program, the International Fund for Agricultural Development (IFAD) has pursued a scaling up strategy over 30 years. The case of ADB's agriculture and rural development program in Cambodia provides an interesting contrast: the evaluation praises the series of policy reform and rural infrastructure projects for their scale and effectiveness, but finds that instead of staying engaged in this areas of successful horizontal and vertical scaling up, ADB switched to other lines of engagement (namely, irrigation development and targeted rural community development programs) with untested modes of intervention, limited scale and relatively poor results. The evaluation recommends that "ADB should build on and consolidate past successes within the sector, up-scaling similar designs and implementation arrangements to other parts of the country particularly in terms of expanding rural infrastructure, firming up maintenance and building institutional capacity." (IED, 2009c)

In West Bank and Gaza, the World Bank was involved over 20 years and stayed engaged even during periods of crisis. Nevertheless, its engagement in some sectors was intermittent. In Bosnia-Herzegovina, the systematic sequencing of projects was identified as a strength of the World Bank program, but the dramatic reduction of its engagement after the initial years of exceptional high financial support was seen as a major risk for the continuity of support. More generally, the World Bank's overall evaluation of engagement in fragile states also expressed concern with its tendency to prematurely

Box 5: The cluster school program in Cambodia

The cluster school concept was a potent idea and proved to be a highly effective vehicle for scaling up. Each cluster school is connected administratively to a group of five to 10 other schools, spread across the four geographical areas of the country. This enabled the program to ultimately reach 30 percent of primary school students in the country. Clusters were used to demonstrate educational changes in policy, management, curriculum, textbooks and teacher training, and to build support among local educators for reforms. In a country with precious few human and material resources, the cluster model provided a way to share expertise among several schools and a way to disseminate ideas and training through an organized school network. It also provided an authorized means for villagers, teachers and students to meet each other and was an approach that was supported by the sector's authorities. The cluster system may have also served to redevelop the bonds within the country's network of educators, which undoubtedly had been weakened during the conflict and whose strengthening likely contributed to the project's sustainability.

Other important elements of success included the involvement of communities in the project, effective communication with civil society, provincial and local authorities, a technical rather than political focus around reform, solid institutional development with transparent results metrics and effective M&E, and input from international experts who could bring to bear their experience with comparable efforts from elsewhere in the region

Source: Annex 3

down-scale resources once an immediate post-conflict situation had passed. While longer-term engagement is critical for successful scaling up in fragile settings, donors appear to find it difficult to stick with programs for extended periods. Among the many challenges highlighted in a number of country evaluations are the frequent turnover of project team leaders and the lack of adequate field presence of expert staff.

Drivers

The project case studies provide a broad range of examples and insight into the role played by drivers in the scaling up process.

Ideas and models

Ideas and models can be powerful drivers for scaling up. The idea can be new, as was the case with the education program in Ethiopia sponsored by Save the Children Norway, which propagated an alternative basic education concept for children out of school, or the UNICEF cluster school approach in Cambodia (Box 5). Alternatively, ideas may be imported from specific examples in other countries, such as the case of the World Bank micro-finance program in Afghanistan, which drew upon the experience of a similar, success program in Bosnia-Herzegovina. Or the idea can be one that reflects a globally agreed approach, as in the case of the GFATM supported programs.

Leadership

Many successful scaling up experiences can be traced back to strong leadership. In some cases this can be a single individual as in the case of the Yemen Social Fund (Box 6), or in others, a small group of leaders as in the case of the Kenya Adolescent Reproductive Health Program. The Yemen case shows that there are also risks associated with strong leadership.

Box 6: The Yemen Social Fund: An example of leadership in action

The *Social Fund for Development (SFD)* is an independently funded quasi-governmental institution, which was established in 1997. It works with local communities across Yemen to finance various social development projects, assist local institutions to develop their capacity and efficiency, and create new employment opportunities co-financed with the private sector. SFD has proven itself to be an enormously successful vehicle for scaling up across several dimensions (and a way for DFID to positively influence some of Yemen's weakest sectors). SFD's scaling up successes have taken place in an otherwise very difficult environment.

While operating partially outside and parallel to the structure of government, it is directed by the deputy prime minister (who is also the minister of planning and international cooperation). Strong political leadership has undoubtedly been an important driver in SFD's expansion and its success in scaling up particular interventions across the country. At the same time, SFD's governance structure is very reliant on the influence and support of political appointees, including the deputy prime minister. There is a clear danger that the SFD could get caught up in Yemen's political crossfire. However, it is hard to imagine any governance structure which would make SFD immune from political developments in the country, given the country's vast governance challenges and SFD's willingness to work closely with the government's development plans.

Source: Annex 3

Incentives and accountability

Perhaps the most important driver is that of incentives and accountabilities for scaled up results. There are various examples of such incentives among our case studies. The most obvious incentive is financial support for specific community projects, as provided under a number of projects that serve the development of local capacity and community participation (e.g., the Seila program in Cambodia, the Mozambique Decentralized Planning and District Finance Program, and the Tajikistan Community Program). In the case of the Peru area development project, the grants were combined with a competitive element, which provided incentives for innovative and effective service delivery at the community level. In the case of Afghanistan Basic Health Services Program, incentives were provided to NGOs contracted for service delivery in underserviced areas (Box 7). Finally, a number of programs—among them those funded by the GFATM—relied on project performance metrics monitored under the programs' M&E system to serve as indicators of progress which in turn determined future funding levels.

Community demand

Programs that respond to community demands have a built-in driver for achieving scale. As the pilots are rolled out, communities not yet benefiting may learn about the beneficial impact of a program, especially if supported by an effective communication strategy. This in turn means that pressures will be put on program administrators to move faster and further in scaling up the program concerned. The Cambodia Education project, the Primary Health Project in Ethiopia and the Peru Highland Area Development Program are prime examples of such "community driven" scaling up processes.

Spaces

The program and project evaluations highlight the important role of spaces in the scaling up process.

Fiscal and financial space

The World Bank and ADB country program evaluations highlighted fiscal and financial space as a major issue in all countries, particularly in terms of the ability of projects and programs to continue in the absence of external aid. The World Bank's efforts to support increased fiscal and financial capacity in most cases did not show much progress over the time horizon of the programs evaluated. The ADB's evaluation of engagement in fragile states called for systematic effort to help close the fiscal and financial resource gap.

Our two project examples from Kenya demonstrate the challenge of scaling up in a country with a politicized and unpredictable budget process. In 2010, a successfully scaled up deworming project in Kenya was undermined by the inability of the Kenyan government to adequately account for its use of donor funds. By contrast, part of the success of the Kenya Adolescent Reproductive Health Program was attributed to the deliberate effort to minimize the project's fiscal cost and to use an interagency process to build support for the initiative across government.

Institutional space

Institutional space was also perceived as a major constraint in virtually all projects. According to the country program evaluations, the World Bank expended significant effort to strengthening institutional capacity in all cases, but with mixed success. Institutional space was also identified as a major constraint by the ADB, with weak human resources judged an important factor.

Box 7: Performance incentives under the Afghanistan Basic Packages of Health Services program

Under the BPHS, actual health service delivery is contracted to NGOs and private actors, based on a bidding process facilitated by the three main donors (USAID, European Commission, World Bank) that results in signed, time-limited 'Performance-based Partnership Agreements'. NGOs are selected competitively, with credible sanctions in case of poor performance. Although the Ministry of Public Health (MoPH) remains responsible for health service delivery in some provinces, its primary role is to develop strategies, goals and objectives, set indicators, and to monitor, supervise and control the performance of the implementing partners. By giving NGOs a fair degree of autonomy but holding them accountable for achieving national priorities, the MoPH has addressed serious constraints, such as scarce human resources, lack of physical facilities and logistical challenges. Carrying out regular, independent and rigorous M&E of health sector performance is expensive. However, it has allowed MoPH to identify problems, act quickly to resolve them and track whether progress has actually been achieved.

Around 30 NGOs are involved countrywide in delivering BPHS services through vertical programs. As of 2005, approximately 70 percent of districts were covered by the BPHS, providing primary health care to 50 percent of the Afghan population. Between 2002 and 2007, there was a 136 percent increase in the number of functioning primary health care facilities from 496 to 1,169. The health management information system indicates that there has been nearly a four-fold increase in the number of outpatient visits from 2004 to 2007. Independent assessments confirm that the quality of health care and health outcomes have also improved significantly over the period, despite a worsening security situation.

Source: Annex 3

Significantly, the projects and programs which witnessed the most successful scaling up could attribute part of their success to design innovations which enabled sufficient institutional space to be seized, despite the generally weak institutional environment. The case studies point to two broad categories of innovation: approaches which succeeded in quickly building institutions from an initially low base; and approaches which made full use of existing institutional capacity.

The former includes the Cambodian cluster school program where the cluster model enabled the rapid roll-out of new policies, management approaches, curriculum, textbooks and teacher training across schools, and the rebuilding of the network linking together the country's teachers and other stakeholders in the education sector. Another example is the Yemen Social Fund, which despite standing partially outside of the government's formal structure, has devoted much of its resources to strengthening the capacity of local institutions and to co-financing their own projects and programs, rather than going it alone.

An example of the second type of innovation is the Alternative Basic Education for Children out of School program in Ethiopia. The program was designed to make full use of each "kebele", the smallest administrative unit in Ethiopia, to build schools and allocate land for the school premises and to hire instructors, provide teacher training, set schedules and design the curriculum. The program also relied on "weredas" (administrative divisions made up of multiple kebeles) to roll out the program and to monitor its quality. A further example is provided by the Afghanistan Basic Packages of Health Services program (BPHS) which relied on the capacity of NGOs already active throughout much of Afghanistan to provide service delivery, thus relieving the burden on government whose capacity was palpably low and whose role was sensibly limited to policy, procurement and oversight.

Political space

Political space was a significant problem in many of the countries. For example, the fragmented and highly contentious ethnic politics in Bosnia-Herzegovina made program design and implementation highly complex for the World Bank. In the West Bank and Gaza, political tensions between Israel and the Palestinian Authority were a continued source of disruption, as were intra-Palestinian divisions. In Ethiopia, the evaluation concluded that the World Bank should have done more by way of political analysis to inform its programs. The cross-country evaluation urged more political analysis as an input into the World Bank's strategies for work in fragile states. Political space presented a mixed experience for the ADB: in Cambodia, political stability was cited as a positive factor, while in Nepal, political instability and labor strikes were a major obstacle.

The Alternative Basic Education for Children out of School program in Ethiopia provides an example of how political space can be created even in a hostile environment. Save the Children Norway overcame the government's distrust of, and poor relationship with, the NGO sector by scaling back its own presence in the country to demonstrate that it did not wish to take on extensive responsibility within the education sector itself and by designing a program which reinforced the role of local level government structures.

Cultural space

Cultural space was identified as an issue in Timor-Leste, where the World Bank failed to consider the local conditions adequately in grafting the successful

Indonesian model of community-driven development onto Timor-Leste's different social, cultural and political traditions. In Ethiopia, an evaluation concluded that the World Bank would have been helped by more familiarity with, and analysis of, local cultural conditions.

Partnership space

Partnership space also created challenges. The World Bank addressed these in various ways: at the level of overall aid coordination, where it often took a lead, by creating bilateral partnerships (e.g., with the African Development Bank in Uganda); and through the creation of trust funds and co-financing at the program and project level. In most cases, the evaluations give the World Bank credit for pursuing a partnership agenda with increasing degrees of intensity and success. In Timor-Leste, however, the lack of effective partnership-building led to duplication, lack of cooperation, and lack of success and sustainability of the programs supported. The World Bank's overall evaluation saw good donor cooperation at the international level, but this was not generally reflected at the country level. The ADB also made significant efforts to partner with other institutions (especially the World Bank) and to participate in or lead donor coordination efforts, but according to evaluations, intentions were generally stronger than actual achievement on the ground.

Among the creative partnerships developed to achieve scale were the collaboration of the Yemen Social Fund with local government entities and the formalization of NGO roles in the delivery of primary healthcare under the Afghanistan BPHS. The latter is a good example of how the contracting out of core state functions and services may serve to enhance rather than undermine, state legitimacy, if implemented effectively (World Bank, 2009).

Ownership space

Many of the case studies point to the importance of early government involvement and ownership for successfully scaled up programs (e.g., the Kenya adolescent reproductive health program, the Seila program in Cambodia, the Mozambique and Tajikistan local and community development programs, and the Timor-Leste Transition Support Program). Box 8 summarizes the role that ownership played in two programs in Ethiopia. Ownership was identified as a significant problem limiting the success of World Bank funded country programs in fragile states. In West Bank and Gaza, the World Bank's evaluation pointed to a positive shift in ownership from the first half of the decade to the second. In those cases where successful project results were achieved and successful scaling up was evident, ownership was clearly present and an important driver.

Security space

The World Bank's evaluation of engagement in fragile states observed that where conflict is in progress, security concerns create serious obstacles for its work. For example, in West Bank and Gaza the violence of the intifada and recurrent conflict between Israelis and Palestinians proved very disruptive over the 20 years of World Bank engagement. In Afghanistan the three project evaluations reviewed here (basic health, TB-DOTS and microfinance) noted problems created by insecurity, particularly in the worst conflict-affected areas of the country. Service delivery was either seriously impaired or impossible in those regions. However, despite these difficulties, all three programs registered successful results at a national level. Security was a significant concern for the ADB early in the Cambodia program and throughout the period of evaluation in Nepal. The ADB's evaluation of engagement in fragile states flagged that problematic security particularly affects the ADB's ability to deploy consultants on the ground.

Box 8: Program ownership in Ethiopia

In Ethiopia, Save the Children Norway (SCN) provided long-term support for the development, field testing and scale up of an alternative model for the first cycle of primary school—*Alternative Basic Education for Children out of School (ABECS)*—that is cheaper, more flexible and more compatible with local conditions than the formal schools. From the mid-1990s onward, SCN scaled down its own capacity to implement projects and relied more on partnerships with government agencies as well as a number of local NGOs for operational capacity. Specifically, it stressed integration with existing institutions; the promotion of community ownership and response to community needs; and government ownership and leadership in the scaling up process at the national and local levels. Government buy-in was enhanced by the ABECS design, which reinforced the role of government structures. ABECS schools are typically built on land allocated by local authorities, which are now the focal point for local development efforts and share in the supervision of ABECS instructors.

The government of Ethiopia's *Productive Safety Nets Programme (PSNP)* was established in 2005, initially for a five-year period, with the support of DFID and the World Bank. The PSNP reaches over 7 million people across seven of the country's 10 regions. It aims to 'graduate' people from food insecurity through an equal combination of food and cash transfers, and to build community assets via associated public works schemes. While some donors, including DFID, initially insisted on a phased approach, involving an initial piloting of the program, the Ethiopian government decided on an immediate national-level roll-out. Ownership of the PSNP is strong and helped drive the program's introduction at scale. There is a high level of political commitment to the PSNP within the Ethiopian government, including at the top levels of political leadership, where there is a belief that this program can help stimulate rural growth.

Source: Annex 3

Design and implementation, sustainability and risk

Design and implementation. While the country program evaluations in fragile states show that the conditions prevailing in these settings have created serious obstacles in terms of drivers and spaces, and significant challenges for the operational modalities of donors, specific project evaluations and some of the country program experiences suggest that these obstacles could be overcome through careful program design and delivery of the programs with a clear scaling up focus, with effective learning mechanisms, and through close partnership and sustained engagement of governments, communities and multiple foreign partners.

Each of the successful case studies relied on creative design elements that identified how scale could be brought about given the local context. For instance, the Kenya Adolescent Reproductive Health Program benefited from an inter-agency, multi-sectoral approach with integrated capacity building elements. As another example, the Alternative Basic Education for Children out of School program in Ethiopia was designed to empower local communities, to leverage existing capacity and to reinforce the role of government structures. Moreover, both these projects relied heavily on monitoring and evaluation and adaption as part of their overall implementation. A clear focus on up-front analysis and on well-designed M&E during implementation, consideration of political factors and realities, extended time-horizons and a stick-with-it attitude were all key elements of the design and implementation approach in successfully scaled up projects.

While the World Bank and ADB country program evaluations do not specifically address the need for a systematic approach to scaling up, they allude implicitly to the benefits of a scaling perspective in design and implementation. For example, the overall evaluation of World Bank engagement in fragile states concluded that there is a need for selectivity and prioritization and for a long-term perspective. Moreover, "appropriate sequencing of reforms and sufficient time to implement them are crucial for achieving results without overwhelming country capacity" (IEG, 2006a). The ADB reports contain expressions of concern about lack of program selectivity, proliferation of project numbers and declining or small project size. They call for "step-by-step" sector strategies, road maps, "gap analysis", umbrella TA operations, multi-tranche financing facilities and programmatic approaches. Furthermore, they stress the need for a long-term perspective, careful sequencing and continuity in engagement. The Cambodia country program evaluation calls for a "road map for the ADB assistance that traces the chain of results from ADB support to well-defined national targets and results" (IED, 2009b)—an apt summary of what scaling up means in practice.

