



UNITED NATIONS
UNIVERSITY
UNU-WIDER

WIDER Working Paper 2022/160

Evidence on aid (in)effectiveness in highly fragile states

A synthesis of three systematic reviews of aid to Afghanistan, Mali, and South Sudan, 2008–21

Christoph Zürcher*

December 2022

In partnership with



United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: This working paper provides a summary of three systematic reviews on the effectiveness of aid in Afghanistan, Mali, and South Sudan between 2008 and 2021. These three countries, like all other highly fragile countries, suffer from bad governance, lack of capacity, and violence. The systematic reviews provide robust evidence that aid interventions in precisely those fields are not effective. Aid cannot improve governance, build capacity for central governments, or stabilize the situation. The international aid community can no longer ignore this evidence. We need a change of paradigm in how we support the most fragile states. The prevailing ‘grand vision’ of fixing failed states needs to be replaced with a much more modest approach that aims for local, tangible, small gains.

Key words: foreign aid, effectiveness, systematic review, Afghanistan, Mali, South Sudan

JEL classification: F35

Acknowledgements: The author acknowledges the support of the Policy and Operations Evaluation Department (IOB), Ministry of Foreign Affairs, Netherlands, which commissioned the systematic review on Mali and South Sudan, and the German Federal Ministry of Economic Cooperation and Development (BMZ), which commissioned the review on Afghanistan. The research was supported by Patrick Labelle (research librarian), Luan Borges, Kolby Hoare, Mir Ahmed Javid, Kathryn Kavanagh, Sabrina Sarna, and Emily Woolner (research associates). Caspar Lobrecht of the IOB provided extensive, valuable feedback throughout the research process.

This paper is a revised, extended version of Christoph Zürcher, with P. Labelle, L. Borges, K. Hoare, M.A. Javid, K. Kavanagh, S. Sarna, and E. Woolner (November 2022), ‘Impact of Aid in Highly Fragile States: A Synthesis of Three Systematic Reviews of Aid to Afghanistan, Mali and South Sudan, 2008–2021’, The Hague: IOB, available at: <https://english.iob-evaluatie.nl/publications/sub-studies/2022/11/10/synthesis-of-systematic-reviews-in-afghanistan-mali-and-south-sudan> (accessed 17 November 2022).

* University of Ottawa, Canada; christoph.zuercher@uottowa.ca.

This paper is published under the UNU-WIDER project *The state and statebuilding in the Global South—international and local interactions*. It is released in connection with the *International Research Conference on the Effectiveness of Development Cooperation*, jointly organized by UNU-WIDER, the Global Partnership for Effective Development Cooperation (GPEDC), and the European Commission (DG INTPA) on 17–18 November 2022 in Brussels, Belgium.

Copyright © Authors 2022

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-293-5

<https://doi.org/10.35188/UNU-WIDER/2022/293-5>

Typescript prepared by Luke Finley.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

The views expressed in this paper are those of the author(s), and do not necessarily reflect the views of the Institute or the United Nations University, nor the programme/project donors.

1 Introduction

Fragile and conflict-affected states cause great harm to their own population. They require and deserve help. Afghanistan, Mali, and South Sudan are among the most fragile and underdeveloped countries in the world. In 2021, they ranked 16th, 9th, and 3rd on the Fragile States Index, and 169th, 184th, and 185th (out of 189) on the Human Development Index.¹ Between 2008 and 2020, the international community provided US\$70.75 billion in official development assistance (ODA) to these three countries. This working paper is concerned with the effectiveness of aid in this context. To what extent can aid be effective in highly fragile situations? Are there differences among aid sectors? What prevents aid effectiveness, and where are there pockets of success?

To provide evidence-based answers to these questions, three systematic reviews of evaluations of aid to Afghanistan, Mali, and South Sudan were conducted. Systematic reviews are exercises in learning. To the best of our knowledge, this is the first time that the instrument of a systematic review has been used to assess aid effectiveness within fragile states. The results should be of interest to donors and policy-makers engaged in fragile states.

The systematic review on aid to Afghanistan was commissioned by the German Federal Ministry for Economic Cooperation and Development (BMZ) and was released in 2020. The systematic reviews on Mali and South Sudan were commissioned by the Policy and Operations Evaluation Department of the Ministry of Foreign Affairs of the Netherlands (IOB) and were released in 2022.

The present report is a synthesis of the main findings from these three systematic reviews. Readers are invited to consult the main texts of the systematic reviews, listed below, for more information, including full bibliographies of all evaluation reports included.

1.1 Afghanistan

The findings on Afghanistan are presented in six stand-alone reports. All reports are accessible at: <https://christophzuercher.weebly.com/papers--reports.html>.

- C. Zürcher (2020), ‘Meta-Review of Evaluations of Development Assistance to Afghanistan, 2008–2018: Chapeau Paper/Summary’, Berlin and Bonn: BMZ.
- C. Zürcher, M.M. Shah, and E. Sylvester (2020), ‘Summary Report of Eleven Bilateral Country-Level Evaluations’, Berlin and Bonn: BMZ.
- C. Zürcher with A. Coon, M. de la Torre Ugarte, P. Labelle, B. Li, R. Masad, H. Popal, R. Saraya, M. Shah, M. Swenson, E. Sylvester, A. Vanderkooy, and M. Wang (2020), ‘Systematic Review of Impact Evaluations of Development Aid in Afghanistan, 2008–2018’, Berlin and Bonn: BMZ.
- H. Popal and C. Zürcher (2020), ‘Summary Report of Selected SIGAR Reports, Afghanistan, 2008–2018’, Berlin and Bonn: (BMZ).
- R. Saraya and C. Zürcher (2020), ‘Summary Report of Evaluation Reports by the Asian Development Bank, 2008–2018’, Berlin and Bonn: BMZ.

¹ The Fragile States Index is accessible at <https://fragilestatesindex.org/data>. Higher ranks mean higher levels of fragility. The HDI is accessible at <https://hdr.undp.org/en/data>. Lower ranks means a lower development score.

- R. Saraya and C. Zürcher (2020), ‘Summary Report of Selected Evaluation Reports by Multilateral Organizations and NGOs, 2008–2018’. Berlin and Bonn: BMZ.

1.2 Mali

The findings on Mali are presented in four stand-alone reports. All reports are accessible at <https://english.iob-evaluatie.nl/publications/sub-studies/2022/11/10/systematic-reviews-mali>.

- C. Zürcher with P. Labelle, L. Borges, K. Hoare, M.A. Javid, K. Kavanagh, S. Sarna, and E. Woolner (2022), ‘Impacts of Development Aid to Mali 2008–2021. Part 1: Summary Paper’, The Hague: IOB.
- C. Zürcher with P. Labelle, L. Borges, K. Hoare, M.A. Javid, K. Kavanagh, S. Sarna, and E. Woolner (2022), ‘Impacts of Development Aid to Mali 2008–2021. Part II: A Synthesis of Country-Level Bi- and Multilateral Evaluation Reports’, The Hague: IOB.
- C. Zürcher with P. Labelle, L. Borges, K. Hoare, M.A. Javid, K. Kavanagh, S. Sarna, and E. Woolner (2022), ‘Impacts of Development Aid to Mali 2008–2021. Part III: A Synthesis of 86 Program and Project Evaluations’, The Hague: IOB.
- C. Zürcher and P. Labelle (2022), ‘Impacts of Development Aid to Mali 2008–2021: Methods’, The Hague: IOB.

1.3 South Sudan

The findings on South Sudan are presented in four stand-alone reports. All reports are accessible at <https://english.iob-evaluatie.nl/publications/sub-studies/2022/11/10/systematic-reviews-south-sudan>.

- C. Zürcher with P. Labelle, L. Borges, K. Hoare, M.A. Javid, K. Kavanagh, S. Sarna, and E. Woolner (2022), ‘Impacts of Development Aid to South Sudan 2008–2021. Part 1: Summary Paper’, The Hague: IOB.
- C. Zürcher with P. Labelle, L. Borges, K. Hoare, M.A. Javid, K. Kavanagh, S. Sarna, and E. Woolner (2022), ‘Impacts of Development Aid to South Sudan 2008–2021. Part II: A Synthesis of Country-Level Bi- and Multilateral Evaluation Reports’, The Hague: IOB.
- C. Zürcher with P. Labelle, L. Borges, K. Hoare, M.A. Javid, K. Kavanagh, S. Sarna, and E. Woolner (2022), ‘Impacts of Development Aid to South Sudan 2008–2021. Part III: A Synthesis of 58 Program and Project Evaluations’, The Hague: IOB.
- C. Zürcher and P. Labelle (2022), ‘Impacts of Development Aid to South Sudan 2008–2021: Methods’, The Hague: IOB.