Sustainability. The program and project evaluations clearly highlight the pervasive challenge of sustainability. Perhaps the most frequently cited obstacle to sustainability is the dependency of programs on aid resources. The World Bank's country program evaluations gave near universally low ratings to sustainability of programs and outcomes. The ADB ranked sustainability between likely and less likely on average, with six out of 12 cases closer to less likely. Two of the ADB evaluations note especially problems of sustainability of infrastructure investments as a result of poor maintenance. The challenge of sustainability is even clearer among the specific project evaluations. Whether it is the primary or specialized health programs supported by the World Bank or USAID, the education programs of UNICEF or Save the Children Norway, the local and community development programs and social funds supported by UNDP, DFID and others, or the agricultural and rural development programs of ADB—all suffer from the same problem: uncertainty as to whether governments are able and willing to support these large-scale programs in the absence of continued engagement by donors. This concern is particularly notable around the programs supported by vertical funds, which have received criticism for too-readily employing parallel delivery systems and putting in place incentives for workers that cannot be afforded by governments.

Other challenges to sustainability include poor governance, weak institutions, limited skilled manpower and in some countries, recurrent political instability and conflicts. Box 9 summarizes the challenges to sustainability in two particular cases. A focus on sustainability in the design and implementation of scaling up pathways and an explicit consideration of how the eventual disengagement by donors from a particular program area will affect the future sustainability of a program is therefore critical.

Risk. The evaluations that we reviewed amply demonstrate the many risks which donors face in their efforts to scale up in fragile settings. They do not, however, permit an in-depth assessment of whether and how donors have addressed these risks. The fact that many programs were scaled up successfully suggests that donors and their local counterparts were not always put off by the risks they faced and that these risks were managed through the scaling up pathways, whether explicitly or implicitly. Moreover, the case studies demonstrated that some of the common assumptions made concerning risk turn out to be incorrect. For instance, contrary to the assumption that large scale interventions incurs greater risk, DFID assessed its support to the Yemen Social Development

Box 9: Challenges to the sustainability of scaled up programs in Mozambique and Kenya

The *Decentralized Planning and District Finance Program* in Mozambique provided support for the development of planning capacity at the district and sub-district level, combined with a funding mechanism for local, small-scale, participatory public service investments. Over the years, it moved from pilot testing in a few districts to replication in many more, with the possibility of a nation-wide application under consideration by the government. However, the program in Mozambique faced significant sustainability challenges: uncertainty about future sources of financing for the substantial costs of implementation and for the district development fund in the absence of donor support; sustained engagement of qualified personnel in a context of overall human resource shortages and weak institutional capacity; and sustainability of community participation, especially if material incentives are to be reduced over time.

In one of the first randomized impact evaluations in international development, a small-scale *deworming project* in western Kenya, underwent a careful impact assessment. The evaluation showed that deworming had even greater public-health benefits than previously thought due to the positive externalities on untreated children and that the intervention presented an extremely low-cost way to boost school enrollment in western Kenya. This evidence, combined with an international campaign to disseminate results and to provide technical assistance in adopting and implementing the project, led to a massive and rapid scale-up of the intervention in 2009. However, later that year, when an audit reported missing funds from the broader multi-donor program of which the deworming project was just one part, donor support was withdrawn and the project suspended.

The case study provides a valuable lesson in the difficulties of translating evidence into policy in fragile states. In a country like Kenya characterized by bad governance but relatively high capacity, the scaling up process itself, from pilot through to large scale replication, may be less of a challenge than the sustaining of an intervention. This provides a strong case that efforts to improve governance and institutions should be incorporated into the scaling up of projects and that strong political economy analysis can support the successful completion of scaling up pathways.

Source: Annex 3

Fund as having the lowest risk rating of any of its projects in the country, despite it being the biggest project within DFID's country portfolio and the broadest in scope and objectives. More research, drawing on specifically designed case studies and an analysis of donors' institutional practices, operational policies and processes will be needed to explore this important dimension of the scaling up challenge more conclusively.

Summary of scaling up results

According to the evaluations we reviewed, all of the projects and programs achieved some notable successes in scaling up interventions to levels that generated real benefits at a regional or national scale in the countries concerned. This resulted in increased adoption of particular models of service provision, of local government and community planning and implementation, of incentive mechanisms and of learning along scaling up pathways.

That said, overall development outcomes were mixed for the World Bank country programs, with satisfactory or moderately satisfactory results recorded in three countries, and unsatisfactory or moderately unsatisfactory results recorded in three others (see Box 10 for two contrasting examples). The ADB's country program evaluations were generally positive, with a roughly equal amount of successful and partly successful ratings, according to the IED's overall evaluation.

In terms of outcomes by *sectoral* areas of engagement, the two areas that consistently proved the most difficult for the World Bank were governance and private sector development, with anti-corruption measures and improvements in the business climate proving particularly hard to secure. In contrast, macroeconomic stability, education, health and infrastructure were sectors in which positive outcomes were mostly achieved. Agriculture and rural development showed mixed results. The World Bank's evaluation of engagement in fragile states noted relatively successful outcomes in macroeconomic stability and infrastructure interventions. It also noted that community driven development projects tend to disburse fast, but generally have only limited community development impact.

Box 10: Two contrasting cases of World Bank scaling up experience

The World Bank program received a satisfactory rating in Bosnia-Herzegovina, a country that emerged in 1996 from a deadly and protracted civil war that left great destruction and deep political divisions, reflected in a fractured governance structure agreed on in the Dayton peace settlement. The World Bank's large scale engagement systematically built programs at scale in various sectors over time. Its careful and ambitious co-financing and partnership agreements, leveraging the World Bank's resources, were a significant factor making this outcome possible.

In contrast, is the unsatisfactory outcome rating for the World Bank program in Timor-Leste, albeit under very difficult conditions. Project sequences failed to be grounded in effective learning and partnerships, reinforcing the importance of designing and implementing effective scaling up pathways.

Source: Annex 4

For the ADB, governance and private sector development were rated as successful or partly successful, in contrast to the World Bank experience. The successful score for the ADB's governance efforts in Cambodia may in part be explained by the fact that the agency did not try to tackle corruption issues, which the evaluation acknowledges as being endemic. Particularly noteworthy are the successful ratings for rural credit support in Cambodia and Nepal. Agriculture and rural development showed mixed results, while support for education programs has generally been successful. ADB's support for infrastructure in fragile states also has a mixed track record; nonetheless, the overall evaluation assesses infrastructure as a sector that is particularly promising for the ADB support in fragile settings.

While the country and project evaluations are not by themselves conclusive, we believe they do provide strong evidence that scaling up is possible and an appropriate objective in fragile states, that the framework which we have adapted and used for assessing scaling up in fragile setting is applicable, and that explicit consideration of scaling up pathways and of the key drivers and spaces along such pathways is critical for success. The project cases also demonstrate that a scaling up perspective helps address common tensions which donors face in supporting the recovery of fragile states, by allowing a focus on immediate and quick results to be reconciled with the longer-term developmental and capacity building objectives.

CONCLUSION

Fragile states present one of the greatest challenges to global development and poverty reduction.

Scaling up represents a valuable objective and strategic approach for donors committed to helping these countries develop. As an objective, scaling up reinforces the logic that the scale of the challenges found in fragile states demands interventions that are commensurate in purpose and equal to the task. As a strategy, scaling up encourages donors to identify and leverage successes, and to integrate institutional development more explicitly into their projects and programs. In addition, scaling up can assist donors in addressing the priority areas of improved project design and implementation, sustainability and effective risk management.

Scaling up successes in fragile states almost certainly occur less often than is possible and do not always involve a systematic approach. Donors should look to more systematically pursue scaling up in fragile states and evaluate their performance with specific reference to this objective. The tailored scaling up framework presented in this paper provides a starting point for thinking about how donors can adopt a more deliberate and informed strategy for scaling up in these settings.

Contrary to expectations, there are compelling examples of successful scaling up in fragile states. While the conditions prevailing in fragile states create serious obstacles in terms of drivers and spaces, and in terms of the operational modalities of donors, these can be overcome through the careful design and delivery of the programs with a clear scaling up focus, and through close partnership and sustained engagement of governments, communities and foreign partners.

ANNEX 1: FRAGILE STATE CLASSIFICATION

We classify countries as fragile in any one year which appear either in the "Alert" group of the Fund for Peace's Failed States Index, or which score below 3.2 on the World Bank's Country Policy and Institutional Assessment (CPIA).

Countries are classified for each of the five years from 2005 to 2009.

We classify the following 33 countries as fragile throughout the entire period studied:

Afghanistan	Angola	Bangladesh
Burundi	Central African Republic	Chad
Comoros	Congo, Dem. Rep.	Congo, Rep.
Côte d'Ivoire	Djibouti	Eritrea
Ethiopia	Guinea	Guinea-Bissau
Haiti	Iraq	Kiribati
Korea, Dem. Rep.	Liberia	Myanmar
Nigeria	São Tomé and Principe	Sierra Leone
Solomon Islands	Somalia	Sudan
Togo	Uganda	Uzbekistan
West Bank and Gaza	Yemen, Rep.	Zimbabwe

In addition, we classify the following 31 countries as fragile in at least one of the five years studied:

Bhutan	Bosnia and Herzegovina	Burkina Faso
Cambodia	Cameroon	Colombia
Dominican Republic	Equatorial Guinea	Gambia, The
Georgia	Guatemala	Iran, Islamic Rep.
Kenya	Kyrgyz Republic	Lao PDR
Lebanon	Malawi	Mauritania
Nepal	Niger	Pakistan
Papua New Guinea	Rwanda	Sri Lanka
Syrian Arab Republic	Tajikistan	Tanzania
Timor-Leste	Tonga	Vanuatu
Venezuela		

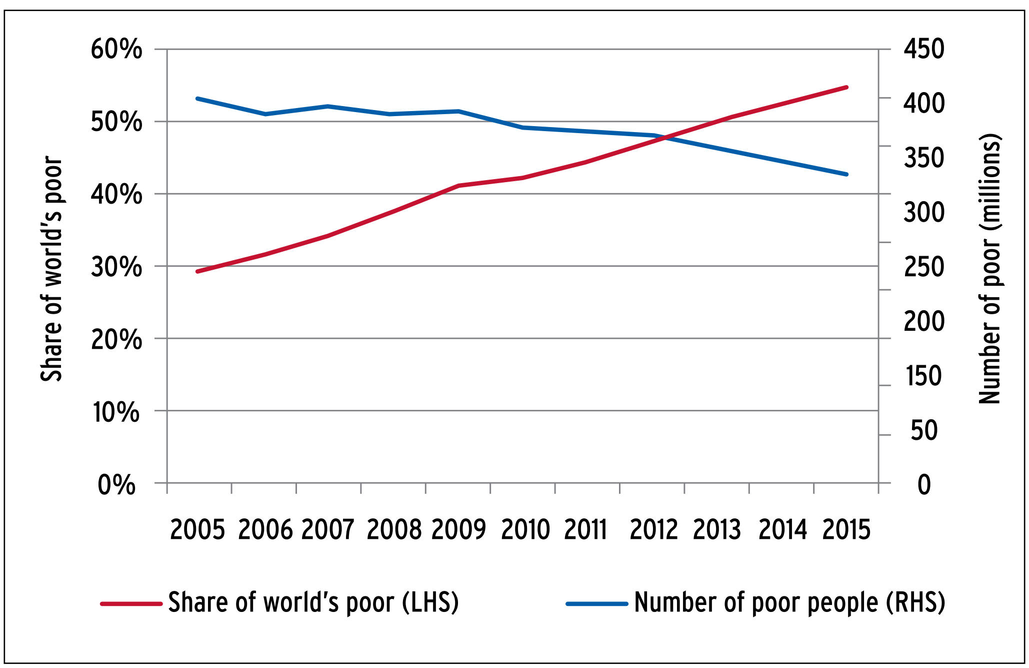
ANNEX 2: BACKGROUND ANALYSIS ON AID TO FRAGILE STATES

Table 2: Geographical distribution of fragile states[29]

	Number of fragile states	Population of fragile states (millions)	Regional share of global fragile state population	Share of regional population living in fragile states
East Asia and Pacific	6	82	6%	4%
Europe and Central Asia	3	39	3%	10%
Latin America and Caribbean	1	10	1%	2%
Middle East and North Africa	5	136	11%	41%
South Asia	5	411	32%	26%
Sub-Saharan Africa	27	608	47%	72%
World	47	1,286	100%	26%

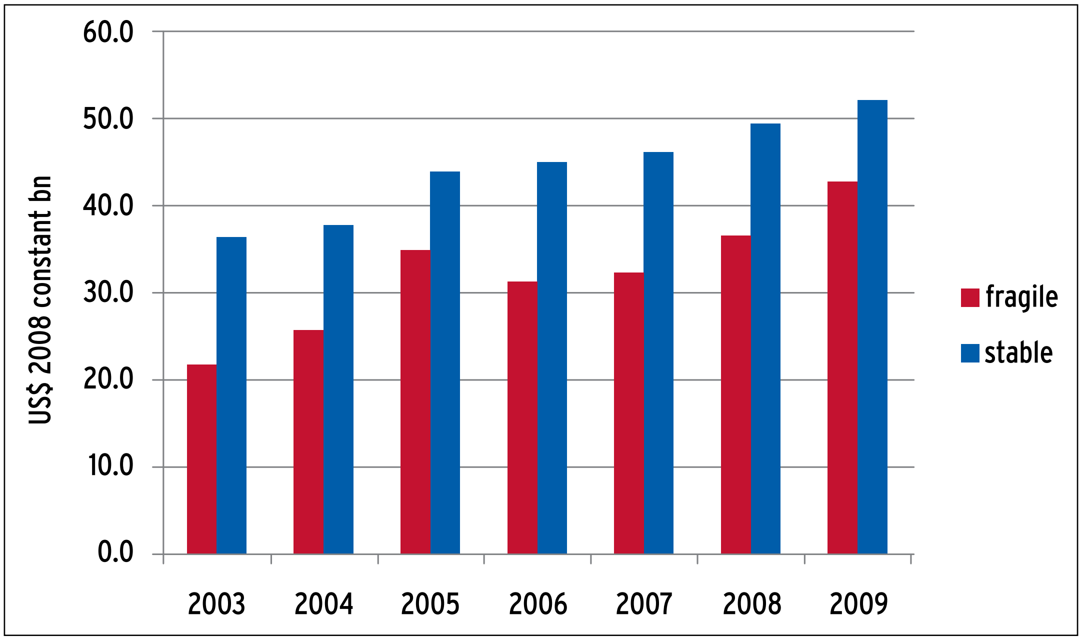
Source: Authors' calculations

Figure 5: Share of world's poor living in fragile states[30]



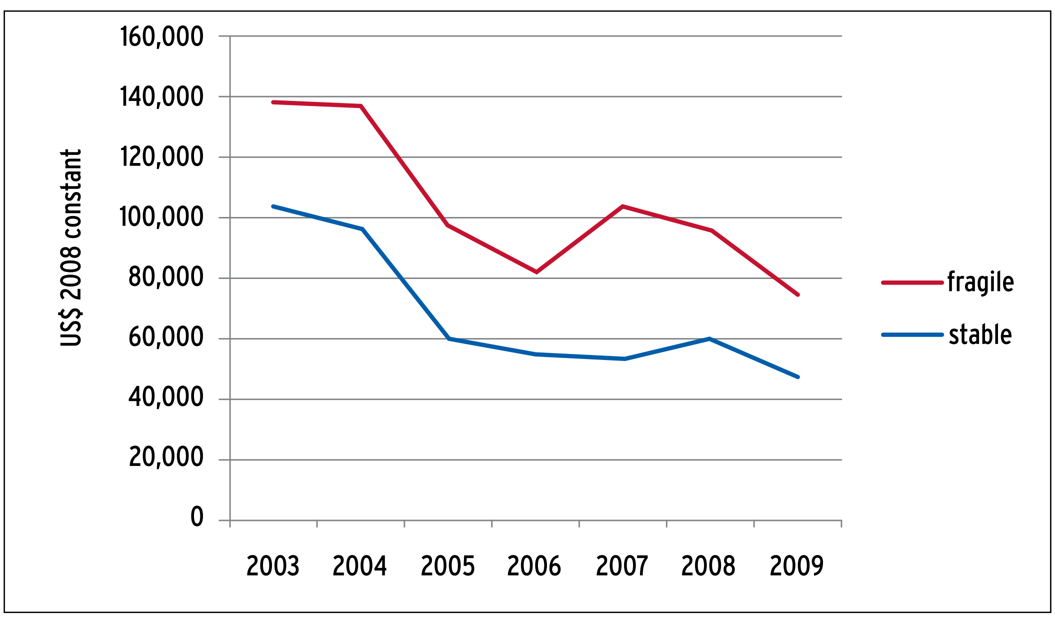
Source: Authors' calculations based on Chandy and Gertz, 2011[31]

Figure 6: ODA to fragile and stable countries[32]



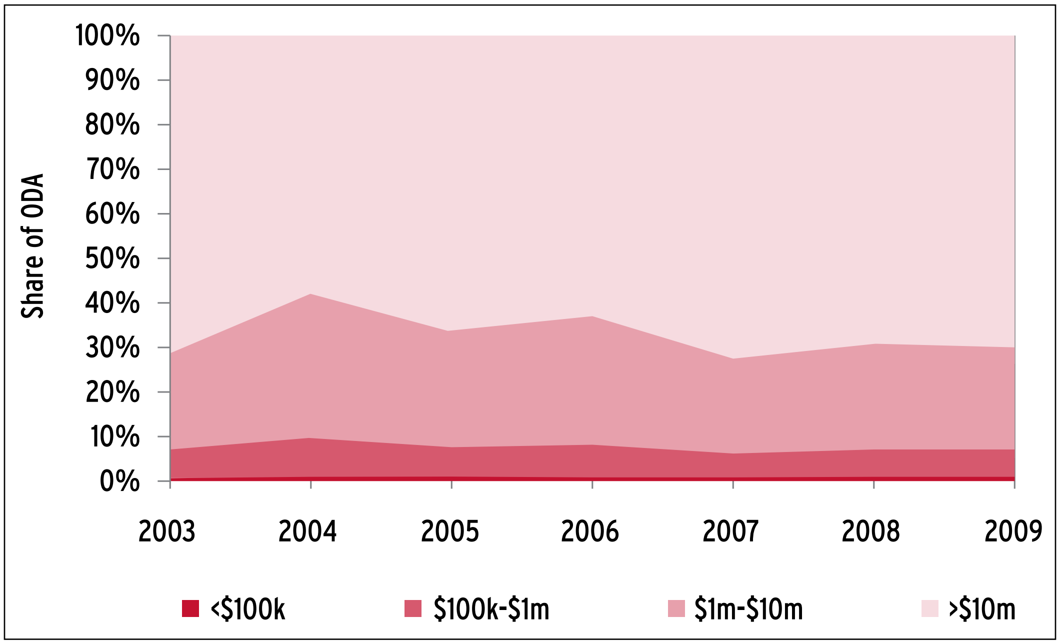
Source: Authors' calculations based on OECD DAC 2011c

Figure 7: Median aid activity size



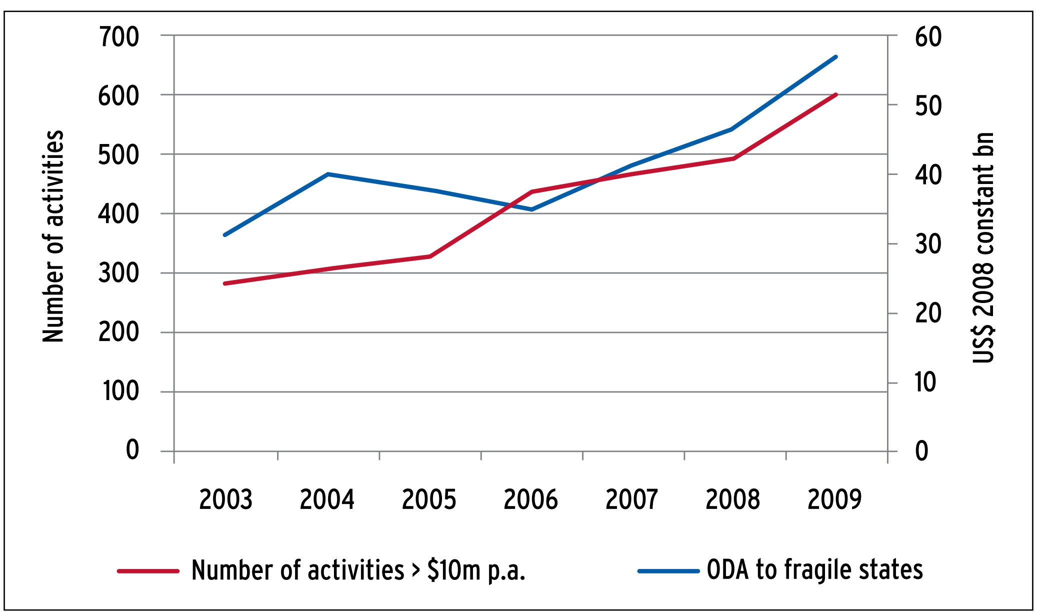
Source: Author's calculations based on OECD DAC 2011b

Figure 8: Distribution of activity size in fragile states



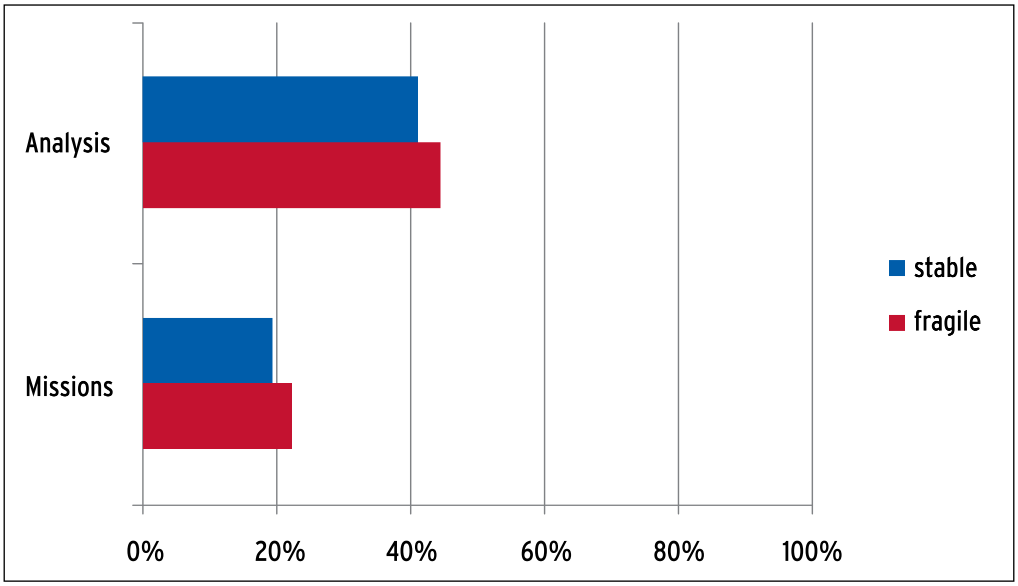
Source: Author's calculations based on OECD DAC 2011b

Figure 9: Number of very large activities versus ODA growth in fragile states



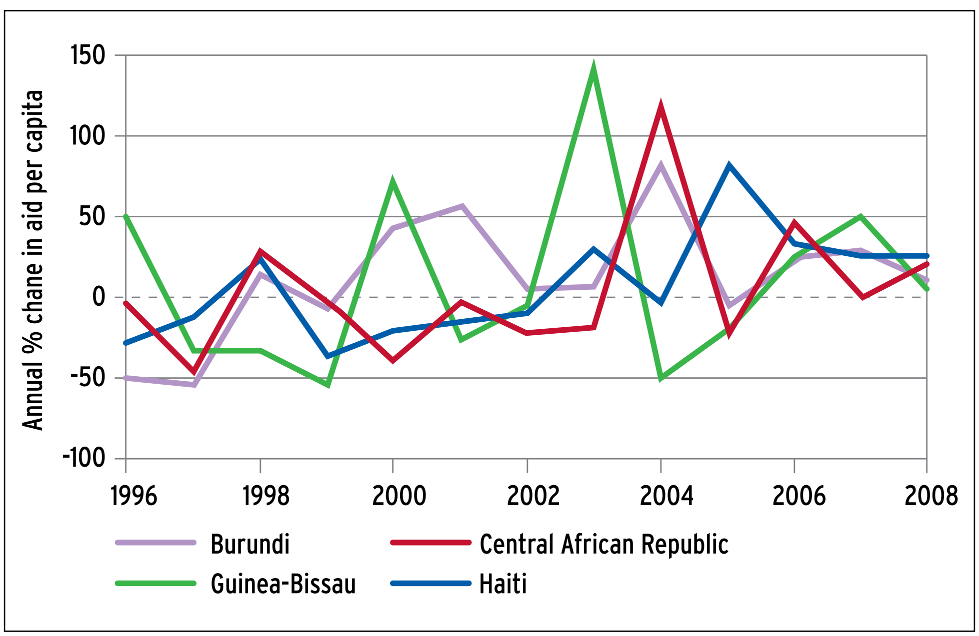
Source: Author's calculations based on OECD DAC 2011b

Figure 10: Share of donor activities that are coordinated



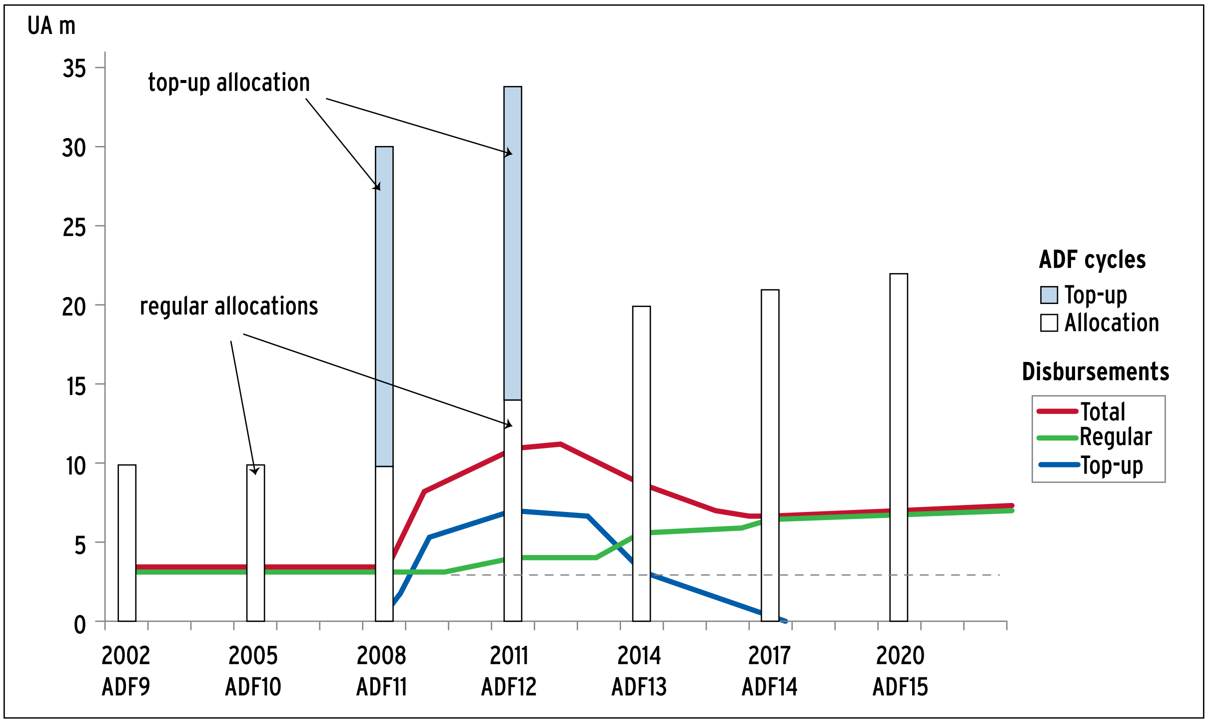
Source: Authors' calculations based on OECD DAC 2008a

Figure 11: Examples of volatile aid flows for four fragile states



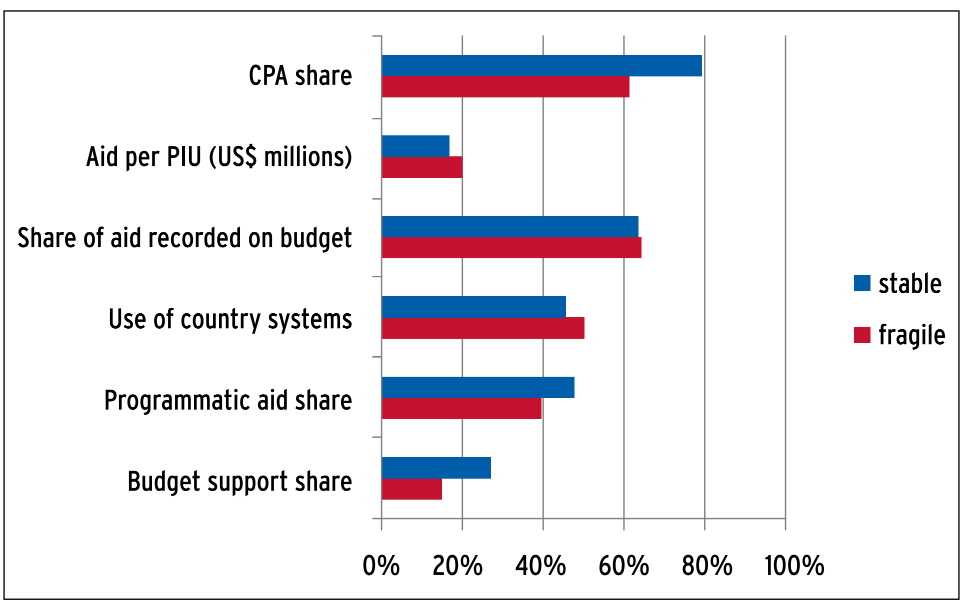
World Bank, 2011

Figure 12: Volatile financing: the case of the African Development Fund



Source: African Development Fund, 2008

Figure 13: Aid sustainability and transfer to government control



Authors' calculations based on OECD DAC 2008a and OECD DAC 2011c

ANNEX 3: SUMMARIES OF PROJECT CASE STUDIES

This Annex contains descriptions of the 17 project case studies drawn upon for this paper. The descriptions are themselves summaries of more detailed accounts, which are clearly referenced in the case study titles. Those interested in understanding more about the projects are encouraged to seek out the original sources.

Health

1. Afghanistan—Basic Packages of Health Services with World Bank support[33]

During the 1980s and 1990s in Afghanistan, much of the health services in rural areas (where 80 percent of the population lives) were provided by NGOs working cross-border, mainly from Pakistan. Funded from various external sources, the NGOs' services were often of good quality, but their coverage was modest, with only about one functioning primary health care facility per 50,000 people. The activities of the NGOs were generally uncoordinated and unfocused; successive governments showed little interest or ability in coordinating their activities and besides, the NGOs were keen to keep their distance from the government.

In 2001, the greatest change in the sector was the establishment of a comprehensive national health plan, enabling all actors to work toward common health objectives and to expand existing health services. The Afghan Transition Authority (ATA) started the process of making essential health services available to all Afghans through the introduction of a Basic Packages of Health Services (BPHS) in 2002. This process was supported by the World Bank, following a World Bank/Asian Development Bank/UNDP preliminary assessment in December 2001/January 2002.

Recognizing the lack of capacity within the Afghan Ministry of Public Health (MoPH), the ATA looked for an opportunity to harmonize NGO resources with government efforts. Under the BPHS, actual health service delivery is contracted to NGOs and private actors, based on a bidding process facilitated by the three main donors (USAID, European Commission and the World Bank) that results in signed, time-limited 'Performance-based Partnership Agreements' (PPAs). NGOs are selected competitively with credible sanctions in case of poor performance. Although the MoPH remains responsible for health service delivery in some provinces, its primary role is to develop strategies and objectives, set indicators, and monitor, supervise and control the performance of the implementing partners.

Around 30 NGOs are involved countrywide in delivering BPHS services, organized both horizontally (into regions) and vertically (into core programs such as malaria and tuberculosis). As of 2005, approximately 70 percent of districts were covered by the BPHS, providing primary health care to 50 percent of the Afghan population—a substantial and important achievement. The cost of delivering the basic health package is about $4 per capita per year, which is comparable with experience in other low-income countries. There was a 136 percent increase in the number of functioning primary health care facilities from 496 in 2002 to 1,169 in 2007. The health management information system indicates that there has been nearly a fourfold increase in the number of outpatient visits from 2004 to 2007. Independent assessments confirm that the quality of health care and health outcomes have also improved significantly over the period, despite a worsening security situation.

The BPHS has proven to be a well designed and largely effective tool for scaling up health services in the country. In an environment crippled by low capacity and a lack of formal institutions, scaling up moved much faster than would normally be expected. Scaling up was driven by the political and moral necessity to expand service coverage rapidly. The speed of scaling up has caused concerns, including over the longer-term financial sustainability of the program. Initial caution about the PPAs was summed up as "too fast, too soon, on too large a scale, with limited experience to draw from" (Daly, 2003).

A clear success of the BPHS design which allowed rapid scaling up was the strategy of co-opting all existing service delivery capacity in the country. NGOs have ultimately been transformed from independent developmental agents into service delivery contractors, while private actors play an increasing role without being subjected to regulation or control of the Afghan government. By giving NGOs a fair degree of autonomy but holding them accountable for achieving national priorities, the MoPH has addressed serious constraints such as scarce human resources, lack of physical facilities and logistical challenges.

Another important element in the scale up of BPHS was the role of M&E. Regular, independent and rigorous M&E of health sector performance has allowed MoPH to identify problems, act quickly to resolve them and track whether progress has actually been achieved. Such M&E has incurred large administrative costs for the ministry, but are seen as crucial to the program's success. By contrast, measurements of PPA performance criteria are believed to be unreliable due to the scarcity of baseline and projected population data.