2 Methods

Searches were conducted in relevant databases, using a specifically developed search string. We also searched the websites of important bilateral and multilateral donor and non-governmental organizations (NGOs) and repositories of evaluations in international development. Finally, we contacted the evaluation departments of multi- and bilateral donors and major NGOs, with a request to provide us with evaluation reports. Search strings and data sources are provided in the appendix.

The searches were conducted between June and September 2021 for Mali and South Sudan, and between October and December 2018 for Afghanistan. Title and abstract screening was conducted with the screening software Covidence. Two researchers needed to independently agree or disagree on whether a study met all required inclusion criteria. Conflicting cases were resolved by the principal investigator.

The systematic reviews are convergent mixed-methods reviews (see Noyes et al. 2022), which include both statistical/experimental and qualitative evaluations in the final synthesis. While well-executed statistical/experimental evaluations are often thought to produce more-robust results than qualitative evaluations, good qualitative evaluations can also provide valuable insights. Furthermore, the majority of available evaluations are qualitative. If we had not taken into account solid qualitative evaluations, a lot of valuable information would have been lost. Also, restricting the study to statistical/experimental studies only could have introduced bias, since these types of studies tend to be much more prevalent in sectors such as health and nutrition, and less so in other sectors. For these reasons, it was decided to also include solid qualitative evaluation reports.

While the present report synthesizes all evidence, in the underlying reports we keep the evidence from statistical evaluations separate from the evidence from qualitative studies. This allows readers to see how the evidence from quantitative and that from qualitative studies ‘speak to’ and mutually reinforce each other.

Studies were included when they met the following criteria:

1. The study was published in English between 2008 and 2021 (2018 for Afghanistan), and published in French between 2008 and 2021 (only for Mali).
2. The study provided an assessment of the outcomes and impacts of projects, programmes, multi-sectoral programmes, and country-level assistance; studies which reported only outputs were not included.
3. Studies met one of two thresholds for methodological quality:
 - a. Either they were what we call ‘rigorous impact evaluations’, with a logically or statistically measured value for the counterfactual; typically, rigorous impact evaluations use one of the following research designs:
 - randomized controlled trials (RCTs);
 - regression discontinuity designs;
 - natural experiments;
 - non-randomized studies with pre-intervention and post-intervention outcomes data in treatment and comparison groups;
 - difference-in-difference designs;
 - interrupted time series;
 - non-randomized studies with control for observable confounding, including various matching design;

- regression designs, including repeated cross-sectional regressions.
- b. Or they were largely qualitative but still meet a certain quality threshold. We call this type of study ‘good-enough evaluation’.² While rigorous evaluations are *likely* to capture the impacts of an intervention, it is *possible* that good-enough evaluations capture the impacts. We used the following four necessary criteria for good-enough evaluations:
 - the study must explicitly intend to assess the outcomes or impacts of one or several specific interventions;
 - the study must contain adequate primary data (typically quantified measures of outcomes, or data from interviews);
 - the study must demonstrate that it is plausible that the data are suited to attributing observed outcomes to the interventions;
 - for interventions with a complex causal chain, the assumed theory of change/causal mechanisms/interventions logic are mentioned.

These criteria were treated as ‘necessary’. Only studies which met all four criteria were included. We recognize that these criteria are, like all existing appraisal tools for non-experimental studies, open to interpretation. In order to minimize the effects of individual bias when assessing the quality of a study, the research team repeatedly applied these criteria to random subsamples of studies. Once all researchers had given their assessments, we discussed the reasoning for our assessments. By repeating this process, we worked toward a shared understanding of how to apply the criteria, resulting in more consistency when the criteria were applied by individual members of the research team. For the actual screening process, two researchers had to independently arrive at the same decision. The principal investigator acted as tie-breaker.

4. Also included were country-level evaluations of bi- or multilateral donors. Such studies were included by default and did not have to meet a threshold for methodological quality. Such evaluations are typically a lengthy and highly condensed synthesis of many programme and project evaluations, and it is in our view not possible to develop a valid quality threshold. Nevertheless, these studies contain a wealth of aggregated information which we did not want to leave out of our analysis.

² This criterion was consistently applied for studies on Mali and South Sudan. For Afghanistan, a slightly less formalized approach was used, insofar as only the principal investigator decided whether a study met the threshold for ‘good enough’.

Data extraction for evaluations was based on a predefined template. Extracted data included population, intervention(s), comparator, outcome, methods, and moderators. In addition, we also extracted data for sustainability, efficiency, and barriers (factors which were said to hinder better implementation). The extracted data served as the basis for the level III reports listed in Section 1 above.

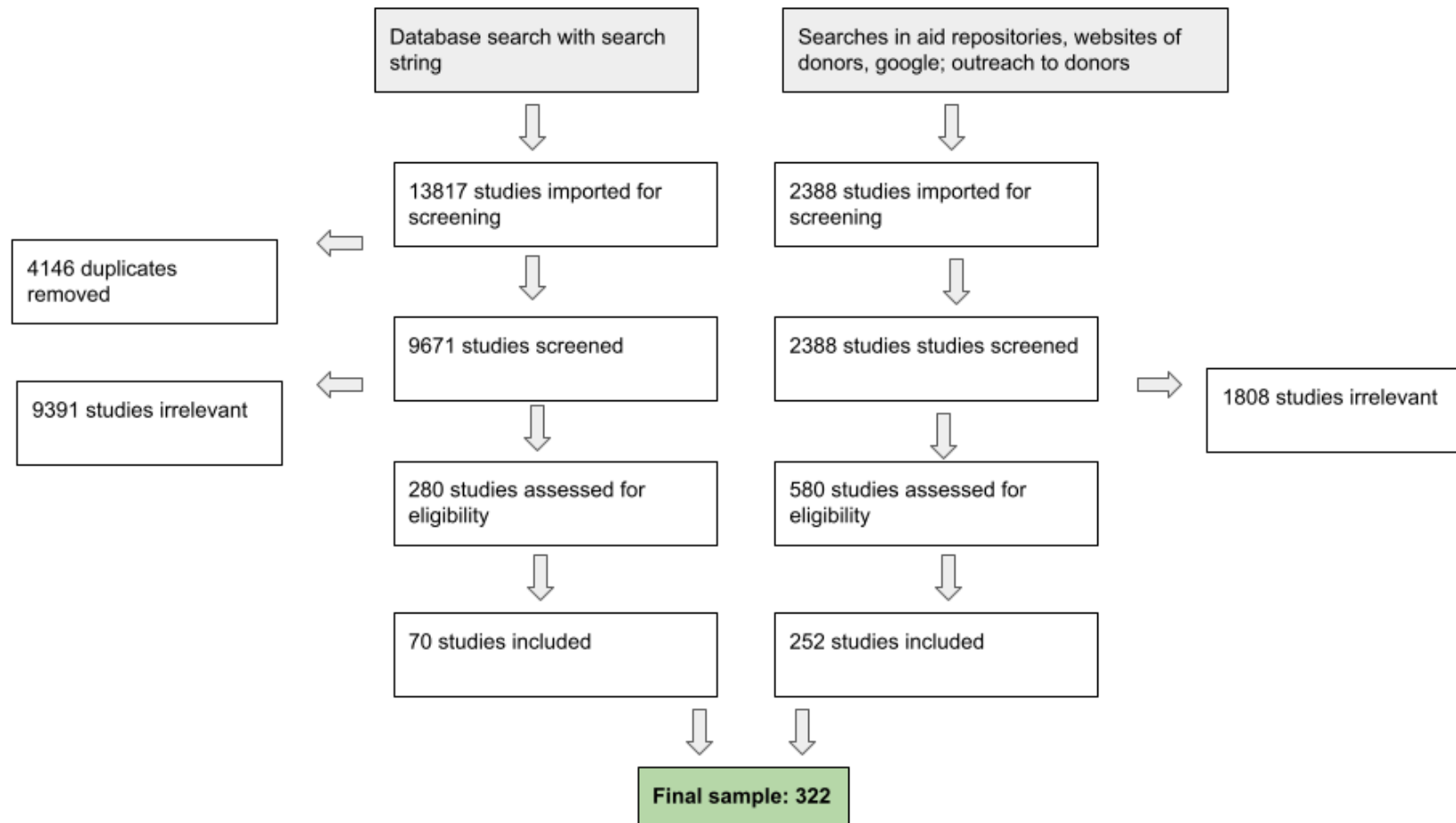
In order to provide an additional in-depth analysis of capacity-building, we used qualitative data analysis software (ATLAS.ti) to identify and extract relevant information.