2. *Afghanistan—The DOTS program with GFATM support*[34]

Afghanistan has a very high incidence of tuberculosis. According to one evaluation:

> "Despite the destruction of the National Tuberculosis Program (NTP) and basic health services by war and an uncertain security situation, the NTP, with assistance from many partners and REACH, increased the number of patients receiving DOTS by 136 percent in four years (from 9261 cases in 2001 to 21,851 in 2005), with an 86 percent treatment success rate. By focusing on rapidly expanding the number of facilities capable of providing tuberculosis (TB) diagnostic and treatment services and involving community health workers in case detection, referrals and home-based DOTS, REACH showed a 10-fold rise in the number of facilities providing TB services and a 380 percent increase in the number of sputum smear-positive pulmonary TB cases detected in two years (from 251/month in 2004 to 818/month in 2006) in 13 provinces." (Ahmadzai et al., 2008, p. 180)

The program started with a clear vision (a TB-free Afghanistan by 2050) and a strategy (use of the WHO DOTS approach) and included well-defined intermediate progress benchmarks for key inputs, outputs and results (the two-year operational plan developed in 2005) with monthly monitoring of progress. The program operated at the community level, specifically involved women, had the support and engagement of the Ministry of Public Health and was carried out under the direction of the National TB Institute. It also engaged a multi-donor group consisting of multilateral and bilateral agencies, as well as civil society organizations, which worked very closely with the government. The program dramatically expanded the coverage of DOTS treatment from 2001 to 2005, both geographically and in terms of the percent of health facilities applying the treatment (Table 3).

Table 3: Scaling up DOTS treatment in Afghanistan

Year	Health facilities	Health facilities applying DOTS n (%)
2001	1013	36 (3.5)
2002	1013	79 (7.7)
2003	1013	131 (12.9)
2004	1013	202 (20.0)
2005	1115	466 (41.8)

Source: Ahmadzai et al., 2008

Key constraints that the program has faced include the widespread destruction of health facilities, limited human resource capacity, climatic and geographic obstacles and the bad security situation in some parts of the country. The evaluation concludes that the key lessons of the Afghanistan DOTS program are: (a) the focus on improving key delivery points works well; (b) training of health facility staff is a key ingredient; and (c) community health workers should be involved in the detection and treatment of TB at the community level.

The source does not permit a complete analysis of the drivers and spaces for scaling up, but it appears that the program effectively created the political, policy, institutional, ownership, partnership and learning space required to scale up successfully. For longer-term sustainability at scale, the question of fiscal space (sustainability) will have to be addressed and is likely to require either sustained donor engagement or steady access to sufficient budget resources. Problems in the security space were beyond the scope of the program and served as a serious constraint in some areas of the country.

3. Timor-Leste—health sector rehabilitation with World Bank lead support [35]

In 2000, a multi-donor health mission, led by the World Bank and the Interim Health Authority of Timor-Leste, developed a sector-wide approach to address the country's serious health and health system problems and to develop appropriate health policies and institutions. It was to combine the provision of basic health services with building national health system capacity.

The World Bank prepared two projects in sequence. The first retained international NGOs as principal service providers (while placing government in control of coordination and policy), whereas the second shifting implementation responsibility to the government. In addition, the second project went beyond the first by stressing greater health service utilization and higher quality care.

The program succeeded in establishing a functioning health system over a period of about five years, relying entirely on Timorese administrators by the program's end. Weaknesses remained in the system as of the mid-2000s (including low utilization rates, stagnating performance indicators and a weak hospital system), but overall the large scale program was deemed a success. Key success factors were:

- a good knowledge base at the outset (as a result of a thorough initial assessment);
- a strongly articulated horizontal and vertical scale, including a strong policy reform agenda;
- a large and sustained financial engagement by the donors;
- flexibility and speed;
- a strong sector-wide framework for donor coordination; and

- relative political stability and in that context, a strong partnership with a politically strong and cohesive government and with the local expert community, resulting in strong ownership.

4. *Kenya–Adolescent Reproductive Health Program with USAID support*[36]

In 1999, the Population Council's Frontiers in Reproductive Health Program (FRONTIERS) and the Program for Appropriate Technology in Health (PATH) collaborated with three government of Kenya ministries to design and implement a multi-sectoral project to introduce HIV/AIDS education in schools. This was the same year when HIV/AIDS was declared a national disaster by the government, opening the doors to public campaigns and discussion about the pandemic.

The project had the following three goals: to improve knowledge about reproductive health and encourage a responsible and healthy attitude toward sexuality among adolescents; to delay the onset of sexual activity among younger adolescents; and to decrease risky behaviors among sexually active adolescents._

Three interventions were piloted in two districts in the Western Province of Kenya over a period of 30 months. Key findings from the pilot project demonstrated that the three ministries could successfully implement the interventions with minimal support from FRONTIERS and PATH.

The positive results of the pilot phase prompted a 20-month phase of adaptation and expansion of KARHP throughout the two pilot districts to enable the ministries to gain experience of implementing the services at the district level. Pilot materials and tools were revised and inter-sectoral committees set up at district and provincial level.

The approach was then further scaled up throughout the remaining six districts of Western Province from June 2005 to May 2006. This province-wide scaling-up experience led to a further 13-month phase of replication, during which the model was introduced in four districts within Eastern and Nyanza provinces in June 2006 to May 2007. This was then followed by province-wide expansion by USAID-funded partners. From June 2007 to May 2008, KARHP was introduced in Nairobi and Central province.

Despite the challenges of working with the public sector in a country known for its bad governance, this program proved that multi-sectoral approaches that build the capacity of government ministries to mainstream services can lead to wide-scale expansion and the sustainability of effective pilot models.

In sum, a few key observations can be made:

- This case followed a "classical" scaling up pathway— starting with an innovative idea, implementing experimentally through pilots, evaluating the pilots before replicating the intervention geographically in multiple sequential steps, and sustaining these over a long period to reach national scale.
- The approach involved both horizontal scaling up (geographic replication) and vertical scaling up (advocacy and policy engagement at the national level). No functional scaling up (e.g., expanding the coverage of health areas or target age groups) appears to have been attempted.
- The interventions were implemented through existing governmental structures, networks and systems. Working with and through the ministries contributed to the sustainability of the interventions and the integration of activities within routine operational plans, protocols and policies. This approach also ensured ownership and prioritization of adolescent sexual

health issues—including budgetary allocations—by the government. Key to the ministries' adoption of KARHP was their initial involvement in the planning and design of the model.

- Capacity building was effectively built into the project using a "cascade approach" following existing government supervisory structures, in which each level of ministry staff would train and supervise those immediately below them.

- Partnerships helped to identify opportunities for leveraging of resources and funding. This worked well in the expansion and replication phases, when the government ministries funded much of the expansion, while USAID funding supported the technical assistance provided through FRONTIERS. The involvement of three ministries also represented an effective partnership, best evidenced through the supportive role played by inter-ministry committees in overseeing the project's roll-out.

- The use of evaluation evidence was critical to both convince the ministries and USAID of the effectiveness of the model and to improve and adapt the original interventions within existing government structures.

- Institutional, ownership, learning, capacity and partnership spaces were effectively created to allow scaling up to take place. Special care was taken to minimize project cost in recognition of the difficulty of securing fiscal resources (or space) in a setting where the budget process is highly politicized and unpredictable.

- Multi-sectoral approaches for addressing adolescent sexual health and HIV/AIDS were proven to be effective in improving knowledge, attitudes and behaviors.

- The challenges of poor governance were evident in the project in terms of high ministry staff turnover and complex budgeting and planning systems within the public sector. Nevertheless, robust project design and strong ownership meant that these constraints did not undermine the project.

5. *Tajikistan—HIV/AIDS, tuberculosis and malaria program under GFATM, implemented by UNDP*[37]

The disintegration of the Soviet Union and of the Soviet health system brought about a serious weakening of the Tajik health system and a low capacity to respond to epidemics, including HIV/AIDS, TB and malaria. UNDP serves as the management authority for GFATM's support program to the Tajik health sector. The program started in May 2003 and has grown rapidly from $1.3 million expenditure in 2003 to $10 million in 2010. Substantial progress has been made in the three key areas of HIV/AIDS, TB and malaria. For example, over four years, malaria cases dropped from an estimated 6,000 to 165. Box 11 summarizes progress with scaling up the model of voluntary HIV/AIDS counseling and testing. A focus on scaling up is built into the approach.

UNDP focused on capacity development from the outset, initially concentrating on implementing partners and more recently, in line with GFATM policy, shifting attention to broader capacity improvement across the health system in the three priority areas. According to UNDP staff, while UNDP is supporting local capacity building, the Direct Implementation Model (DIM), as opposed to implementation by the government itself, continues to be needed in Tajikistan because of weak national administrative capacity.

The GFATM approach involves the setting of long-term and intermediate result targets that define the scale and key dimensions of the scaling up pathway and provide benchmarks for intensive monitoring. While the GFATM funding model is known for being quick and unbureaucratic, according to UNDP staff and government clients, GFATM has recently become more like a traditional donor, with a large number of headquarter staff and increasingly burdensome reporting requirements.

Key aspects and issues in terms of drivers and spaces for scaling up are summarized in Table 4.

The GFATM program's scaling up approach in Tajikistan is characterized by the following aspects and challenges:

- Pathways for scaling up (including defining the ultimate scale, the sequence of interventions and intermediate targets) in each of the three illness areas are well-defined, with monitoring along the way.
- Financial/fiscal sustainability depends on continued GFATM support, and/or greater government's fiscal capacity and priority in the longer term. The long-term financial trajectory for sustainability of programs at scale has yet to be defined.
- As long as UNDP remains as the principal recipient and administrator for the country, the implementation mode will be DIM. Greater attention is required on enabling the capacities of the government and civil society to take over the implementation of grants or certain grant components in the future. Early thought will have to be given as to how to manage this transition and ensure that natural disincentives for hand-off do not prolong the DIM approach unnecessarily.
- Issues of broader health sector strategy and GFATM engagement are still to be addressed. This is part of a larger debate regarding GFATM programs and to what extent they should be

Box 11: Voluntary HIV Counseling and Testing

A good example of scaling up in HIV-related interventions is the introduction of routine provider-initiated testing, as well as voluntary counseling and testing (VCT) in all public health facilities and service delivery points. The table below shows the growth in the number of people tested, the number of labs and counseling facilities, and the number of new HIV cases registered. In 2007, the Ministry of Health piloted the integration of provider-initiated testing and counseling (PITC) into reproductive health centers and maternity homes in eight districts, as part of the GFATM grant. Currently, 30 percent of territorial districts in Tajikistan offer PITC to pregnant women. Experience has shown that agreement to undergo HIV testing when routinely offered is particularly high among pregnant women. HIV testing among pregnant women has increased from 19,801 women tested in 2007 to 40,171 in 2008 and 76,297 in 2009.

Routine testing does not replace the need for VCT and in non-clinical settings there is a need to strengthen the capacity of health workers for VCT provision among most at risk population groups. A planned comprehensive communication campaign is expected to further improve the uptake of VCT and reduce HIV-related stigma.

	2005	2006	2007	2008	2009	2010
People reached with VCT	58,899	93,791	93,264	148,255	210,179	>280,000
Number of labs	9	9	13	13	21	21
Points providing VCT services	82	97	108	122	231	231
New HIV cases registered	189	204	339	383	431	1004

Source: Drafted by UNDP Tajikistan Country Office Staff; quoted from Linn, 2011b

Table 4: Drivers and spaces for HIV/AIDS, tuberculosis and malaria program

Drivers		Spaces	
External	Health crisis	Fiscal/ financial resources	GFATM funding critical; without it the program is un-sustainable
	Global Fund	Policy	policy dialogue for ATM subsector reform
	International/ local NGOs	Institutional/ Organizational	Currently depending on UNDP; effort to build up government capacity a long-term goal
Internal UNDP	GFATM funding opportunities	Political	Advocacy for ATM
Internal government		Partnership	Implementing partners; government
Incentives	Global fund financing strong driver	Natural resources	N.A.
		Cultural	Adaptation to local customs/sensitivities required

fully integrated and supportive of comprehensive health program design and implementation.

Education

6. *Cambodia—education project supported by UNICEF*[38]

Shortly following the comprehensive peace settlement in Cambodia, UNICEF worked closely with the country's Ministry of Education to design and implement a basic education reform program. The program had three main strategies: a focus on quality of instruction and access equity for all children, supported by empowerment of local communities; analytical work provided through long-term technical assistance to support educational reform from national through local levels; and the design of new policies to institutionalize reform of technical and operational systems.

Program implementation was arranged into four components:

- The designation of 36 demonstration cluster schools, each connected administratively to a group of five to 10 other schools, spread across the four geographical areas of the country. The cluster arrangement meant that the program was able to ultimately reach 30 percent of primary school students in the country. Clusters were used to demonstrate educational changes in policy, management, curriculum, textbooks and teacher training, and to build support among local educators for reforms. Clusters also built upon the work of cooperating NGOs, which could be used to absorb greater donor funding.

- Toward the end of the project, pilots were set up to improve in-service teacher training in primary schools. It was originally intended that these pilots would be extended nationwide with a $30 million grant from USAID. However, the coup derailed the program and it was later recast on a much smaller scale and with less funding.

- Curriculum reform carefully developed with the support of parents, teachers and university academics, local and national officials. This led to the development of new textbooks and new rules for testing and evaluation.

- The establishment of a national Education Management Information System (EMIS) for sector monitoring and decision making. This included the introduction of an annual school census in all grades up to grade 12, whose information could then be used to support planning in provinces, districts and cluster schools.

From a scaling up perspective, there are a number of relevant lessons from this project.

First, clusters proved to be a highly effective vehicle for scaling up. In a country with precious few human and material resources, the cluster model provided a way to share expertise among several schools and a way to disseminate ideas and training through an organized school network. It also provided an authorized means for villagers, teachers and students to meet each other and was an approach that was supported by the sector's authorities. The cluster system may have also served to rebuild the bonds within the country's network of educators, which had been undermined during the conflict and whose strengthening likely contributed to the project's sustainability. It is interesting to note that the project contained other scaling up elements that were not effective, the in-service teacher training pilots being a case in point.

Second, the involvement of communities in the project has been identified as a vital element in the project's success. For a country devastated as Cambodia was by a prolonged civil war, education reform required re-establishing the foundations of the education system—a process which had to be consultative and inclusive to be effective.

More broadly, effective communication was required to ensure broad-based support for the project among civil society and beyond. Consultations with civil authorities focused on budget issues, security and support for change. Provincial governors were routinely briefed to solicit protection of construction sites and to share information regarding security. Governors and provincial, district and village leaders were included in community ceremonies to turn over money for construction or repair of schools. For illiterate community members, this provided visible, verbal accountability for money given for education where the account books and purpose for funding were read.

Third, care was taken to ensure that curriculum development did not become politically charged. An initial focus on competencies leading to functional skills provided a platform for discourse about content and teaching methods that was free of political tension. Trust was nurtured among stakeholders before more difficult problems associated with history or geographical content were tackled later on.

Fourth, Cambodia's weak governance meant that education reforms had to be seen as credible and built around meaningful institutional development. Procurement and accountability for money in the sector at all levels were rightly perceived as the foundation of educational integrity, just as much as valid testing and promotion. Projections, receipting, cost standardization, timely payment of salaries, transparent monitoring, involvement of NGOs and parents, and public reporting were all necessary for reconstructing the education sector after the conflict.

Fifth, the project made use of all available technical expertise in an environment that was recognizably weak in capacity. Connections to regional experts (outside of Cambodia) during the initial phase of work and to staff long-term, on-site TA were critically important. Experts from the Philippines and Thailand had done similar work in their own countries, and brought a technical background and cultural sensitivity to how education reform could support the transition from conflict to stability. Their insights provided not only technical expertise but practical strategies for overcoming fear, as well as opening connections for Cambodian educators to technicians in their home ministries. The UNICEF model also drew on the experience of the strong network of educators working in Cambodia's refugee camps, including the work of NGOs, such as Save the Children.

7. *Ethiopia—Alternative Basic Education for Children out of School project, implemented by Save the Children, Norway*[39]

In Ethiopia, Save the Children Norway (SC Norway) developed and field tested an alternative model for the first cycle of primary school—Alternative Basic Education for Children out of School (ABECS)—that is cheaper, more flexible and more compatible with local conditions than the formal schools. At the same time, the model maintains high quality and ensures that children have the opportunity to continue their education in the regular education system.

ABECS was developed in close collaboration with the education authorities in Amhara Regional State. It began with a pilot phase, in which SC Norway helped establish 37 ABECS centers in three weredas in North Gonder in Amhara Regional State. The results in terms of increased enrollment and retention rates, as well as gender equity, were excellent.

The ABECS model was then accepted as the preferred model for non-formal education by the other NGOs working in Amhara Regional State. By early 2003, the nine NGOs working within primary education in Amhara Region had set up 151 centers in 17 weredas. The regional authorities later encouraged the expansion of ABECS to all 114 weredas of the region. Over 1,000 ABECS centers were established throughout the region in the second half of 2003, with an estimated enrollment of over 120,000 children. SC Norway continues to support the ABECS program, concentrating its efforts on maintaining the quality of the ABECS education throughout the scaling-up exercise. Thus far, the outcome of the ABECS program seems promising and its potential impact seems to exceed the comparatively modest funding that ABECS has received.

In a country like Ethiopia, with low school enrollment levels as a point of departure (22 percent in 1991–92), the goal of Education for All by 2015 was exceptionally ambitious and demanded a massive effort. The government accepted the main responsibility for education through the formulation of its 20-year Education Sector Development Plan, but also provided opportunities for other actors to provide support. By the end of the first phase of the ESDP (1997–2002), enrollment had increased dramatically but there were still large parts of the country where primary education was not available. In 2003 over 38 percent of all school age children remained out of school. By the time the ESDP moved into its second phase (2002–2005), it was evident that the Education for All goals could not be reached through formal education alone.

ABECS is one of several alternative education models developed by NGOs in Ethiopia. Key elements of the ABECS model are: the establishment of a local school committee to run the school, usually appointed by the "kebele" (the smallest administrative unit in Ethiopia); the building of the school with local materials by the local community; the hiring of instructors from local out-of-work school leavers; teacher training through formal courses during school breaks and through in-service training; a school schedule that allows the school committee to adjust the schedule to local needs (for instance, the local agricultural calendar, when children are required to help with the harvest); and a carefully designed curriculum, developed by SC Norway.

There are four clear strengths to SC Norway's approach:

Local knowledge: A dominant feature of social life in Ethiopia is the hierarchically organized, highly centralized and authoritarian nature of the Ethiopian state. The NGO sector in Ethiopia is young and limited in numbers. Relations with the government have been uneasy. NGOs have been treated with mistrust by the government and their work has been circumscribed by a detailed and cumbersome regulatory framework. This framework seems to effectively prevent Ethiopian NGOs from becoming membership organizations, which could evolve into civics movements. Local NGOs primarily see their role as encompassing service delivery.

SC Norway has been working in Ethiopia since 1969. In recognition of the difficult operating environment presented by Ethiopia, SC Norway scaled down its own capacity to implement projects from the mid-1990s onward, and relied more on partnerships with government agencies as well as a number of local NGOs for operational capacity.

Integration with existing institutions: The ABECS model was designed to integrate effectively with the formal school sector, both for students and teachers. Children who complete the ABECS program are accepted into the second cycle of the formal schools and have the option of continuing their education. After a two-year training period, ABECS instructors may qualify as formally certified primary school teachers.

Promoting community ownership and responding to community needs: The ABECS model promotes the role of the local community. It gives the local school committee a large measure of disciplinary control over school instructors. In addition, it has responded directly to community concerns, by for instance, reducing the distance girls have to travel to attend school and adjusting the school schedule to the community calendar.

Government ownership: Ultimately, the government led the scaling up process of the ABECS model. The acceptance of the ABECS model depended on an amenable policy environment within the Ministry of Education. SC Norway was fortunate to find well-placed allies in the system who had an interest in testing and promoting the models proposed by SC Norway. SC Norway has developed good relations with the education authorities at both regional and wereda levels, partly by providing some support to the formal schools (in the form of libraries or laboratories) and to the administration itself.

Government buy-in was enhanced by the ABECS design which reinforced the role of government structures including kebeles and weredas. ABECS schools are typically built on land allocated by a kebede (which often includes extra land that can be used to generate income for the school, e.g., by renting out the land on share-cropping contracts). Weredas, who are now the focal point for local development efforts, share in the supervision of ABECS instructors.

Despite the project's undoubted successes, there are concerns regarding the sustainability of the sector's rapid expansion under the ABECS model. It seems unlikely that communities will be able to fully cover costs when NGOs withdraw, despite attempts by the communities to contribute to various cost-sharing schemes. Teachers' salaries are already a major issue at wereda level, consuming 90 percent or more of wereda education budgets and even if ABECS instructors are salaried at lower rates than formal teachers, the number of additional instructors needed will create serious problems. Furthermore, one of the driving forces of the success of the pilot phase was the clear expectation by ABECS instructors that they would be hired as formally qualified primary school teachers once they had successfully completed their training. The first compromise reached on this issue is that qualified ABECS instructors will be favorably treated when new positions in the formal schools become available, but that they will be paid at the former rates for as long as they remain ABECS instructors.

8. *Kenya—deworming in Kenya with World Bank and DFID support*[40]

Intestinal helminths, or worms, infect roughly a fourth of the world's population and contribute to anemia, malnutrition and other illnesses, disproportionately affecting school-aged children. Fortunately, simple and cheap drugs have been shown to be highly effective against many types of worm infections. Since 1987, the World Health Organization has endorsed mass school-based deworming programs in areas with high helminth infections.

Despite this endorsement, it remained unclear whether poor countries should accord priority to deworming programs. In one of the first randomized impact evaluations in international development, Ted Miguel and Michael Kremer tested the effect of school deworming using evidence from a small-scale deworming project in southern Busia, a poor and densely-settled farming region in western Kenya. The project covered 75 project schools with a total enrollment of over 30,000 pupils between ages six to eighteen.

The evaluation showed that deworming had even greater public-health benefits than previously thought due to the positive externalities on untreated children, and that the intervention presented an extremely low-cost way to boost school enrollment in Western Kenya. This evidence, combined with an international campaign to disseminate results and to provide technical assistance in adopting and implementing the project, led to a massive scale-up of the intervention.

Within Kenya, the education and health ministries formed a joint team to develop a plan for implementing deworming in all the regions of the country where worms were a major health problem. This was incorporated into the national education plan, which was approved by the government and funded by donors and the Kenyan government.

In 2009 the Kenyan government successfully implemented a nationwide school-based deworming project reaching over 3.5 million high-risk children at a total cost of just Sh. 70 million ($836,000). Donors, including DFID and the World Bank, supported the project with technical and financial assistance.

In late 2009, an audit reported that Sh. 5.5 billion ($63 million) of funds for primary education had gone missing. When donors demanded that the Kenyan government account for the missing funds, the education minister was sacked, only to be reinstated later. The permanent secretary was suspended, later reinstated and then transferred to another ministry. Unsatisfied with the government's response, donor support for the broader multi-donor program of which the deworming project was just one part, was suspended. As a result, the deworming project was not fully implemented in 2010. It is unclear whether the project will be fully reinstated.

The case study provides a valuable lesson in the difficulties of translating evidence into policy in fragile states. Even when randomized trials—the gold standard in evidence-based policy formation—identify cost-effective interventions in small-scale pilots, other weak links from ministry financing to project implementation may hinder successful scale-up of proven interventions. Compelling evidence, combined with strong recipient demand and ownership, may not be enough to assure project success in poorly governed countries. Furthermore, in a country like Kenya characterized by bad governance but relatively high capacity, the scaling up process itself, from pilot through to large scale replication, may be less of a challenge than the sustaining of an intervention. Whereas the scaling up phase may only last a few months, during which time the project may only encounter a small number of obstacles, sustaining an intervention covers an indefinite period when countless challenges may emerge.

The case study provides a strong case that efforts to improve governance and institutions should be incorporated into the scaling up of projects and that strong political economy analysis can support the successful completion of scaling up pathways. In addition, the case study highlights the collateral damage to scaling up that can be caused by weak donor-government relationships.

Local and Community Development and Social Funds

9. Cambodia—The Seila (local development) program with UNDP support[41]

The Seila program started in 1996 in a context of post-conflict and weak governance. The program sought to pool donor and government resources for investments in local infrastructure with community and local participation. Subsequently, the program incorporated an explicit focus on supporting sub-national participatory planning and budgeting. The intended outputs of the program were improved local services, stronger local institutions, and enhanced national policies for, and regulation of, decentralized governance structures.

Building explicitly on the experience with and lessons of early program initiatives, the Seila program initially focused on four provinces, but was systematically extended by the mid-2000s to cover all provinces in Cambodia. The program involved a combination of funding sources organized through a Provincial Investment Fund, which provided incentives for participation by the local authorities and communities. It also engaged actively in capacity building. The program was supported by UNDP in a long-term engagement (exceeding 15 years) and by a group of donors who provided pooled financial support. The implementation approach created space for learning, by "continuous monitoring and evaluation, both internal and external, and through operational feedback … several far-reaching evaluations of Seila have declared the program an unusual success in the Cambodian context." (Hughes, 2007, p. 95-6)

Key elements of the program's success were:

- the program's approach, which allowed for flexibility, experimentation and feedback;
- the deliberate attempt to scaling up from small to large, with long-term engagement by the key partners; and
- successful capacity building efforts and en-

gagement with government, supported by strong incentives for the key stakeholders at the local and provincial level.

Important challenges for the sustainability of the program include:

- concerns about the long-term sustainability of local services funded under the program;
- uncertainty about long-term donor funding sources and eventual integration into the national budget;
- how to manage the salary supplements granted to program staff in the long term;
- how to maintain the local participatory process which is seen as burdensome by some participants; and
- maintaining the transparency and good-governance aspects of the program.

10. Mozambique—Decentralized Planning and District Finance with support from UNCDF and UNDP[42]

Similar to the Seila program in Cambodia, this project in Mozambique provided support for the development of planning capacity at the district and sub-district level, combined with a funding mechanism for local, small-scale, participatory public service investments. Starting in 1998, the UNCDF and UNDP developed a pilot in one province of Mozambique of supporting planning and local investment activities in 14 of 18 districts. After a midterm evaluation in 2000, the program was extended until 2006 with bilateral (Dutch and Norwegian) financial support. As part of the extension, the program was scaled up to reach all 18 districts in the original province as well as to seven additional provinces. There were indications that the government expected to take the program nationwide.

Key aspects of program design were:

- careful prior analysis of the sub-national governance situation as well as of the experience with past donor engagement;
- an effective learning process and flexible adaptation from the pilot phase through the scaling up process;
- the establishment of a district development fund, which provided financing and incentives at the local level;
- the creation of government ownership by working with and through government institutions, rather than setting up a parallel institutional framework; and
- vertical scaling up by reforming national policies and legislation in support of district planning and financing.

As in the case of the Seila program, the program in Mozambique faced significant sustainability and impact challenges, including:

- reliance on external financing as the substantial costs of implementation were covered by UN agencies and the district development fund depended heavily on donor support;
- sustained engagement of qualified personnel in a context of overall human resource shortages and weak institutional capacity;
- sustainability of community participation, especially if material incentives were to be reduced over time; and
- questions over the poverty impact of the infrastructure financed under the program, which seemed to primarily benefit better off households in the districts.

11. Yemen—The Social Fund for community development, supported by DFID[43]

The Social Fund for Development (SFD) is an independently funded quasi-governmental institution, which was established in 1997. It works with local communities across Yemen to finance various social development projects, assist local institutions to develop their capacity and efficiency, and create new employment opportunities co-financed with the private sector.

SFD was originally supported by the World Bank to help mitigate the impact of structural adjustment. Over many years, it has grown in size and the scope of its work has broadened both sectorally and geographically. SFD's education support alone has delivered almost 6,000 new classrooms since 2004. Some of SFD's initiatives, such as the capacity building program, have been introduced countrywide in all 333 districts of Yemen. Today, SFD is now considered one of the most visible and effective organizations for poverty reduction and community development in Yemen.

SFD receives support from a number of international donors, including DFID. SFD has been supported by DFID since 2003 and has evolved into a flagship program in DFID's country program. Spending on the SFD made up the largest percentage of expenditure in DFID's country portfolio (some 60 percent of total DFID spend in 2010) until the 2011 moratorium.

While operating partially outside and parallel to the structure of government, it is directed by the deputy prime minister (who is also the minister of planning and international cooperation). Furthermore, SFD supports the Yemeni government to implement programs at community level, including through co-financing, and provides significant support to build the capacity of government institutions. It is also a major proponent of the government's Yemen National Development Plan.

The following observations can be made about SFD:

- SFD has proven to be an enormously successful vehicle for scaling up, across several dimensions. SFD's scaling up successes have taken place in an otherwise very difficult environment for donors to work. Given its success, SFD has developed a reputation as being a highly convenient instrument for absorbing and spending aid money relatively quickly.
- Despite being the biggest project within DFID's portfolio and the broadest in terms of its scope and objectives, SFD also has the lowest risk rating of any project and provides a way for DFID to positively influence some of Yemen's weakest sectors.
- Strong political leadership has undoubtedly been an important driver in SFD's expansion and its success in scaling up particular interventions across the country. At the same time, SFD's governance structure is very reliant on the influence and support of political appointees, including its director, the deputy prime minister. There is a danger that the SFD could get caught up in Yemen's political crossfire in the future. However, it is hard to imagine any governance structure which would make SFD immune from political developments in the country, given the country's vast governance challenges and SFD's willingness to work closely with the government's development plans.
- Although SFD's degree of alignment with the work of government ministries has evolved (and strengthened) over time, SFD's existence outsight the government structure has raised concerns over the sustainability of its work. SFD has a close and effective partnership with the Ministry of Education, but less so with other

ministries. Alignment has been particularly problematic between SFD and local government. Another sustainability concern arises from SFD's use of different salary scales to those used in Yemen's public sector. Advocates of SFD's approach argue that it presents the least damaging approach to operating parallel systems.

12. *Tajikistan—District planning as part of a broader community development program, supported by UNDP*[44]

In Tajikistan, as in many other fragile states, districts have weak planning capacities. UNDP has developed a district-level planning methodology together with the Ministry of Economic Development and Trade, which is currently being implemented in 15 districts. Eventually the ministry plans to take over sole responsibility for managing this initiative and to roll it out to all districts country-wide. UNDP expects that, in the future, it will support the strengthening of the planning capacity of sub-district local authorities.

The district planning effort is a component of a broader Community Development Program (CP) in Tajikistan, which provides financing for small-scale local public service investments, micro-credit financing and employment creation schemes. UNDP developed and has been supporting this program for over 10 years and is planning to support it for at least another five years. The CP represents a strong platform for capacity building and strengthening community-based and local development, and for attracting funding from other donors for these purposes. The CP relies on the well-established capacity of five UNDP decentralized area offices in various parts of the country, supported by technical and project leadership capacity in the UNDP country office in the capital. Under CP, UNDP also engaged in vertical scaling up by providing assistance for the development of national legislation on decentralization and sub-national government structure and functions. The district planning component represents a functional scaling up of the original CP design and relies on a systematic scaling up approach.

As in the case of Cambodia and Mozambique, the sustainability of CP and the district planning component are subject to concerns regarding sustainability, especially if and when donors were to withdraw their financial and technical support. The fact that the government has been fully engaged in the development of the district planning method and intends to mainstream it though its own administrative channels, is a strong foundation for eventual sustainability, provided that further support from UNDP and other donors is specifically geared to assure long-term sustainability.

Agriculture and rural development

13. *Cambodia—Agriculture and rural development projects supported by the Asian Development Bank*[45]

In 2009, the Asian Development Bank's Independent Evaluation Department (IED) evaluated the entire program of ADB's 12 agriculture and rural development (ARD) projects and 31 technical assistance projects in Cambodia approved since 1995. The projects supported four types of interventions: (i) ARD policy and sector management projects, (ii) rural infrastructure projects, (iii) irrigation projects, and (iv) targeted rural development projects. During 1996-2000, parts of Cambodia were still suffering from post-conflict insecurity and the institutional and human resource capacity of the ARD sector was very low.

While the completed projects were rated successful or partly successful and ongoing projects were generally being implemented satisfactorily, the IED evaluation concluded that the experience was quite varied across the four types of intervention:

- "ARD policy and sector management programs and projects have been highly relevant, effective, mostly less efficient, are likely sustainable, and have contributed substantially to countrywide impacts.

- Rural infrastructure projects have also been highly relevant, effective, efficient, and are likely sustainable with potential resolution of maintenance issues. They have also contributed significantly to countrywide impacts.

- Large irrigation projects have been less relevant, are likely to be considered less effective when evaluated, have generally been less efficient, and are likely to be sustainable but only if special efforts and inputs are provided. They will make little contribution to countrywide impacts.

- Targeted rural development projects were relevant in concept but not in design. As a result, they are likely to be less effective, less efficient and of uncertain sustainability. They are not likely to contribute significantly to either region-wide or country-wide impacts." (IED, 2009c, p. iii)

In effect, the first two categories of projects were countrywide in scale, sequenced to build systematically on their predecessors and focused on a combination of policy reforms and sector investments. According to the IED evaluation, the ADB-funded interventions in these two program areas substantially reformed the ARD policy framework in Cambodia, including a far-reaching reform of the land titling system, which enabled the subsequent provision of land titles for over 1 million properties out of about 3 million, with World Bank support. In addition, they led to improved rural transport access, and improved health, education, agricultural extension and water-sanitation services in 11 out of 24 provinces. On the other hand, the irrigation projects fell short in reaching their limited target areas, and the targeted rural development projects involved a complex set of many small investments in local services and income generation activities, whose scale will likely remain limited, serving only a small fraction of rural communes countrywide.