All studies included were then distributed among ten predefined aid sectors: women's rights; health; rural development and climate change; rule of law; stabilization; education; sustainable economic development; nutrition; humanitarian assistance; and good governance. These sectors were defined based on discussion with the commissioning agency IOB.³ When distributing studies, we typically followed the designation given by the studies themselves. This explains why some types of intervention can be found in more than one sector. For example, WASH (water, sanitation, and hygiene) interventions can be found in humanitarian aid, but also in health; support for village savings associations can be found in rural development, humanitarian aid, and stabilization.

Next we present the PRISMA diagram (Figure 1) and a breakdown of the final sample by country and sector (Table 1).

³ In the case of Afghanistan, a slightly different sectoral classification was used. For the present report, we reclassified studies to ensure consistency.

Figure 1: PRISMA diagram for search on Afghanistan, Mali, and South Sudan



Source: author's illustration based on own work.

Table 1: Evidence base

	Afghanistan (2008–18)		Mali (2008–20)		South Sudan (2008–20)		Total per sector	(Of which 'rigorous')	Note: * not included in
	Included studies, total	(Of which 'rigorous')	Included studies, total	(Of which 'rigorous')	Included studies, total	(Of which 'rigorous')			
Country-level programme evaluations	15	n/a	18	n/a	12	n/a	45	na	
Women's rights	14	2	9	4	4		27	6	
Health and nutrition	21	9	35	27	13	7	69	43	
Rural development, climate	5		26	8	7	2	38	10	
Rule of law	2				1		3	0	
Stabilization	12	11	4		12	1	28	12	
Education	14	3	5	3	3		22	6	
Sustainable economic development	29	2			3	1	32	3	
Humanitarian assistance, refugees, migration*			3	1	7	2	10	3	
Good governance	18		4		8		30	0	
Afghanistan Reconstruction Trust Fund	11						11	0	
Other	7						7	0	
Total (per country)	148	27	104	43	70	13	322	83	
Total studies							322		
(Of which rigorous)							83		

Afghanistan review; ** applies only to Afghanistan.

Source: author's construction based on own work.

3 Country contexts

Afghanistan, Mali, and South Sudan are among the most fragile countries in the world, suffering from the typical symptoms of fragility, such as low growth and few economic opportunities, widespread poverty, low life expectancy, high population growth especially in rural areas, low education levels, internal migration, and so on. In addition, all three countries suffer from violence that exacerbates the developmental problems. Afghanistan has suffered through four decades of war. After the victory of the Taliban in August 2021 violence ebbed, but the country is far from a stable place. Mali is caught in intertwined conflicts, with secessionist Tuareg and Islamist rebellions in the north and inter-ethnic violence in centre of the country. South Sudan plunged into an inter-ethnic civil war only three years after it gained its independence. By 2018 the most intense fighting was over, but violence remains endemic.

As in all highly fragile states, real political power in these three countries resides not with formal political institutions but rather in networks of patronage which transcend the formal political institutions. This creates a political economy where elites have little incentive to strengthen formal state institutions. Instead, they need to nurture the patronage networks which are the basis for their authority and their political and often physical survival. Rent-seeking, widespread institutionalized corruption, and intense competition between rival networks are the typical symptoms of such a political economy of fragile statehood. One important implication of this political economy is that elites are not interested in political reforms which would endanger this mode of governing. Donors therefore often find that their partners show no political will to implement policies which might lead to more accountability, more-robust formal institutions, or democratic procedures.

Despite many commonalities stemming from the fact that all three countries belong to the group of extremely fragile states, there are some differences in how these states became fragile and how the international community reacted.

3.1 Afghanistan

Afghanistan, more than any other country, exemplifies the shattered dreams of externally led state-building. After nearly two decades of massive international engagement, Afghanistan today remains one of the poorest countries in the world, again ruled by the Taliban.

The most recent engagement of the international community in Afghanistan began in 2001. Backed by the US, the so-called Northern Coalition defeated the Taliban and the international community installed Hamid Karzai as interim president of Afghanistan. In 2004, a new constitution was adopted that was meant to pave the way towards a democratic and self-sustainable Afghan state. Once the formal trappings of statehood were in place, the international community began to bankroll Afghanistan's attempts at state-building. During the first decade of their engagement (2002–12) alone, international donors allocated a staggering US\$47.2 billion in ODA. This vast sum reflects both the dire needs of the country and the ambitious vision of the donor community.

The civilian reconstruction went hand in hand with military support. The NATO-led International Security Assistance Force (ISAF) was established under a UN mandate. At its height, the force was more than 130,000 strong, with troops from 51 NATO and partner nations. In addition, there was a sizeable US force which operated outside ISAF, so-called Operation Enduring Freedom.

In the early days of the engagement, optimism was high. The Afghan population welcomed the international engagement and its promise of stability and development. The Taliban appeared to have been defeated, and there were some quick gains enabled by their ousting and the improvements in everyday security. Starting from a very low level, access to basic health services and education improved, roads and other transport infrastructure were rehabilitated, some development aid started to reach rural regions, and the nascent government in Kabul received generous support. In 2014, Ashraf Ghani replaced Hami Karzai as president. The fairness of the election was widely disputed, but the international community nevertheless saw the regime change as a sign of the gradual progress of Afghanistan's transition.

At the same time, there were clear signs that the donors' vision of a stable, democratic, and self-sustainable Afghanistan might be out of reach. The constant stream of external funding had built a burgeoning rentier state with high levels of corruption, entrenched patronage networks, and little capacity or will for reforms. Despite massive investments in government capacities, the Afghan state remained weak.

Just as troublesome was the fact that the Afghan government was not able to earn legitimacy among its population. The new government, despite the massive support from donors, could not build a social contract with the citizens. Instead, the Afghan population increasingly saw the government as ineffective and corrupt. Similarly, the reputation of international donors suffered. Without tangible progress in poverty alleviation and security, the Afghan government and its foreign backers were losing the battle for the hearts and minds of the Afghan people.

Donors continued their financial support. While ODA peaked in 2011 at nearly US\$6 billion, it was still at US\$3.68 billion on average for the years 2012–18.⁴ The vision of a self-sustainable and democratic Afghanistan was reiterated at a series of large pledging conferences.

Regarding the overall aid portfolio, there were few adjustments. Notably, the share of aid for 'government and civil society' averaged an annual 49 per cent of total ODA between 2008 and 2020, and was still at 49 per cent in 2019. At the same time, funding for humanitarian aid and rural development remained low, at 6.1 per cent and 10.2 per cent of total ODA respectively for the period 2008–20. These numbers suggest that donors continued to treat Afghanistan as a 'normal' developing country (unlike South Sudan, where spending for 'government and society' shrank to 6 per cent; see Section 3.3), despite the lack of progress in fields such as state capacity and good governance.

By 2018, it became evident that the US wanted a fast exit option. Talks between the US and the Taliban led to the peace agreement in Doha of 29 February 2020. This agreement was negotiated between the US and the Taliban only, leaving out the government of Afghanistan as well as the wider international community. Following the signing of the agreement, the US announced the withdrawal of the remaining US troops by September 2021. The Taliban, seizing the opportunity, increased its military pressure on the Afghan security forces. After capturing many major cities, often without a fight, the Taliban gained control of Kabul in August 2021.