While giving credit to ADB for its early engagement and successful sequencing and scaling up of ARD operations in Cambodia, the IED evaluation faults ADB for not maintaining momentum in areas where it had been most successful and consolidating progress with further follow-up projects. Instead, ADB moved to other project areas beyond its core competencies, with complex design and limited scale, and without adequate learning from prior experience. The IED evaluation therefore recommends the following for the future:

- "ADB should build on and consolidate past successes within the sector, upscaling similar designs and implementation arrangements to other parts of the country particularly in terms of expanding rural infrastructure, firming up maintenance and building institutional capacity.

- ADB should look for ways to support the continuing and unfulfilled demand for rural credit particularly building synergies with financial sector operations, identifying needs and partnering with other institutions with experience in successful microfinance programs.

- Greater use should be made of the sector development program modality particularly undertaking rural infrastructure, water resource management and strengthening related institutional arrangements." (IED 2009c, p. vi)

In sum, this set of ADB-funded programs suggests the following:

- Successfully scaled up programs involved project design that benefited from relatively focused interventions, drawing on the core competencies of the agencies supporting them (both government and donor agencies), and representing sequential projects that built systematically on the successes of preceding ones.

- Less successful programs involved limited vision of scale from the start, a lack of learning from prior experience, discontinuities in governmental agencies and complex project design involving many small and disconnected interventions.

- Continuing on successful scaling up pathways with long-term engagement is a key requirement for high impact and lasting success.

14. Peru—Rural Highland Area Development Program with IFAD support[46]

Since 1980 until today, IFAD has supported the development of smallholder farming in the highlands of Peru, initially in a context of conflict and post-conflict. Over this period, IFAD systematically scaled up a program of rural area development through eight successive project loans and active engagement with a range of Peruvian partners. Aside from horizontal scaling up to reach a wider geographic scale (eventually covering 1,610 communities, or about 30 percent of Peru's poor highland community and households), the program also scaled up functionally by broadening its scope beyond farming to non-farm activities, employment and institutional capacity building at the sub-national level. More recently, the program involved vertical scaling up by engaging Peru's Ministry of Agriculture and helping to develop a national program based on the IFAD-supported approach.

The program started out as a bottom-up, community based initiative, driven by an alliance of national experts, community activists, and the IFAD country program manager, who helped formulate and push forward the concept and its implementation. Competitions among communities for the best programs and resources created a powerful incentives system, reinforced by the program's internal M&E and by intermittent external evaluations. Fiscal constraints were not binding since project costs were deliberately kept low. Political, ownership, institutional and learning spaces were created in a process of continuous learning and adaptation. The project involved active outreach and networking among national program partners and experts, the establishment of a supportive legal framework and efforts to garner the support of the national government. An important element of the program involved a flexible response to Peru's diverse cultural space:

> "One of the key ingredients of success of the Peru program was the great cultural compatibility of the rural development model chosen with the cultural norms of the highland population. The community based approach was particularly well suited to the cultural traditions of the Andean population in the south, less so in the north, where there are fewer Andean communities. As a result, the SIERRA NORTE project adapted the operational model to rely more on municipal and private sector agents (including the mining sector), rather than insisting on the model of community driven development that was so appropriate in the south." (Linn et al., 2010, p. 15/6)

Other sectors

15. Afghanistan—micro-finance program supported by the World Bank[47]

Based on a 2002 joint needs assessment by the Asian Development Bank, UNDP and the World Bank, the World Bank took on the task of supporting the setting up a national apex institution for the development and strengthening of the fragmented microcredit system in Afghanistan. The program consisted of the establishment of a performance-based funding mechanism for individual retail microfinance institutions, technical assistance and training. Over time the funding was intended to shift from predominantly grant financing for retail organizations to loan financing— a goal on which substantial progress was made over the period 2003-2008. The program aimed to achieve national scale and coverage quickly and good progress was made toward that goal, except in the south of the country where insecurity and military conflict made its operation impossible. Resources were channeled through a total of 11 client microfinance institutions (MFIs), the largest of which was the local branch of BRAC, which in 2006 accounted for 80 percent of total borrowers and 60 percent of microloans outstanding.

One relevant dimension of the Afghanistan microfinance program was that it built explicitly on the experience of a similar apex institution approach that had been broadly successful in the post-conflict conditions of Bosnia-Herzegovina. Key elements which carried over to Afghanistan were that:

- microfinance was identified as a main priority in the reconstruction planning process;
- multi-donor cooperation took place from the beginning;
- a performance-based funding mechanism provided incentives for retail MFIs;
- grant-financed technical assistance and capacity building was a core component of the program;
- the program supported legislative and regulatory reform for the MFI sector (vertical scaling up); and
- the program was designed and managed to stay non-political.

The Afghanistan program therefore represents an example of transnational scaling up.

While successful overall, the program encountered a number of challenges:

- an inflationary context created difficulties for the microfinance operations;
- the lack or weak infrastructure in the country combined with insecurity and conflict in some parts made access costly, difficult and in some cases, impossible;
- staffing constraints created obstacles;
- religious and cultural sensitivities had to be respected in the design and implementation of the program;
- tensions developed between the goal of quick delivery of resources to microfinance clients on the one hand and the strengthening of institutional capacity on the other; and
- tensions also developed between the objective of achieving social development goals through microfinance activities versus the development of a financially sound and self-sustaining microfinance system.

16. Timor-Leste—the Transition Support Program, supported by the World Bank and other donors[48]

The TSP represented a pooling of donor resources in support of Timor-Leste's post-conflict transition, which helped to foster the viability of the state and to establish a mechanism for longer term development planning. The funds were intended as bridging finance, to cover recurrent expenditures through budget support, until oil and gas revenues later came on stream.

Annual negotiations of yearly "tranches" between the government and the donor group were arranged to agree on key aspects of the national development plan to be implemented and on policy measures to be taken. Based on these agreements an action matrix was developed, specifying policy measures and activities to be undertaken alongside target completion dates.

The TSP made a positive contribution to development outcomes across many key areas, including improving governance and the delivery of basic services to the poor, addressing the negative spillover effects from violent conflict and disease, and increasing economic growth and reduced poverty. Over time the focus of the program shifted more attention to security and job creation and from short-term transition management to longer-term development. Overall, the TSP is viewed as a successful way of supporting Timor-Leste's post-conflict transition to stability and effective self-government.

The key elements of this success were:

- the program was based on a careful diagnosis and a realistic assessment of capacity constraints and what could be expected from capacity building efforts;
- the intensive engagement of a cohesive and pragmatic government with sound developmental goals, which in turn helped create a strong sense of national ownership for the TSP;
- the program provided a mechanism for development planning, and hence for sectoral scaling up initiatives;
- it explicitly incorporated security aspects (for instance, reintegration of combatants);
- it was designed for flexibility and adaptation; and
- it facilitated donor coordination.

17. *Ethiopia– Productive Safety Nets Program with World Bank and DFID support*[49]

The government of Ethiopia's (GoE) Productive Safety Nets Program (PSNP) was established in 2005, initially for a five-year period. The PSNP operates on an annual budget of around $300 million and reaches over seven million people across seven of the country's 10 regions. It aims to 'graduate' people from food insecurity through an equal mixture of food and cash transfers and to build community assets via associated public works schemes. There are two transfer modalities: a public works (conditional) transfer and an unconditional transfer. Around 6.2 million people or 80-90 percent of program participants are expected to contribute their labor in return for the transfers. Those who qualify for the unconditional transfers not only suffer from chronic food insecurity but also lack labor and other sources of support.

The PSNP was initially designed by the World Bank and has been supported by a number of donor agencies. The program is run by the Food Security Coordination Bureau, located within Ethiopia's Ministry of Agriculture.

The PSNP is generally perceived to have been a successful program to date. Although there is little evaluation data as yet, early results include some evidence: that levels of food insecurity among recipients is declining; that a majority of recipients were able to avoid selling of assets; that around a quarter were able to use funds to gain new assets and skills; and that PSNP cash is enabling households to invest in education and health (Devereux et al, 2006; Regional Hunger and Vulnerability Programme, 2007). Its extension for at least one further five-year period seems to be assured.

The following observations can be made about PSNP:

- PSNP is an example of a project that moved immediately to scale without piloting or phasing. While some donors, including DFID, initially insisted on a phased approach as a 'non-negotiable' aspect of the program design, including with an initial piloting of the program, the GoE pushed back against these demands and decided on an immediate national-level roll-out.

- Ownership of the PSNP is strong and helped drive the program's introduction at scale. There is a high level of political commitment within GoE to the PSNP, including at the top levels of political leadership, where there is a belief that this program can help stimulate rural growth. However, GoE also sees PSNP as a temporary measure that will become obsolete once growth becomes more productive and pro-poor. This is contrary to donor perceptions of the PSNP as a permanent program which can support Ethiopia's social development.

- DFID has committed resources and time to promoting the PSNP within GoE. Much of DFID's efforts have been focused on expanding the policy space for the promotion of social protection, and improving the information basis for designing and evaluating the program. DFID's framing of social transfers and social protection as 'an African success story', particularly via study tours, has been an important factor in persuading government officials to increasingly adopt what was originally seen as a donor-driven agenda.

- One weakness of PSNP's design is in its financing arrangements. The PSNP does not fall within the normal framework for financial management of public services but operates through a unique grant guideline process, known as a federal specific grant. This means that rather than each region gaining its allocation in accordance with universal principles, regions instead nominate the number of beneficiaries for the PSNP on an annual basis and use this to negotiate with the federal government. This discretionary process has the potential for patronage politics to come into play, prevents the program from being an entitlement-based approach, and allows for accusations of regional bias.

ANNEX 4: SUMMARIES OF COUNTRY PROGRAM CASE STUDIES

The World Bank

In recent years the World Bank Group's Independent Evaluation Group (IEG) carried out six evaluations of country programs in fragile states, each covering between five and eight years of World Bank program results: Bosnia-Herzegovina (1996-2003), Ethiopia (1998-2006), Cambodia (1999-2006), Timor-Leste (2000-2005), Uganda (2001-2007) and West Bank and Gaza (2001-2009). In the case of Timor-Leste, the evaluation covered a series of projects in two sectors only (community development and agriculture), not the entire World Bank Program. In 2006, IEG also carried out an overall evaluation of World Bank engagement in fragile states.

None of these evaluations carried out a systematic assessment of whether and how the World Bank pursued a scaling up approach in it programs or projects. However, the evaluations offer some examples of more or less successful scaling up (type 1 and type 2 errors). They also provide insights on how the conditions of these fragile states might have supported or impeded potential scaling up in terms of drivers and spaces. They also offer a perspective on the World Bank's operational modalities as they would support or impede scaling up. Finally, they allow some assessment of outcomes in terms of the criteria used by IEG, which also allow some indirect judgment on the scalability of the World Bank supported operations.

Scale and dimensions of the World Bank operations

In all six countries, the World Bank operated at the national, provincial and local levels. Even where projects were limited in terms of investments to specific localities, it virtually always attempts to scale up vertically because of its focus on policy and institutional reform.

Drivers and spaces

The country program evaluations do not permit a systematic assessment of *drivers* of scaling up. However, they do allow an assessment of the role of a number of *spaces* in permitting or constraining World Bank operations at scale.

- *Fiscal and financial space* was identified as an issue in virtually all cases, particularly in the ability of projects and programs to continue in the absence of external aid. The World Bank's efforts to support increased fiscal and financial capacity in most cases did not show enough progress over the time horizon of the programs evaluated to provide significant fiscal and financial space.
- *Institutional space* was a major constraint virtually everywhere; the World Bank expended significant effort to strengthen institutional capacity in all cases, but at best with mixed success.
- *Political space* was a significant problem in many of the countries. For example, in the case of Bosnia-Herzegovina, the fragmented and highly contentious ethnic politics made program design and implementation highly complex; in the case of West Bank Gaza, political tensions between Israel and the Palestinian Authority were a continuing source of disruption, as were intra-Palestinian divisions. In Ethiopia, the evaluation concluded that the World Bank should have done more by way of political analysis to inform its programs. The overall evaluation urged more political analysis as an input into the World Bank's strategies for work in fragile states.

- *Partnership space* also created challenges in all cases. The World Bank addressed these in various ways, at the level of overall aid coordination where it often took a lead, through bilateral partnerships (e.g., with the African Development Bank in the case of Uganda), and through the creation of trust funds and extensive co-financing at the program and project level. In most cases, the evaluations give the World Bank credit for pursuing the partnership agenda with increasing degrees of intensity and success. In the World Bank's engagement in Timor-Leste, however, the lack of effective partnership building led to duplication, lack of cooperation, success and sustainability of the programs supported. The overall evaluation saw good donor cooperation at the international level, but this was not generally reflected at the country level.

- *Cultural space* was identified as an issue in Timor-Leste, where the World Bank failed to consider the local conditions adequately in trying to apply the successful Indonesian model of community driven development onto Timor-Leste's different local social, cultural and political traditions. In Ethiopia, the evaluation concluded that more familiarity with and analysis of the local cultural conditions would have been helpful.

- In all cases except for Uganda, lack of *government ownership* was identified as a significant problem limiting the success of World Bank funded programs. In the West Bank and Gaza, the evaluation pointed to a positive shift in ownership from the first half of the decade to the second. In those cases, where successful project results were achieved and in particular in the successful scaling up examples, ownership was clearly present. The overall evaluation of engagement in fragile states also found that that lack of common goals between governments and the World Bank was a serious obstacle.

- *Security space* was not a significant concern in the six countries, except in the case of West Bank and Gaza, where the violence of the Intifada and recurrent conflict between Israelis and Palestinians proved very disruptive over the 20 years of World Bank engagement. The overall evaluation of engagement in fragile states observed that where conflict is in progress, security concerns create serious obstacles for the World Bank's work.

Modalities of World Bank engagement

The evaluations, and the few success stories which they report on specifically, offer some insights into the nature of World Bank operational modalities and to what extent they shaped the degree of success, sustainability and scalability of outcomes.

- *Early engagement* during and immediately after a major conflict was identified as major positive factor in Bosnia-Herzegovina and the West Bank and Gaza, but in the case of Timor-Leste, this did not prevent a failure of World Bank programs and performance.

- *Analytical work* was highlighted in all cases as an important area of significant contribution by the World Bank, although with varying degrees of impact. In Ethiopia, the World Bank should have done more social, political and cultural analysis. The overall evaluation of engagement in fragile states stressed the importance of political analysis.

- *Funding arrangements* were highlighted in most evaluations. In post-conflict situations, the World Bank generally established trust funds and searched for co-financing by other donors of its program in the interest of reaching greater scale. The evaluations generally applauded these efforts. Budget support was a frequently used tool with mixed success, as were SWAps, which were successful in Uganda but less so in Ethiopia.

- *Streamlined operational procedures* for processing emergency operations in post-conflict situations were a major factor in allowing the World Bank to respond much more quickly than under ordinary circumstances in both Bosnia-Herzegovina and Timor-Leste. In the former case, this was not to the detriment of the quality of operations and fiduciary aspects, while in the latter case the evaluation attributed significant quality and fiduciary problems to the fact that shortcut procedures were applied. The overall evaluation observed that the World Bank's standard processing procedures often created serious delays.

- *Project Implementation Units* were the preferred mode of operation in all six countries. Evaluations attributed this to the weak local institutional capacities, but highlighted the problems this created in terms of further weakening state institutions and sustainability of programs. The evaluations recommended early integration of PIU functions into ministerial structures.

- *Staffing* issues are highlighted in a number of country evaluations, including the frequent turnover of project team leaders and the lack of adequate field presence of expert staff. In Ethiopia and Uganda, the presence of expert staff in-country was cited as a reason for successful outcomes. The overall evaluation of engagement in fragile states expressed concern about staff numbers, quality and incentives.

- *Continuity of engagement* is highlighted as an issue in some cases. In Timor-Leste, the World Bank's engagement in a sequence of projects over four years in two sectors (community development and agricultural development) should have been a plus, but the absence of any clear pathway between the interventions led to poor and unsustainable outcomes. In the West Bank and Gaza, the World Bank was engaged over a 20 years and stayed engaged even during periods of crisis, yet its engagement in some sectors was intermittent. In Bosnia-Herzegovina, the systematic sequencing of projects was identified as a strength of the program, but the dramatic reduction of the World Bank's engagement after the initial years of exceptionally large financial support was seen as a major risk for longer term impact. In the cases of successful scaling up, sustained engagement was clearly a major factor. The overall evaluation of engagement in fragile states also expressed concerns over premature down-scaling of resources after an immediate post-conflict situation has passed.

- *Monitoring and evaluation* was highlighted as weak in a number of cases, especially in the projects of Timor-Leste, where the evaluations identified insufficient learning from one project to the next as a major problem. The overall evaluation of engagement in fragile states called M&E in fragile states "negligible."

- *Tensions and dilemmas* were noted in a number of cases. In post-conflict settings including Bosnia-Herzegovina, the West Bank and Gaza and Timor-Leste, the tension between the quick delivery of financial support and the pursuit of long-term development outcomes led to the conclusion that too much attention was devoted to crisis response with detrimental impacts on long-term development outcomes. In Timor-Leste, the evaluation attributed this to the intense pressure on the World Bank by local and international politicians, as well as by other donors, to show quick impact in the face of humanitarian emergencies and the destruction of essential infrastructure.

Outcomes

The outcome ratings of the program evaluations are shown in Table 5. Overall, the outcomes were mixed with "satisfactory" or "moderately satisfactory" ratings in three countries, and "unsatisfactory" or "moderately unsatisfactory" ratings in three countries.

Table 5: World Bank country program ratings

	Bosnia-Herzegovina 1996-2003	West Bank/ Gaza 2001-2009	Timor-Leste 2000-2005	Cambodia 1999-2006	Ethiopia 1998-2006	Uganda (joint with AfDB) 2001-2007
Overall	Satisfactory	Moderately unsatisfactory	Unsatisfactory for CEP; moderately satisfactory for ARP 1; moderately unsatisfactory for ARP2.	Moderately satisfactory	Moderately unsatisfactory	Moderately satisfactory
Macroeconomic		Unsatisfactory		Moderately satisfactory	Moderately satisfactory	Satisfactory
Governance		Moderately satisfactory		Unsatisfactory	Moderately unsatisfactory	Moderately satisfactory (but anti-corruption moderately unsatisfactory)
PSD	Unsatisfactory	Moderately Unsatisfactory		(Problem area)	Moderately unsatisfactory	Moderately satisfactory
Agriculture			Mixed ARP 1: Moderately satisfactory; ARP 2: Moderately unsatisfactory	Moderately unsatisfactory		Moderately satisfactory
CDD			unsatisfactory			
Health	Satisfactory	Moderately satisfactory		Satisfactory	Moderately satisfactory	Moderately unsatisfactory
Education	Satisfactory	Moderately Unsatisfactory		Satisfactory	Moderately satisfactory	Moderately satisfactory
Infrastructure	Satisfactory	Moderately unsatisfactory		Satisfactory	(Roads OK)	(Roads OK)
Sustainability	OK at project level, but problem overall due to funding drop	No (political, financial)	CEP: unlikely ARP: unlikely	Signifcant risk	High risk	Substantial risk

Particularly notable is the "satisfactory" rating of Bosnia-Herzegovina, which emerged in 1996 from a deadly and protracted civil war that left great destruction and deep political divisions.[50] The World Bank's large scale engagement, which systematically built programs at scale in various sectors over time with careful and ambitious co-financing and partnership agreements leveraging the World Bank's resources, was a significant factor making this outcome possible.

The "unsatisfactory" ratings were particularly disappointing in Timor-Leste since they were attributed to an unsatisfactory performance of the World Bank, albeit under very difficult conditions. The fact that the project sequences under evaluation failed to be grounded in effective learning and partnerships reflects the importance of an effective approach to design and implementation.

In terms of outcomes by sectoral area of engagement, the two areas that consistently were the most difficult are governance and private sector development, with anti-corruption measures and improvements in the business climate particularly elusive. In contrast, macroeconomic stability, education, health and infrastructure were sectors in which positive outcomes were regularly achieved. Agriculture and rural development showed mixed results. The overall evaluation of World Bank engagement in fragile states noted relatively successful outcomes in macroeconomic stability and infrastructure interventions. It also noted that community driven development projects tend to disburse fast, but generally have only limited community development impact.

The evaluations gave near universally low ratings to the sustainability of programs and outcomes.

Special aspects of scaling up

The evaluations recorded a few cases where successful scaling up of projects or programs was demonstrated. The cases span a wide range of examples, including:

- *an education* support program in Cambodia, which involved testing new educational approaches, systematically scaling them up horizontally and vertically, with good government ownership, effective partnerships, and supported by locally-based expert staff;
- the *demobilization* program in Ethiopia achieved the virtually complete demobilization and integration of almost 150,000 combatants over a 5-year period and helped provide jobs, housing and services for the ex-combatants and their families;
- a sequence of projects in Uganda by the World Bank and the African Development Bank in the *road sector* over some eight years; and
- a *solid waste* project in West Bank and Gaza, where the persistent efforts by the World Bank team to engage Israeli and Palestinian counterparts at all levels led to the establishment of a large solid waste treatment plant which served an important region in the West Bank, with support maintained by the World Bank throughout the Intifada period and with demonstration effects that led to the establishment of cooperation among local government leaders in tacking the solid waste problems of the territory.

The overall evaluation concluded that the World Bank ought to practice greater selectivity and prioritization and adopt a long-term perspective. Moreover, "appropriate sequencing of reforms and sufficient time to implement them are crucial for achieving results without overwhelming country capacity."(IEG, 2006a, p. x)

Judging from the evaluations reviewed, the World Bank's weakness in the countries concerned typically did not involve type 1 errors (insufficient pursuit of scaling up), but more generally involved type 2 errors (where scaling up is pursued, but done in the wrong way). In cases with unsatisfactory and unsustainable outcomes, the programs and projects failed to pursue systematic scaling up pathways in terms of creating ownership, overcoming fiscal, political and governance constraints, assuring effective learning from experience and building lasting partnership. In some cases, wrong models were pursued, as in the case of the voucher privatization program in Bosnia-

Herzegovina, the community development program in Timor-Leste, or of the fertilizer reform program in Ethiopia. The cases in which programs were successfully sustained and carried to scale invariably involved the development of sector strategies, strong ownership and partnerships, sustained engagement and effective deployment of technical staff in the field.

The Asian Development Bank

In recent years, the ADB Independent Evaluation Department (IED) carried out three evaluations of country programs in fragile states: Cambodia (1998-2008), Nepal (2004-2009), and Pakistan (1985-2006). In addition, IED evaluated the overall experience of ADB in fragile and conflict affected states during the period 2005-2009.[51] IED also carried out an evaluation of ADB's agricultural and rural development (ARD) sector program in Cambodia.

As in the case of the World Bank, none of these evaluations carried out a systematic assessment of whether and how ADB pursued a scaling up approach in it programs or projects. However, the Cambodia country program evaluation has assessments and recommendations linked closely to the scaling up concept and the Cambodia ARD sector evaluation offers some examples of more or less successful scaling up (type 1 and type 2 errors). And as is the case with the World Bank, the IED evaluations provide various insights into scaling up in terms of the environments described, the operational modalities employed and the outcomes achieved.

Scale and dimensions of the ADB operations

In all three countries, the ADB operated at the national, provincial and local levels. The ADB also looked to scale up vertically, wherever possible, to embed broader policy and institutional reforms.

Drivers and spaces

As in the case of the World Bank, ADB country program evaluations did not permit an assessment of *drivers* of scaling up. However, among the *spaces* needed for successful outcomes, virtually all created difficulties to varying degrees:

- *Fiscal and financial space* was identified as an issue in virtually all evaluations. The overall evaluation of ADB engagement called for systematic efforts to help close the resource gap.
- *Institutional space* was generally also a major constraint, with weak human resources an important factor.
- *Political space* presented a mixed experience: In the case of Cambodia, political stability was cited as a positive factor, while in the case of Nepal, political instability and labor strikes were a major obstacle.
- *Partnership space* also created challenges in virtually all cases. According to the evaluations, the ADB made significant efforts to partner with other institutions (including, and especially the World Bank) and to participate in or lead donor coordination efforts. But the evaluations generally conclude that intentions were stronger than actual implementation on the ground.
- *Cultural space* was not specifically identified as an issue in the ADB evaluations.
- In Cambodia and Nepal, alignment with country priorities, and hence *ownership space*, was perceived as strong. This is in stark contrast to the evaluations for the World Bank, where in most cases a lack of government ownership was identified as a significant problem limiting the success of programs.

- *Security space* was a significant concern early in the Cambodia program and throughout the period of the evaluation in Nepal. The overall evaluation flagged security issues as a constraint on the ADB's ability to deploy consultants on the ground.

Modalities of ADB engagement

The evaluations offer some insights into the nature of ADB operational modalities and to what extent they shaped the degree of success, sustainability and scalability of outcomes.

- *Early engagement* by ADB in Cambodia after the domestic conflict was identified as having been helpful.
- *Analytical work* generally played a role in supporting ADB's engagement in fragile states, and joint assessments of post-conflict reconstruction needs were seen as useful.
- *Funding arrangements* were highlighted in most evaluations. SWAps were successful in Cambodia (education), Nepal (education), Solomon Islands (infrastructure) and PNG (roads, health). Falling project size was noted as an issue in Cambodia and a proliferation of many small projects was identified as a problem in Pakistan. In Pakistan, the evaluation recommended multi-tranche financing facilities and programmatic approaches.
- *Streamlined operational procedures* for processing loans in post-conflict situations were identified as a problem in Afghanistan and Timor-Leste in terms of the quality of project preparation and the effectiveness of fiduciary standards.
- *Staffing* constraints were highlighted as an issue in the overall evaluation. In addition, insufficient delegation of responsibility to resident mission offices was generally seen as a problem.
- *Continuity of engagement* was perceived as a general feature of ADB involvement in fragile states. However, the failure to continue with successful interventions was identified as a problem in agricultural and rural development projects in Nepal, although other donors picked up some of these.
- *Monitoring and evaluation* was generally not a focus of the IED evaluations. The Cambodia and Nepal evaluations characterized M&E as weak and insufficient, respectively.
- *Tensions and dilemmas* were noted mostly in terms of the tradeoff between speed of response and quality of preparation.

Outcomes

The country program ratings of the ADB overall evaluation are shown in Table 6. Overall, the outcomes were generally positive, with successful ratings in three current FCAS and in three exited FCAS and with partly successful ratings in the remaining countries. Sustainability was ranked on average between "likely" and "less likely" with six out of 12 cases closer to "less likely". The overall evaluation of ADB engagement and the Nepal evaluation note problems of sustainability of infrastructure investments as a result of poor maintenance.

In terms of outcomes by sectoral area of engagement, governance and private sector development (the two areas that consistently were the most difficult in the case of the World Bank) are rated as "successful" or "partly successful" in the case of the ADB. The "successful" ranking on governance for Cambodia may in part be explained by the fact that the ADB did not try and tackle corruption issues, which the evaluation acknowledges has being endemic. Particularly noteworthy are the successful ratings for rural credit support in Cambodia and Nepal. Agriculture and rural development showed mixed results while support for

Table 6: ADB country program ratings

Item	WA	Rating	Relevance	Effectiveness	Efficiency	Sustainability
CURRENT FCAS	**1.4**	**PS**	**1.9**	**1.5**	**1.0**	**1.4**
Afghanistan	1.5	PS	1.9	1.5	1.3	1.5
Federated States of Micronesia	1.1	PS	1.3	1.0	1.0	1.0
Papua New Guinea	1.5	PS	2.0	1.4	1.2	1.4
Republic of Mashall Islands	0.8	PS	2.0	0.3	0.7	0.7
Solomon Islands	1.7	S	2.0	2.0	1.0	2.0
Timor-Leste	1.6	S	2.0	2.0	1.0	1.5
Tuvalu	1.7	S	2.0	2.0	1.0	2.0
EXITED FCAS	**1.7**	**S**	**2.2**	**1.8**	**1.5**	**1.5**
Azerbaijan	1.8	S	2.0	2.0	2.0	1.0
Tajikistan	2.1	S	2.6	2.2	2.0	1.8
Uzbekistan	1.5	PS	2.0	1.4	1.4	1.4
Lao PDR	1.7	S	2.0	2.0	1.0	2.0
Nepal	1.5	PS	2.3	1.4	1.2	1.3
Overall Average (Current and Exited)	1.6	S	2.0	1.6	1.3	1.5

FCAS = fragile and conflict-affected situations. Lao PDR = Lao People's Democratic Republic. WA = weighted average.

a The ratings follow the ADB. 2006. Guidelines for Preparing Performance Evaluation Reports for Public Sector Operations. Manila. In calculating for overall rating, the weights for the criteria are as follows: relevance (20%), effectiveness (30%), efficiency (30%), and sustainability (20%). Highly successful = WA is greater than or equal to 2.7; Successful = WA is greater than or equal to 1.6 and less than 2.7; Partly successful = WA is greater than or equal to 0.8 and less than 1.6; Unsuccessful = WA is less than 0.8. Ratings for the criteria are rounded to the nearest whole number. Relevance: 3 = highly relevant, 2 = relevant, 1 = less relevant, 0 + irrelevant. Effectiveness: 3 = highly effective, 2 = effective, 1 = less effective, 0 = ineffective. Efficiency: 3 = highly efficient, 2 = efficient, 1 = less efficient, 0 = inefficient. Sustainabliity: 3 = most likely, 2 = likely, 1 = less likely, 0 = unlikely.

Source: IED, 2010

education programs has generally been successful. The ADB's support for infrastructure in fragile states also has a mixed track record. Nonetheless, the overall evaluation of ADB engagement assesses infrastructure as a sector that is particularly promising for ADB support in fragile settings.

Special aspects of scaling up

Throughout the ADB reports, there are elements of the analysis and recommendations which are consistent with a scaling up approach, including:

- Concern about lack of program selectivity, proliferation of project numbers, declining or small project size;
- Calls for "step-by-step" sector strategies, road maps, gap analysis, umbrella technical assistance operations, multi-tranche financing facilities and programmatic approaches;
- Stress on long-term perspective, careful sequencing and continuity in engagement.

The most explicit statement is from the Cambodia country program evaluation, which calls for a "road map for ADB assistance that traces the chain of results from ADB support to well-defined national targets and results" (p. 15).

The ARD sector evaluation for Cambodia presents a particularly striking case study of scaling up, in which there were a sequence of loans in various ARD subsectors, creating both type 1 and type 2 errors (see Annex 3).

The Global Fund

The Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM) has supported scaling up of programs to fight HIV/AIDS, tuberculosis and malaria in developing countries since 2002. A recent evaluation published in *Global Health Governance* in 2010 reports on the performance of 198 grants to 41 fragile states with a total disbursement of $4.9 billion (Bornemisza et al, 2010).[52] The evaluation considers six criteria:

- percentage of targets reached;
- disbursement rating;
- continued funding for Phase Two;
- continued funding beyond Phase Two (five years);
- M&E Ratings; and
- independent third party ratings of data quality and reporting systems.

Contrary to earlier evaluations that had found no significant differences between GFATM operations in fragile and stable states (in particular, the GFATM's own report (GFATM, 2005)), this evaluation concludes that there were significant differences between the two types of country programs.

a. While roughly the same number of grants was made to both fragile states as to other countries, the per-country grant size was almost two-thirds higher in fragile states.

b. Two-thirds of programs in fragile states were administered by multilateral organizations (principally the UNDP), rather than by national bodies, as compared with only one-third in other states.

c. Grants in fragile states achieved a high rate of their targets at 83 percent, but this was (statistically) significantly less than the rate for other countries (88 percent).

d. Grant disbursement was rated satisfactory for 79 percent in fragile states, as compared with 85 percent in others.

e. Seven grants (4 percent) did not get extended to the second phase in fragile states, as against three (1 percent) for others.

f. Fourteen grants (18 percent of total applications) were extended beyond the first five years in fragile states, compared with 42 (32 percent) for other countries.

g. Ratings for the quality of M&E was substantially lower in fragile states, with only 47 percent rated satisfactory, compared with 67 percent in other countries.

h. Independent onsite verification of data quality and reporting systems showed 64 percent satisfactory in fragile states, as against 78 percent in other countries.

i. For all indicators the performance of GFATM grants was worse in fragile states "affected by humanitarian crisis", than other fragile states.

With 40 percent of its funding going to fragile states, the GFATM is heavily invested in these countries. Overall, GFATM programs in fragile states perform remarkably well in terms of reaching their targets, their disbursement rate and their extension rate from Phase One to Phase Two. Performance has been less strong for M&E, for data quality, and especially for grant extension beyond the first five years. In all performance categories, GFATM programs performed significantly less well in fragile states than those in stable environments, but the difference was modest. A similar difference can be found between programs in fragile states affected by humanitarian crises and those unaffected, the latter performing marginally better. A majority of grants in fragile states have to be administered by external (multilateral) organizations, reflecting the relatively weak development of local institutions, presumably reflecting the weakness of local institutions.

Thus, while fragile states, as expected, create somewhat greater challenges than other countries for the scaling up strategy of the GFATM, successful scaling up does take place in four-fifths of fragile states. Three areas are identified as requiring special attention in fragile states: data quality and reporting systems, M&E, and country ownership and administration by national agencies.

REFERENCES

Ahmadzai, H. and F. Kakar, M. Rashidi, P. Suarez, O. Ameli, A. Hartman (2008). "Scaling up TB DOTS in a fragile state: post-conflict Afghanistan." *The International Journal of Tuberculosis and Lung Disease*, 12-2, 180–185.

African Development Fund (2008). "Strategy for Enhanced Engagement in Fragile States." Tunis.

Baird, M. (2010). "Service Delivery in Fragile and Conflict-Afflicted States." World Development Report 2011 Input Paper, The World Bank.

Bornemisza, O. and J. Bridge, M. Olszak-Olszewski, G. Sakvarelidze, J. Lazarus (2010). "Health Aid Governance in Fragile States: The Global Fund Experience." *Global Health Governance,* Volume IV, No. 1.

Bulir, A. and A. J. Hamann (2006). "Volatility of Development Aid: From the Frying Pan into the Fire?" International Monetary Fund Working Papers, 06-65.