⁴ ODA numbers here and in the following sections are from the Creditor Reporter System (CRS), Development Assistance Committee (DAC), Organisation for Economic Co-operation and Development (OECD): <https://stats.oecd.org/Index.aspx?DataSetCode=crs1> (accessed 2 November 2022).

3.2 Mali

The landlocked country of Mali faces a multitude of development challenges, including high population growth rates and an undiversified economy dependent on subsistence agriculture. Mali also experiences recurring natural hazards such as droughts, irregular rains, and floods. In addition, there is endemic violence and insecurity in the north and the central regions. Northern Mali is also a drug trafficking hub for cocaine from Latin America to Europe. The competition over lucrative smuggling routes fuels political instability and further complicates the search for peace.

Mali has been on a sharp downward trajectory since 2012. Prior to 2012, Mali was a 'donor darling', with foreign governments keen to provide support for the democratic transitions that began in the 1990s. A toxic brew of drought, insurgencies, and a military coup ended the positive trajectory as well as the optimistic outlook of donors. By 2012, Tuareg and Islamist secessionist rebellions in the north had destabilized the country. The Tuareg rebellion crystallized around heavily armed Tuaregs who had fought in the Libyan civil war and returned to Mali after NATO's intervention in Libya.

In reaction to the ineffective response of the government to the new threats, the military overthrew the democratically elected President Touré. Unable to restore its authority in the rebel-held areas, the Malian government requested international support. French forces intervened to stabilize the northern regions, followed by the UN Multidimensional Integrated Stabilization Mission in Mali (MINUSMA), established in 2013 with a broad mandate to support security and stabilization.

Unrest gripped not only the north. In 2013, violence also erupted in central Mali. The conflict was fought mainly between agricultural and pastoral communities, fuelled by competition for land and water and exacerbated by climate change.

In June 2015, a peace agreement was brokered between the government and an alliance of various rebel groups. However, that peace deal did not include the Islamist insurgency, and insecurity in the north remained high. Important elements of the peace process were the promise of decentralization, giving more authority to the regions, integrating former rebels into a national army, boosting the economy in the north, and introducing national reconciliation through dialogue and attention to justice.

The implementation of this agreement proved to be extremely difficult. Constitutional reforms which were promised under the peace agreement remained on hold, and the peace process lacked popular support.

In August 2020, and then again in May 2021, two more military coups d'état took place. The international community condemned these, and tension between the government of Mali and Western countries rose to new heights.

In hindsight, the escalation of the conflict in northern Mali and the breakdown of the democratic order in 2012 was a watershed moment for the country and for its relations with donors. During the decade prior to 2012, Mali had been treated by the donor community as a positive example of democratic governance. From 2002 to 2011, annual ODA steadily increased from US\$610.35 million to \$1.14 billion. A substantial part of the aid was channelled through the government in the form of budget support. Budget support was credited with an overall positive effect on overall public finances in Mali, and specifically with positive impacts in education and health, but also with having significantly contributed to achieving goals in the national poverty and growth strategy, macroeconomic management, and public investment in the social sectors.

Donor engagement in Mali changed in reaction to the rupture of 2012. While ODA still consistently grew, budget support was drastically reduced, less aid went to the public sector, and a larger share was channelled through multilateral organizations and NGOs. Humanitarian assistance, with food security as the most important spending line, dramatically increased.

It is also noteworthy that in reaction to the breakdown of the democratic order, donors invested more aid in sectors such as government and elections, probably because they assumed that increasing aid to these sectors would help the return to democracy. Despite an overall investment of US\$5.245 billion in ODA since 2013, Mali today is less secure, less democratic, and more fragile than it was before the breakdown of the democratic order in 2012.

3.3 South Sudan

South Sudan came into existence in 2011. It was born as an extremely fragile state, and it has remained so ever since. Between 2011 and 2020, South Sudan received US\$11.43 billion in ODA.

South Sudan had been fighting for autonomy from the north since 1983. The violent struggle for independence from the north went hand in hand with violence between political factions and ethnic groups within the south. In 2005, South Sudan signed a peace agreement with the government in Khartoum. This so-called Comprehensive Peace Agreement (CPA) called for a power-sharing agreement between the north and the south and for the formation of a semi-autonomous government in the south. It was also agreed that there would be a transition period of six years, after which the final status of South Sudan would be determined by a popular referendum. Donors rushed to provide support for the CPA as well as for the new institutions of the Government of South Sudan. It was assumed that such support would help to make the fragile peace permanent, and to prevent South Sudan from pushing for independence. However, in January 2011, 99 per cent of the South Sudanese population voted for independence in a referendum, and South Sudan became an independent country. Donors continued to provide financial and political support for the newborn country.

Despite high hopes and big efforts, the peace did not hold. In 2013, internal conflicts within the ruling Sudan People's Liberation Movement led to an acute political crisis that soon turned into a full-fledged civil war fought along ethnic lines. After five years of devastating violence, the two parties agreed on a ceasefire and then a peace agreement that called for a power-sharing structure. Between 2013 and 2018, the war had displaced four million South Sudanese (out of a population of 11 million). An estimated 400,000 people were killed, the economy collapsed, and food shortage was widespread. South Sudan had become a major humanitarian disaster.

Between independence in 2011 and the resurgence of war in 2013, donors provided massive levels of aid for state-building. ODA almost tripled between 2011 and 2013, from US\$351 million to \$1,083 million. Donors, however, grew increasingly frustrated with the lack of political will on the side of the South Sudanese government to engage in the state-building project.

The outbreak of the civil war in 2013 marked the beginning of a new phase in donor engagement in South Sudan. Donors now predominantly engaged in emergency assistance, food security, and local-level peacebuilding, hoping to prevent a humanitarian catastrophe and to contribute to stability and eventually peace. Support for government institutions was scaled down. ODA was still climbing, from US\$1,083 million in 2013 to \$1,751 million in 2017, the year when ODA peaked. While aid for state-building shrank, humanitarian aid dramatically increased. In 2011, humanitarian aid had amounted to US\$110 million. That number rose to \$ 1154 million in 2017. Relative to overall ODA, the share of humanitarian aid rose from 31 per cent in 2011 to 66 per cent in 2017.

4 Aid for core state functions: stability, capacity-building, and good governance

How effective can aid be in highly fragile and conflict-affected countries, and does effectiveness vary among aid sectors? The findings from the three systematic reviews provide comprehensive and evidence-based answers to these questions.

We first discuss the effectiveness of aid in addressing the core deficiencies of fragile and conflict-affected states, namely a lack of security, state capacities, and institutions and procedures which would enable ‘good governance’.

4.1 Stabilization

Stabilization is a key objective of donors in fragile and conflict-affected states, and aid is supposed to play an important role. Unfortunately, stabilization is a vague concept, and donors subsume a very wide variety of activities under the term, as the examples from our sample demonstrate.

Interventions which are categorized by donors as ‘stabilization’ can be separated into five groups. A first group of stabilization projects and programmes aims to quickly restore basic services to the population in the hope that this will encourage them to work with the government and to loosen ties with insurgents. Such stabilization initiatives rely on what can be called a peace dividend approach: by re-establishing services such as education, water, sanitation, transportation, and health services, which had been interrupted during war, donors hope to win over the population. This in turn will lead to more good will for the government and its foreign backers, which will then eventually help to build a more legitimate and capable government. In a nutshell, peace dividends will help to win the hearts and minds of the people.

A second group of stabilization projects aims to provide economic opportunities, for example by providing training, investing in job creation, or providing access to credits. Such economic opportunities are also supposed to provide a peace dividend, but in addition they are assumed to make it more expensive for insurgents to recruit and pay fighters. When labour is cheap and economic opportunities scarce, insurgents find it easy to recruit fighters. By providing sources of legitimate income, the costs for insurgents rise. All of this should in theory lead to more stability.

A third type of intervention aims to directly strengthen the mediation and conflict resolution capacity of communities and political actors. Donors have supported conflict management and peacebuilding processes at the local level, offered training in dispute resolution, facilitated peace meetings, strengthened local-level justice support, and created and supported so-called peace clubs, where local communities can meet and learn mediation skills. Donors have also created spaces for discussion between opposing parties and provided support for political communication via mass media in support of peace, democratic participation, and reconciliation.

A fourth type of intervention supports political institutions and processes which are seen as contributing to stability. For example, donors have provided support for rebel organizations to transform into political parties, for the Demobilization, Disarmament, and Reintegration (DDR) process by building cantonment for former rebel forces, for the Truth, Justice, and Reconciliation Commission (CVJR) in Mali, and to legislatures in all three countries. In Afghanistan and Mali, donors have also pushed for political and fiscal decentralization in order to defuse regional competition.

Finally, support for civil society actors was also seen as a possible contribution to stabilization. Since unresponsive and often corrupt governments were seen as one source of instability, it was

assumed that more-capable civil society actors would be better placed to hold the government more accountable, thereby helping to rebuild a social contract.

The findings from the three systematic reviews strongly suggest that aid for stabilization in fragile states is not effective.

Before we document this finding, three clarifications are needed.

First, donors often (re)label ‘regular’ developmental interventions as stabilization interventions. In fact, stabilization interventions are often very similar to ordinary developmental interventions in sectors such as emergency aid, livelihoods, rural development, access to justice, etc.⁵ In the context of stabilization, it is then assumed that developmental outcomes might also contribute to more stability, often without the complete causal chain being specified (or measured).

Second, not all evaluations of stabilization interventions use a valid measure for the success of stabilization. Some rigorous studies measure success as a reduction of violence, more legitimacy for the government, or a renewed social contract between society and the government.⁶ In our view, this is an appropriate way of measuring stabilization. However, other evaluations (mainly qualitative studies) measure success more upstream. Such studies may, for example, discuss whether stabilization interventions led to more jobs, whether peace clubs remained operational, or whether beneficiaries thought that their mediation capabilities increased. Achieving such immediate outcomes is then often portrayed as a success of the stabilization intervention. This is a false claim, however. Such immediate outcomes may be an important first step towards stabilization, but without measuring the subsequent steps (for example, a reduction of violence) these studies cannot prove that the interventions did indeed lead to more stabilization. It would be advisable for evaluations of stability interventions to always use a valid measure of stabilization outcomes, such as a reduction of violence or the improved legitimacy of the government.⁷

Third, these immediate stabilization outcomes (such as employment opportunities, peace clubs, or mediation training) are typically implemented at the local, communal level. The hope is that such local-level stabilization will trickle up to create more stability at a higher level. The reviewed evaluations show that such immediate local effects were rarely achieved, and hence there could not be a trickling up.

We now return to summarizing the findings. Afghanistan provides the largest evidence base, with 12 project-level evaluations, extensive reports by the Special Inspector General for the Reconstruction of Afghanistan, and assessments of the stabilization programmes of the UK, US, Denmark, and Canada. In sum, there is no evidence that these initiatives had a positive effect on stability.

The US ran by far the largest stabilization programme in Afghanistan. The available evaluations suggest that this programme not only did not dampen violence but even exacerbated intergroup

⁵ Note that we included evaluations under stabilization where there was an explicit objective of reducing violence, or when the study was placed in the stabilization field by the donors themselves.

⁶ We think that a reduction of violence is indeed an appropriate measure for stabilization initiatives, but we also note that it is often not easy to obtain data on this, especially at the subnational level. Studies which employ such violence data often use declassified military data (available mainly for Afghanistan and Iraq) or collect their own data.

⁷ Studies which measure only immediate outcomes can still contribute to our evidence base. If the immediate outcome (the first element in the causal chain) was not achieved, then we know that the stability outcome was also not achieved. This is indeed the case for most studies in the stabilization sample.

tensions and attracted violence. Evaluations at the project level came to similar conclusions. For example, the Commander's Emergency Response Programme (CERP) neither reduced violence nor won the hearts and minds of the local population. CERP was the premier stabilization programme of the US army in Afghanistan. It provided small- to medium-scale projects to communities in the field of humanitarian relief and reconstruction. Other evaluations investigate the effects of various community-driven development programmes and projects aimed at increasing youth employment by offering technical vocational education. None of these interventions reduced violence or the propensity to join the insurgency or increased the legitimacy of the government.

The evidence from Mali and South Sudan corroborates the findings from Afghanistan. In South Sudan, interventions in stabilization were not effective and could not prevent a resurgence in violence. After independence and up to 2014, donors operated under the assumption that a major cause of conflict was a lack of development, and hence they focused on delivering peace dividends, under the assumption that peace and development would reinforce each other. They also supported DDR measures and provided incentives for military organizations to transform themselves into political parties. These endeavours failed because of the South Sudanese government's lack of political will.

After the outbreak of the civil war, donors helped to create spaces for dialogue in society and supported the various peace negotiations between the warring parties, but without discernible effect in terms of violence reduction. Donors also supported conflict management and peacebuilding processes at the local level, including dispute resolution, peace meetings, livelihood programmes, and local-level justice support. There is no evidence that such interventions achieved the desired immediate outcomes, and the reviewed evaluation reports doubt that such local-level interventions had any long-term effect on stability. Furthermore, most interventions were directed at the community level and hence not designed to support conflict management processes at inter-communal or inter-ethnic level. As such, they would have had little effect on the major drivers of conflict. These findings strongly suggest that small-scale local stabilization efforts do not have a trickle-up effect on stabilization.

Findings from Mali are similar. We have evaluations on a variety of interventions in the field of stabilization. At the national level, donors provided support for the Truth, Justice, and Reconciliation Commission (CVJR), for political communication via mass media in support of peace, for political and fiscal decentralization, and for cantonment in the framework of the DDR processes. At the local level, donors supported income-generating activities for the private sector, capacity-building in conflict resolution, and restoring access to basic services. The available evidence does not suggest that interventions in any of these fields contributed to greater stability. Many of the evaluations included could not demonstrate that immediate outcomes were achieved, and none provided evidence that the interventions contributed to the overarching objective of stability in terms of violence reduction or improved legitimacy for the government.

In sum, there is no evidence that aid has contributed to greater stability in Afghanistan, Mali, or South Sudan. Most evaluation reports do not discuss why this might be the case. However, a reading of the recent academic literature on the topic suggests that aid only has a stabilizing effect under rare circumstances, which are unlikely to be present in regions that are still conflict-affected. In order to have a stabilizing effect, aid must be implemented in reasonably secure regions under government control where insurgents lack the capacity to sabotage, loot, or tax aid projects. Furthermore, aid should be implemented in participatory ways, preferably through accepted local authorities; and it should be transparent and not benefit local power-brokers through corruption or nepotism. Absent such rare conditions, aid is unlikely to improve stability but has the potential

to exacerbate corruption and intergroup conflicts, as documented in Afghanistan (Sexton 2016; Zürcher, 2020).

The existing evidence strongly suggests that aid is not a good tool for making very unstable regions less unstable. Aid alone will not chase away insurgents. However, it may still have a stabilizing effect on regions where some modest levels of security and governance are achieved and where insurgents cannot freely move. One implication is that stabilization efforts should not focus on the highest-risk areas but should rather seek to bolster and make more durable existing pockets of stability and build out from there.

4.2 Government capacity-building

A lack of government capacity is at the core of fragile states, and many development actors have sought to build capacity across all levels of government and, in addition, to enhance the capacity of the civil society.

For Afghanistan, the evidence clearly suggests that state capacity-building measures were mostly not successful. In the few instances where progress was made, it remained confined to small silos, not translating into more overall state capacity, or it was mainly borrowed from the so-called ‘second civil service’, consisting of well-paid Afghan returnees or international consultants. There is no clear case of a successful capacity-building programme in the Afghanistan sample, but there are quite a few examples of rather ineffective capacity-building programmes. For example, donors agreed that the massive Afghan Reconstruction Trust Fund (ARTF) had no impact on better governance, nor did it contribute to better capacity in the Afghan government. Likewise, three programmes for decentralization and capacity-building for the subnational administration were not successful at building more capacity for the local administration or providing incentives for a meaningful decentralization process. Also, two reviewed capacity-building programmes for civil servants proved mostly ineffective.

The reasons for these disappointing results vary but are mostly linked to the difficult context: there was little demand from the government for such programmes, the permanent competition for power among various networks hampered co-operation within and among institutions, and the frequent changes of key personnel made institutional learning difficult. Most importantly, capacity-building was not effective when there was no political will to have capacity built, which was often the case in politicized fields such as, for example, decentralization, which the central government opposed.