Burnside, C. and D. Dollar (1997). "Aid, policies, and growth." The World Bank Policy Research Working Paper Series, 1777: 847-868.

Chandy, L. and G. Gertz (2011). "Poverty in Numbers: The Changing State of Global Poverty from 2005 to 2015." Global Views No. 18, The Brookings Institution.

Clemens, M. (2004). "The Long Walk to School: International Education Goals in Historical Perspective." Center for Global Development Working Paper Series, 37.

Collier, P. and D. Dollar (1999). "Aid allocation and poverty reduction." The World Bank Policy Research Working Paper Series, 2041.

Daly, E. (2003). "Performance Based Partnership Agreements in the Transitional Context of Afghanistan." International Committee of the Red Cross Discussion Paper.

DANIDA (2005). "A Joint Evaluation: Humanitarian and Reconstruction Assistance to Afghanistan, 2001-2005 from Denmark, Ireland, the Netherlands, Sweden and the United Kingdom." Denmark.

Desai, R. and H. Kharas (2010). "The Determinants of Aid Volatility." Global Economy and Development Working Paper 42, The Brookings Institution.

Devarajan, S. and D. Dollar, T. Holmgren, eds. (2001). *Aid and Reform in Africa: Lessons from Ten Case Studies*. Washington, D.C.: The World Bank.

Devereux, S. and R. Sabates-Wheeler, M. Tefera, H. Taye (2006). *Ethiopia's PSNP: Trends in PSNP Transfers Within Targeted Households.* Sussex: IDS and Addis Ababa: Indak.

Dollar, D. and V. Levin (2005). "The Forgotten States: Aid Volumes and Volatility in Difficult Partnership Countries." Preparatory Paper, DAC Learning and Advisory Process on Difficult Partnership Countries Senior Level Forum.

DFID, (2004). "Improving the Development Response in Difficult Environments: Lessons from DFID Experience." PRDE Working Paper 4.

DFID (2005). "Why We Need to Work More Effectively in Fragile States." London.

DFID (2009). "Promoting Social Transfers: DFID and the politics of influencing," DFID Working Paper Series 32.

DFID (2010a). "Working Effectively in Conflict-Affected and Fragile Situations." London.

DFID (2010b). "Evaluation of DFID country programmes: Yemen." Evaluation Report EV706.

Frontiers in Reproductive Health and Population Council, PATH, Kenya Ministry of Education, Kenya Ministry of Gender, Sports, Culture and Social Services, Kenya Ministry of Youth and Sports (2008). "From Pilot to Program: Scaling up the Kenya Adolescent Reproductive Health Project."

Ghani, A. and C. Lockhart (2008). *Fixing Failed States*. Oxford University Press.

Global Fund to Fight AIDS, Tuberculosis and Malaria (2005). "Global Fund Investments in Fragile States: Early Results." Geneva.

Greeley, M. (2007). "Aid Effectiveness and Microfinance: Lessons from Afghanistan." In J. Manor, ed. *Aid That Works: Successful Development in Fragile States.* Washington, D.C.: The World Bank.

Grindle, M. (2004). "Good Enough Governance: Poverty Reduction and Reform in Developing Countries." *Governance*, Volume 17, Issue 4.

Hartmann, A. and J. Linn (2008). "Scaling Up: A Framework and Lessons for Development Effectiveness from Literature and Practice." Wolfensohn Center for Development Working Papers No. 5, The Brookings Institution.

Hestad, I. and M. Focas Licht (2002). *Lessons: Basic Education in Rural Africa*. Oslo: Save the Children Norway.

Hughes, C. (2007). "The Seila Program in Cambodia." In J. Manor, ed. *Aid That Works: Successful Development in Fragile States*. Washington, D.C.: The World Bank.

IED (2007). "Country Assistance Program Evaluation for Pakistan." Asian Development Bank. Manila.

IED (2009a). "Country Assistance Program Evaluation for Nepal: Delivering Assistance in a Challenging Environment." Asian Development Bank. Manila.

IED (2009b). "Country Assistance Program Evaluation for Cambodia: Growth and Sector Reform." Asian Development Bank. Manila.

IED (2009c). "Sector Assistance Program Evaluation for the Agriculture and Rural Development Sector in Cambodia." Asian Development Bank. Manila.

IED (2010). "Special Evaluation of Asian Development Bank's Support to Fragile and Conflict-Affected Situations." Asian Development Bank. Manila.

IEG (2004). *Bosnia and Herzegovina Post-Conflict Reconstruction and the Transition to a Market Economy*. Washington, D.C.: The World Bank Group.

IEG (2006). *Engaging with Fragile States*. Washington, D.C.: The World Bank Group.

IEG (2006). "Timor-Leste Project Performance Assessment Report." Report No. 36590, The World Bank Group.

IEG (2008). *Ethiopia: Country Assistance Evaluation, 1998–2006*. Washington, D.C.: The World Bank Group.

IEG (2010a). *Cambodia: An IEG Country Assistance Evaluation 1999–2006*. Washington, D.C.: The World Bank Group.

IEG (2010b). *The World Bank Group in the West Bank and Gaza, 2001–2009: Evaluation of the World Bank Group Program*. Washington, D.C.: The World Bank Group.

IEG and OPEV (2009). *Uganda: Country Assistance Evaluation*. Washington, D.C.: The World Bank Group.

International Dialogue on Peacebuilding and Statebuilding (2011). "The Monrovia Roadmap on Peacebuilding and Statebuilding." Monrovia.

Kharas, H. (2008). "Measuring the Cost of Aid Volatility." Wolfensohn Center for Development Working Paper Series No. 3, The Brookings Institution.

Kharas, H. and K. Makino, W. Jung (2011). "Overview: An Agenda for the Busan High Level Forum on Aid Effectiveness." In H. Kharas, K. Makino, W. Jung , eds. *Catalyzing Development: A New Vision for Aid*. Washington, D.C.: The Brookings Institution.

Kulipossa, F. and J. Manor (2007). "Decentralized District Planning and Finance in One Mozambican Province." In J. Manor, ed. *Aid That Works: Successful Development in Fragile States*. Washington, D.C.: The World Bank.

Leader, N. and P. Colenso (2005). "Aid Instruments in Fragile States." PRDE Working Paper 5.

Linn, J. (2011a). "Scaling Up with Aid: The Institutional Dimension." In H. Kharas, K. Makino, W. Jung, eds. *Catalyzing Development: A New Vision for Aid*. Washington, D.C.: The Brookings Institution.

Linn, J. (2011b). "Scaling Up Development Interventions: An Exploratory Review of UNDP's Country Program in Tajikistan." United Nations Development Programme.

Linn, J. and A. Hartmann, H. Kharas, R. Kohl, B. Massler (2010). "Scaling Up the Fight against Rural Poverty." Global Working Papers No. 39, The Brookings Institution.

Manor, J. (2007). "Introduction: Synthesizing Case Study Findings." In J. Manor, ed. *Aid That Works: Successful Development in Fragile States*. Washington, D.C.: The World Bank.

McGillivray, M. (2005). *Aid Allocation and Fragile States*. Discussion Paper 2006/01, UNU-WIDER. Helsinki.

Miguel, E. and M. Kremer (2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica*, Vol. 72, No. 1.

Moreno Torres, M. and M. Anderson (2004). "Fragile States: Defining Difficult Environments for Poverty Reduction." DFID.

Mosko, S. (2003). *Village Learning through Children's Schooling*. Addis Ababa: Save the Children Norway – Ethiopia Programme.

NORAD (2004). "Study of the impact of the work of Save the Children Norway in Ethiopia: Building Civil Society Evaluation." Study 6. Oslo.

OECD DAC (2007). "Principles for Good International Engagement in Fragile States and Situations." Paris.

OECD DAC (2008a). "2011 Survey on Monitoring the Paris Declaration." Paris.

OECD DAC (2008b). "Service Delivery in Fragile Situations: Key Concepts, Findings and Lessons." Paris.

OECD DAC (2011a). "Aid Risks in Fragile and Transitional Contexts." Paris.

OECD DAC (2011b). Creditor Reporting System. Accessed March 2011.

OECD DAC (2011c). DAC Annual Aggregates Database. Accessed March 2011.

OPM and IDL (2008). "Evaluation of the Implementation of the Paris Declaration: Thematic Study – The Applicability of the Paris Declaration in Fragile and Conflict-Affected Situations." DFID.

Pallage, S. and M. Robe (2001). "Foreign Aid and the Business Cycle." *Review of International Economics*, 9(4).

Porter, D. and M. Andrews, J. Turkewitz, C. Wescott (2010). "Managing Public Finance and Procurement in Fragile and Conflict Settings." World Development Report 2011 Input Paper, The World Bank.

Pritchett, L. and F. de Weijer (2010). "Fragile States: Stuck in a Capability Trap?" World Development Report 2011 Input Paper, The World Bank.

Regional Hunger and Vulnerability Programme (2007). "Lessons from Ethiopia on a Scaled-Up National Safety Net Programme." Wahenga Brief No. 14. Johannesburg.

Rosser, A. (2007). "The Transition Support Program in Timor-Leste." In J. Manor, ed. *Aid That Works: Successful Development in Fragile States*. Washington, D.C.: The World Bank.

Sandefur, J. (2011). "Held Hostage: Funding for a proven success in global development on hold in Kenya." Global Development: Views from the Center Blog, Center for Global Development.

Save the Children Alliance (2002). "Learning for Life: Save the Children's Experience with Alternative Basic Education in Ethiopia, Addis Ababa."

Schaffer, R, and R. Ashkenas (2005). *Rapid Results! How 100-Day Projects Build the Capacity for Large-Scale Change*. San Francisco: Jossey-Bass.

Stewart, F. and G. Brown (2010). "Fragile States: CRISE Overview 3." Centre for Research on Inequality, Human Security and Ethnicity (CRISE), Oxford.

World Bank (2004). "World Development Report: Making Services Work For Poor People." Washington, D.C.

World Bank (2005). "Fragile States –Good Practice in Country Assistance Strategies." Washington, D.C.

World Bank (2009). "Contracting out core state functions and services in fragile and post-conflict situations: A transitional arrangement or an option for long-term delivery?" Discussion Note, Fragile and Conflict Affected Countries Group.

World Bank (2011). "The World Development Report 2011: Conflict, Security, and Development." Washington, D.C.

ENDNOTES

1. This paper is concerned with development activities in fragile and low capacity environments. For the sake of brevity, we refer simply to "fragile states" from hereon, using it as an umbrella term which captures the meaning of both terms. See later section on defining fragility.
2. Note that some donors deliberately counterbalance this with longer term commitments to country engagement. However, programming decisions remain narrowly time-bound.
3. The World Development Report 2011 posits that current donor processes are based on assumptions that don't exist in fragile states—basic security, competitive markets and functioning institutions.
4. This description combines elements of the OECD DAC and World Bank definition of fragility.
5. We judge these two indices to be the most credible, widely-used measures of fragility, and to capture different aspects of this broad term. (The Failed States Index is chiefly concerned with measuring state authority and civil cohesion, while the CPIA measures state capacity and policy orientation; legitimacy can be loosely described as a function of these two elements combined, as well as other independent factors that are hard to accurately measure.) These two indices have the advantage of using absolute (as opposed to relative) measures and being updated annually. The full list of countries we classify as fragile is provided in Annex 1.
6. An associated fallacy is that fragility can be equated with post-conflict status, when many fragile states either do not suffer from conflict or their experience of conflict is only one aspect of a broader development problem.
7. See Annex 2. Note that while overall aid to fragile states is increasing, it is concentrated among a few recipients. In 2009, 51% of aid to fragile states was accounted for by just 6 countries.
8. This echoes the distinction between short and long routes of accountability, which was introduced by WDR 2004 and is often used in discussions of public service delivery in fragile states (e.g., Baird, 2010). The short route is directly from provider to beneficiary, while the long route goes through the central government.
9. Note that the discussion here is concerned with core roles of the state, as opposed to services that might equally be provided by non-state actors without undermining the state's legitimacy.
10. See DRC case study in Baird, 2010.
11. A stronger claim is that donors operating in fragile states should shift towards a more venture capitalist approach to programming decisions that accepts higher risks and a higher failure rate on the basis of potentially greater returns. This would lead to a much expanded scope of potential donor activity.
12. It can be argued that risk tolerance among donors has likely reduced in recent years as a result of both the results agenda and the increased level of donors resources, both of which have put pressure on donors to be capable of reporting positive outcomes (however small) and to avoid negative outcomes at any cost.
13. Examples from Burma and Zimbabwe. See DFID, 2010a.
14. Calculations in this section are based on OECD DAC 2008b, OECD DAC 2011b and OECD DAC 2011c. All dollar values given in 2008 prices. See Annex 2 for additional illustrations of aid trends in fragile states.
15. This underestimates the true number of donors, as it does not include donors who do not report to the OECD DAC. However, some aid may be in the form of silent partnerships with delegated authority.

16. See RCA definition in footnote 17.

17. RCA scores capture the share of each donor's aid that is allocated to a country or sector at least in proportion to the donor's share of global ODA. Data is based on 2009 ODA.

18. The difference appears to depend on the period over which aid flows are studied; Kharas analyses aid between 1970 to 2006, while Levin and Dollar focus narrowly on the period 1992 to 2002.

19. Given that the sample sizes for these scores are relatively small, the differences between fragile and stable country scores should be interpreted with caution.

20. See Annex 2.

21. The description of Hartmann and Linn, 2008 used here borrows from a similar summary provided in Linn, 2011b.

22. For instance, in Nepal, DFID and GIZ publicized Basic Operating Guidelines to protect their activities from insurgents and security forces. The guidelines explained the objectives of donor involvement, outlined the values and standards upheld by the aid programs and set minimum conditions required of conflict parties, which included forbidding the theft or diversion of aid resources, allowing staff to work without threat of physical violence, and upholding the observation of humanitarian law and human rights obligations. According to DFID, "the guidelines were an effective tool in helping to establish space for development" (DFID, 2010a, p11).

23. In fact, it may be better to consider a tripartite accountability framework, which involves the state, citizens and donors, from which donors gradually withdraw over time as their role appropriately declines and ultimately vanishes. We are grateful to our peer reviewer for this observation.

24. This echoes the point made earlier which distinguishes the timeframes over which scale and sustainability can each be achieved. Note, we are not implying here that interventions that have a scaling up component are the same ones that are likely to be assessed as higher risk. For a counter-example, see Yemen's Social Development Fund in Annex 3.

25. Note, we are not implying here that interventions that have a scaling up component are the same ones that are likely to be assessed as higher risk. For a counterexample, see Yemen's Social Development Fund in Annex 3.

26. Subsequent developments may lead to different assessments of experience. These are beyond the scope of this desk study, unless stated otherwise.

27. IFAD recently started including an assessment of scalability and/or scaling up performance in its project evaluations. UNDP is another organization that has recently started to include scaling up as a specific criterion in its evaluation approach.

28. Fragile states account for 34 percent of countries with GFATM programs, 40 percent of grants and 46 percent of disbursements.

29. Fragile states classification as described in the introduction. Data is from 2009, the most recent year for which data is available.

30. Fragile states classification as described in the introduction.

31. Note, the fragile state classification used in this paper is broader than that used in Chandy and Gertz, 2011 and thus the poverty shares are marginally higher.

32. Country specified ODA, excluding debt relief.

33. Baird, 2010; Danida, 2005; Daly, 2003.

34. Ahmadzai et al., 2008.

35. Rosser, 2007.
36. Frontiers in Reproductive Health et al, 2008.
37. Linn, 2011b.
38. Baird, 2010.
39. NORAD, 2004. More complete accounts of ABECS can be found in Mosko, 2003; Ingvild & Focas Licht, 2002; Save the Children Alliance, 2002.
40. Sandefur, 2011 and Miguel and Kremer, 2004.
41. Hughes, 2007.
42. Kulipossa and Manor, 2007.
43. DFID, 2010b.
44. Linn, 2011b.
45. IED, 2009c.
46. Linn et al, 2010.
47. Greeley, 2007.
48. Rosser, 2007.
49. DFID, 2009.
50. Despite this strong rating the evaluation report calls the overall outcome in Bosnia-Herzegovina "disappointing" due to weakness in private sector development, remaining governance issues and the questionable sustainability of development achievements.
51. The overall evaluation of ADB's engagement in fragile and conflict affected states (FCAS) covers seven countries in FCAS status at the time of evaluation and five countries that had exited FCAS status. The ADB definition of FCAS status is more restrictive than the one used in this evaluation, since it uses as the only criterion a threshold level in the ADB's Country Performance Assessment (CAP) rating. The evaluation report recommends consideration of a broader, potentially more inclusive set of criteria.
52. The share of countries with GFATM programs that are fragile states is 34 percent; the share of total grants that go to fragile states is 40 percent; and the share of total disbursements is 46 percent.

The views expressed in this working paper do not necessarily reflect the official position of Brookings, its board or the advisory council members.

ISSN: 1939-9383

BROOKINGS

1775 Massachusetts Avenue, NW
Washington, DC 20036
202-797-6000
www.brookings.edu/global