In Mali, the results are similarly disappointing. There is no evidence that capacity-building at the level of central governance was effective. For example, in 2015, three years after the breakdown of the democratic order, the World Bank noted that objectives in the fields of governance reforms, management of public resources, and fiscal decentralization had not been achieved. The same report noted that the World Bank had overestimated the capacity and political will of the government of Mali for institutional reform.

While capacity-building for the central government was usually not effective, there were some pockets of success, mainly in the health sectors and in very technical, apolitical fields at the subnational level, for example in education, rural development, and health. Several evaluations noted that capacities in the health sectors had improved, resulting in better vaccination coverage, better maternal and child health, better training for midwives and nurses to perform necessary treatments and to address the shortage of doctors in the community health sector, and enhanced capacity of the Ministry of Health personnel.

Regarding South Sudan, there is broad consensus in the evaluation reports that capacity-building measures for the government were rarely effective. Donors overestimated both existing state capacity and the government's political will for reforms. As a result, programming for capacity-building was overambitious, unsustainable, and ineffective. The hope that the government would become a partner in service delivery was never realized.

Least effective were capacity-building measures in 'politicised' fields which offer opportunities for lucrative corruption, and which are therefore vital for maintaining the rentier and patronage state. For example, Norway's support for the natural resource sector did not lead to more accountability and transparency, since there was no political will for it. Support for the legislature and the judicial sector appear to have had little effect either, since there is no sign of improvement in this sector. While it was possible to provide skills to individuals, this did not translate into more-effective and more-accepted political institutions.

To compensate for the lack of government capacities, donors created parallel systems for service delivery, which helped to provide outputs, but these results were not sustainable and did not address the underlying issues of the lack of state capacity.

Like in Mali, some positive results were found in the health sector, where donor support did help to increase individual and institutional capacity, leading to better treatment of malaria, diarrhoea, and pneumonia and to better healthcare provision for mothers and children.

In sum, the available evidence clearly suggests that capacity-building for governments in very fragile states is ineffective, especially when it is applied to fields which can affect political processes in general and the patronage and rent-seeking economy in particular. The obstacles are smaller in strictly technical fields which are not overly politicized, where there is less opportunity for rent-seeking, and when the subnational rather than the national level is targeted.

4.3 Good governance

Good governance is a very broad sector, including public sector and regulatory policy reform, democracy promotion, election support, anti-corruption programmes, and rule of law, among other things. It is, together with stabilization and capacity-building for the government, at the core of what donors are trying to achieve in fragile states. Promoting good governance is seen not only as a normatively desirable goal but also as a prerequisite for stable, self-sustainable, and violence-free statehood. Donors rarely give up on their visions of bringing good governance to fragile and conflict-affected states. One of the few exceptions is South Sudan after 2014, when donors admitted that the South Sudanese government was unwilling to work towards better governance. As a result, donors largely gave up on working with the government for better governance and instead increased humanitarian and emergency aid which bypassed government channels. By contrast, in Afghanistan donors continued to sponsor good governance projects, despite mounting evidence of widespread systematic corruption and flawed democratic institutions.

In Afghanistan, there is no evidence that interventions aiming to improve governance were effective. We mentioned above that capacity-building measures for the central and subnational governments were not successful. Also not effective were initiatives aiming to reform public administration and to create better regulatory frameworks for private sector development and for the agriculture sector.

Evaluations in rule of law suggest that interventions were not successful because they were overly ambitious, were not based on the political-economic realities on the ground, and were ideologically framed by an unrealistic theory of change.

Finally, regarding democracy promotion, the available evidence suggests that development assistance could provide the technical capacities needed for conducting elections; however, projects aimed at democratic awareness or democratic participation had little effect.

In Mali, the evaluations included cover a variety of interventions in various fields, such as management of public resources, fiscal decentralization, political decentralization, rehabilitation of governmental infrastructure, capacity-building for government officials and civil society, strengthening the relationship between citizens and public authorities, and establishing spaces for public debate. There is no example of an effective intervention in any of these fields in our sample.

Prior to 2012, many donors had noted some positive trends in the governance sector in Mali. However, after the breakdown of the democratic order in 2012, the available evaluations increasingly suggest that the governance sector was in a bad state, with ineffective institutions, weak capacities, lack of accountability, endemic corruption, and entrenched patronage. According to a World Bank evaluation of 2015, most objectives in the fields of governance reforms, management of public resources, and fiscal decentralization were not achieved. The same report noted that the World Bank had overestimated the capacity and the political will of the government of Mali for institutional reform.

Also not effective were decentralization interventions in Mali. After the 2015 peace agreement, administrative decentralization was seen as one way to promote stability and better governance. As a result, many donors supported the process of decentralization. However, little was achieved, mainly because there was no political buy-in from the central government for decentralization, which was seen as strengthening the political position of the north.

One field where donors saw some success was in election support. Support from the UN Development Programme (UNDP) for elections in 2013 and 2016 was seen as mainly effective in the sense that this technical help enabled the government to conduct elections. But we also note that conducting elections did not result in improved democracy, since Mali's democracy has been in steady decline since 2008.

Not surprisingly, good governance initiatives in South Sudan were also not effective. After South Sudan gained independence, donors invested in good governance projects by supporting institutions such as the National Legislative Assembly, the Anti-Corruption Commission, the National Audit Chambers (NAC), and the National Elections Commission (NEC). Also supported were reforms in public financial management and budget strengthening. Still other interventions aimed at support for the rule of law, for community policing, for the free media, and for civil society organizations.

However, donors generally overestimated South Sudan's state capacity and its political will to improve governance, leading to overambitious and ineffective programmes. By 2014, a consensus among donors had emerged that governance projects had not been effective, that ownership by the government for such projects remained low, that little capacity had been built, and that the new government of South Sudan lacked not only the capacity but also the political will to become a committed partner for state-building.

In sum, interventions in the field of good governance were not effective in Afghanistan, Mali and South Sudan. Factors which hampered such programmes were entrenched patronage-based practices within the government, a lack of buy-in from the government, donor-driven top-down project design with little regard for the core institutional requirements and the demands of the partner institutions, and the lack of political will of the government.

5 Aid for basic services: health and education

5.1 Health

For Afghanistan, the available studies point to a tangible increase in access to basic healthcare and a massive improvement in many health indicators, especially child and maternal mortality. The reports reviewed suggest that successful interventions took place in midwifery training, antenatal care visits, deliveries attended by health workers, conditional cash transfers for women and community health workers, and improved family planning. Access to healthcare was improved with the implementation of a Basic Package of Health Services (BPHS), providing primary care and enhanced access to in-patient care and eliminating national user fees. Due to limited government capacity to provide health services, all publicly funded health services provided are delivered either through contracting out, in the form of service delivery by NGOs, or through service provision and programming operated directly by the Ministry of Public Health.

For Mali, the evaluation reports agree that there were some positive outcomes, especially regarding maternal, newborn, and child health, access to health for women, and mobilizing resources for the fight against HIV/AIDS. Budget support, which was widely used until 2012, is credited with the improved capacity of the healthcare system and better access to healthcare. After 2012, budget support was reduced. Instead, donors increasingly used multi-donor trust funds, which are credited with strengthening the health system and reducing maternal and infant morbidity and mortality. Effective interventions were also reported in the fields of malaria management, WASH, and child mortality. One evaluation mentioned that telemedicine training for health workers and teleconsulting were used effectively to improve the health of mothers and children.

For South Sudan, the available evaluations also report some positive outcomes, again especially in maternal, newborn, and child health. Notably, the massive Health Pooled Fund (HPF) was credited with significant results in improving health services. One evaluation suggests that the HPF was a success story and that this aid modality should be implemented in other fragile contexts. The evaluations also note that more improvements in health services were hindered by a lack of skilled human resources, difficult logistics, and endemic violence.

Also noteworthy is the finding that many women refused to access specific services because the health service providers were men. In many communities, unfavourable gender roles and societal norms continued to be a barrier to access, particularly in areas with poor literacy and high poverty rates. The 'rigid roles' for men and women were highlighted as one of the biggest barriers to women getting healthcare in South Sudan.

In sum, interventions in the health sector were mostly effective in all three countries. That said, the sustainability of many of the effective health practices relies on continued funding from donors. Results are therefore often not sustainable without external support.

5.2 Education

For Afghanistan, the available evaluation reports agree that substantial progress has been made regarding better access to primary education for boys and girls. Much of the progress was enabled by the collapse of the Taliban regime and the improving security situation after 2003. In addition, education was an early priority of the Afghan government, and many donors also made it a priority in their aid portfolios. However, despite massive investments, a large demand for infrastructure remained, the quality of education remained problematic, and evaluation reports warned that many gains might not be sustainable given the enduring insecurity and the lacking financial and

administrative capacity of the Afghan government. Furthermore, by 2015, donors also warned that impressive enrolment figures and number of schools built, in fact, included many ‘ghost’ pupils and schools that did not actually exist.

In Mali, some gains were made in education, leading to improved enrolment and retention. Before the collapse of the democratic order in 2012, donors allocated a substantial part of their aid as direct budget support. Budget support was credited with improvements in the education sector, especially with increasing enrolment rates in basic education. Even after the end of widespread budget support, donors continued to support the educational sector, and evaluation reports suggest that some results were achieved, especially in enrolment and retention. The widely used school feeding programmes had a positive impact on both nutrition and retention. Also noteworthy is the fact that donors recorded the strong commitment of the Government of Mali to the education sector even after 2012. Similarly, there was a high level of buy-in for the educational sector from local administrations as well from communities. However, while access to education increased, the quality of learning did not improve.

In South Sudan, the impact of education projects was relatively low. Evaluations suggest that donors did contribute to some improvements in infrastructure for education and retention rates, but not necessarily in the quality of learning. Support for education, especially basic education, became slightly more prominent after 2014 as donors moved away from state-building and shifted some of their aid to basic services. A very thorough and comprehensive evaluation of the Global Partnership for Education (GPE) found that GPE programmes in South Sudan achieved their target for improving school supervision and for increasing total enrolment. However, according to this report, the interventions did not improve learning outcomes or school leadership or reduce school dropouts. Many evaluations also note that improvements in the education sector were hindered by a high level of insecurity in the country, making it very difficult to provide basic services outside of the relatively safe urban region. Cash transfers for girls as well as school feeding appear to have been effective in supporting girls’ attendance and retention.

In sum, interventions in education appear to have been somewhat effective even under very adverse circumstances. In all three countries, evaluations show that it is possible to improve infrastructure and to increase enrolment and retention: this was most effective in Mali and least so in South Sudan. School feeding and targeted cash transfers can be effective add-ons. It is noteworthy that governments in Mali and Afghanistan both saw education as a priority, and that buy-in for this sector was relatively high. However, in all cases, improving the quality of education has been much more challenging than increasing enrolment.

6 Aid for livelihoods and economic opportunities

The strengthening of livelihoods and the creation of economic opportunities is an important objective of donors in fragile contexts, and donors use a wide variety of interventions in this field. At the macro level, donors may support projects in the fields of macroeconomic policies and financial management, support for the private sector, or investments in sectors such as telecommunications, transportation, and energy. On the more local level, there are interventions in rural development and resilience, including small infrastructure for irrigation, roads, energy, access to credits and savings, and a wide array of community-driven development approaches.

Interestingly, the bulk of the available evidence from our three countries refers to more-localized interventions in rural development. Information on macro-level economic interventions is rather scarce for all three countries.

In Afghanistan, programmes supporting economic development, macroeconomic policies, and financial management capacities achieved some progress in the early stages of reconstruction after 2004. For example, there was initial growth in telecommunications, transport, and construction, but results were not sustainable, nor was it realistic to expect sustainable economic growth, given the insecure environment and the shrinking aid flows after 2013. Interventions aimed at promoting the private sector were rarely effective. Evaluations cited weak institutional infrastructures and procedures, widespread corruption within the Afghan government, political instability, and insecurity as the main reasons. Interventions aimed at regulatory policies for fiscal management and for public administration reform were also rarely effective because institutional absorptive capacity was limited. In general, donors agree that interventions aimed at reducing poverty and creating jobs and incomes for people throughout the country were overall not effective.

Regarding rural development, the evidence suggests that the interventions, often implemented through newly created community-level organizations, helped to build a large amount of small infrastructure such as roads, irrigation, and access to energy. This contributed to improved livelihoods and strengthened coping mechanisms, but it has not led to sustainable economic growth that translated into jobs or income opportunities.

In sum, interventions in sustainable economic development in Afghanistan, despite some progress, have not been able to reduce poverty rates or promote sustainable economic growth.

For Mali and South Sudan, we have evidence only on interventions in rural development.

In Mali, rural development was an important aspect of donors' aid portfolios. Before the onset of violence, donors were optimistic about the long-term opportunities for Mali's agricultural sectors. That changed after the collapse of the democratic order, and interventions in the rural sector increasingly focused on food security, coping mechanisms, and resilience. Results appeared to have been satisfactory, and interventions such as rural credits, cash and in-kind transfers, saving associations, introduction of new crops, and irrigation projects were all effective to some extent. It is also noteworthy that the co-operation between donors and the government in the sector of rural development was satisfactory, and that capacity-building and training for government officials at the subnational level in technical fields relating to rural development was effective. This contrasts with most other fields, where co-operation and capacity-building were rarely effective. The available evaluations also agree that interventions aimed at increasing productivity beyond subsistence farming in rural development were rarely successful. For example, support for agricultural value chains or for small agro-businesses was not effective.

In South Sudan, projects on rural development and climate, including those related to food security, livelihoods, water and sanitation, and small infrastructure, proved to be reasonably effective at improving coping strategies and resilience. However, there is no evidence in the available evaluations that interventions successfully improved productivity beyond subsistence level for the beneficiaries.

In sum, macro-level interventions (for which we have evidence only from Afghanistan) were not effective in promoting sustainable economic growth. At the local and rural level, interventions in all three countries were reasonably effective at increasing the resilience and improving the coping strategies of beneficiaries, but not enough to improve their productivity beyond subsistence.

7 Aid for women's empowerment and gender equality

In this section, we summarize the main findings from evaluations which were categorized by the donors themselves as having a main focus on gender relations and women's empowerment. The scope of these interventions differs. At one end of the spectrum are interventions designed to improve the economic situation and access to services for women (and often men). At the other end are interventions which aim at transforming gender relations, leading to a more gender-equal society. The evidence suggests that the former type of intervention was often effective in achieving the stated objectives, whereas there is no evidence that the latter type of intervention had a transformative effect.

The gender dimension was especially prominent in development co-operation in Afghanistan. In the political narrative of many donors, helping Afghan women and girls was often portrayed as both an important objective and a legitimizing factor for international engagement. The end of Taliban rule in 2002 and the rehabilitation of infrastructure in health and education meant that access to health and education markedly improved for men and women. This does not automatically mean, however, that interventions targeting gender equality were effective.

The effectiveness of gender programming was low in Afghanistan. For example, despite sustained support, donors reported that the capacity of the Ministry of Women's Affairs (MoWA) remained weak. Donors also noted that both the capacity and the political will of the Afghan government and political elites for gender equality programming remained very limited, as prevailing cultural norms made progress difficult.

There were, however, pockets of modest success. Examples include rural literacy for women, increased access to health and education, and better livelihoods in women-specific activities within agriculture, such as mushroom farming and kitchen gardening. In sum, small, modest projects embedded in traditional structures helped to increase access to health and education and modestly improved livelihoods for women. By contrast, larger, more ambitious projects aimed directly at changing gender norms and relations had no discernible impact.

Results are similar for Mali and South Sudan. In Mali, the political and cultural context made it difficult to make tangible progress on women's rights and gender equality. Some positive results were achieved regarding legislative changes. Donors pushed for new laws on gender quotas in parliament, better protection of victims of gender-based violence (GBV), and better representation of women in formal and informal institutions dealing with issues of peace. It remains unclear, however, whether such legal changes contributed to improved gender equality through actual implementation. The support of the government and its willingness to implement new laws and norms appear to be rather low.

At the local level, projects aimed at tangible benefits for women were often effective. For example, village savings and loans associations (VSLAs) helped improve women's economic status. Interventions aimed at better educational outcomes for girls led to concrete results. Finally, reports noted that removing health user fees increased access to healthcare for women.

In South Sudan, interventions that aimed to improve women's rights and gender equality were not effective. While gender equality policies were revised or newly introduced, they were rarely implemented. Also, women's participation in the many unsuccessful peace negotiations rarely yielded results. Project evaluations suggest that interventions aimed at reducing GBV were not effective.

Some small pockets of success can be found in projects in the education sectors. Targeted interventions helped increase female enrolment and retention. Also worth mentioning is the fact that some positive results regarding women's economic empowerment were achieved in aid sectors such as rural development or emergency aid. While community-level projects aimed at enhancing women's social or political empowerment produced few tangible outcomes, projects with an economic component, specifically targeting women's economic situation, fared better, since they had broader acceptance and reach.

In sum, interventions that aimed to improve gender equality were not very effective. Donors were sometimes successful in pushing for institutional reforms, leading, for example, to gender quotas in the parliaments of Afghanistan and Mali, the formation of a Ministry of Women's Affairs in Afghanistan, the formal inclusion of more women in political processes, and new legislation for the protection of women's rights. However, such institutional changes often did not translate into action, as laws were not implemented or ministries were not capacitated. As a result, there was little real change in gender relations.

More effective were smaller, local interventions providing tangible benefits for women, for example in fields such as literacy for rural women, increased access to health and education, support for livelihood projects in women-specific activities within agriculture, better access to microcredits and saving associations for women, or better retention rates for girls in basic education.

Finally, a few evaluation reports also point out that in contexts such as these, gender projects have a large potential to do harm, by creating a backlash against women who participate in projects which are seen by large segments of society as not compatible with traditional values.

8 Humanitarian aid

Humanitarian aid is a vital sector in fragile and conflict-affected states, and nowhere more so than in South Sudan, where on average for 2011–20, around 64 per cent of total ODA was directed towards the humanitarian sector. In Mali and Afghanistan the proportions are considerably lower, with 12 per cent in Afghanistan and 10 per cent in Mali.

The systematic review on Afghanistan did not include evaluation on humanitarian aid, and hence we can report findings only for Mali and South Sudan. Given the importance of the sector, the evidence base is relatively small, with only three dedicated project evaluations for Mali and seven for South Sudan.

In Mali, after the collapse of the democratic order and the onset of war in 2012, most donors increased their humanitarian aid, much of which was channelled through multilateral organizations. Humanitarian assistance, with food and in-kind aid as well as income-generating activities in rural areas, grew in importance and contributed to the rudimentary social safety net. Humanitarian interventions helped to prevent malnutrition in the conflict-affected areas in the north. In addition, school feeding programmes had positive effects on enrolment and attendance, and the increased use of cash-based transfers (for example for school feeding and for nutrition support) increased aid efficiency. Especially promising were programmes which combined food aid (such as cash, voucher, or food transfers) with the provision of assets such as irrigation channels, flood protections, or home gardens.

In South Sudan, humanitarian assistance became the dominant aid sector by 2014. Humanitarian interventions were reasonably effective in preventing the bad from getting worse. A substantial proportion of humanitarian aid was channelled through multilateral pooled funds such as the UN-managed Common Humanitarian Fund for South Sudan. This fund was commended for its capacity to absorb large grants and use contributions in a strategic yet flexible way. Measures implemented through the fund included general food assistance, cash-based transfers, school feeding, and large-scale distribution of livelihood kits. The reports note that cash-based transfers were often more cost-efficient and timely than in-kind transfers, but that this modality could not be rolled out in all regions due to the difficult context. Other reasonably effective interventions were nutrition projects, clean water delivery projects, WASH projects, setting up VSLAs, and cash transfers for small livestock or inputs such as seeds, tools, or fishing kits. Despite these achievements, many evaluations of humanitarian programmes noted that the effects on long-term resilience were probably small.

While the evidence base is relatively small, the evidence suggests that humanitarian aid was reasonably effective. It remains unclear, however, whether interventions also contributed to more resilience and thus produced sustainable results, or whether humanitarian assistance remained a band-aid.

One more observation is worth making: the available evaluation reports in the field of humanitarian assistance almost never discuss the potential of humanitarian aid to do harm. Given that there is a very substantial literature on the topic, and that examples of the destabilizing impact of misused aid are well documented in many fragile and conflict-affected contexts, not least in South Sudan, this seems a glaring omission. We strongly believe that evaluations of interventions in fragile states should pay more attention to unintended negative consequences. The absence of this, in the sector of humanitarian aid interventions, but also in all other sectors included here, is surprising.

9 Why aid is mostly not effective in fragile states

The evidence suggests that there are differences across sectors in aid effectiveness in highly fragile states.

Interventions in education and health were reasonably effective, although it is unclear how sustainable the results are. Also reasonably effective were rural development programmes, which contributed to improved livelihoods and strengthened coping mechanisms. Rural women also benefited from such livelihood programmes. However, rural development programmes did not lead to sustainable economic growth that translated into jobs or income opportunities.

Programmes supporting macroeconomic development, macroeconomic policies, financial management, and support for the private sector were mostly not effective and generally did not contribute to sustainable economic growth. Likewise, interventions aiming to transform gender relations had little impact.

Interventions in the ‘good governance’ sector, capacity-building for central government, and interventions aiming to contribute to stabilization were not effective.

Why is this so? The reasons are manifold, but four intertwined factors stand out from the evaluation reports: first, the distinct political economy of fragile states; second, a lack of capacity of the government; third, endemic violence; and fourth, overambitious interventions.

As mentioned above, in highly fragile states political power resides not with formal political institutions but rather in networks of patronage which transcend the formal political institutions. Rent-seeking, widespread institutionalized corruption, and intense competition between rival networks are the typical symptoms of such a political economy of fragile statehood. One important implication is that elites in such a system are not interested in political reforms which would endanger a mode of governing that is the sole source of authority. Consequently, there is a distinct lack of political will to truly embrace reforms which could lead to more accountability, more-robust formal institutions, and more-democratic procedures. Consequently, the effectiveness of aid programmes is especially low in fields such as good governance, decentralization, anti-corruption, and similar fields.

This situation is without doubt exacerbated by a lack of capacity. Many evaluation reports stress that governmental structures lacked the capacity to deal with aid flows in a productive way. Unfortunately, many donors consistently overestimated state capacity, especially in Afghanistan, and designed programmes based on largely imagined absorptive and administrative capacity. Unfortunately, as we have seen, building capacity rarely works in such contexts, leading to a vicious circle that few fragile countries can escape. In addition, elites in such contexts are typically highly fragmented, with intense and often violent competition among different networks. Foreign donors, therefore, can rarely build reliable partnerships with government officials because there is a high rate of turnover and government officials are often preoccupied with internal power struggles.

The third factor explaining a lack of effectiveness is insecurity and violence. A constant theme in the evaluation reports is that the lack of basic security is a pervasive problem, affecting every aspect of development co-operation. The reports highlight how difficult it is to implement and monitor development projects when sites are not accessible, or when development workers are at risk of being targeted by insurgents. Many reports describe how a lack of security caused delays and cost overruns. Lack of security also necessitated that aid organizations employ security measures, which increased implementation costs.

Finally, a recurring theme in the evaluation reports is that donors often designed overly ambitious projects which were based not on the political-economic realities on the ground but rather on overly optimistic assumptions and ideologically charged theories of change.

Taken together, the distinct political economy of fragile states, a lack of capacity of the government, endemic violence, and overambitious planning go a long way in explaining why aid is rarely effective beyond sectors such as education, health, or livelihoods.

Paradoxically, the orthodox aid approach to fragile states is to use aid as a tool to overcome these blockages. Unfortunately, the available evidence suggests that development aid is not an effective instrument for this. Donors need to come to terms with the fact that while traditional development aid can help to improve basic livelihoods and service provision—albeit to a limited extent and non-sustainably—it has little political transformative capacity in fragile states.

10 Pockets of success

There are some pockets of success in highly fragile states: aid is somewhat effective in health, education, and rural development. For example, the evaluations suggest that basic health indicators improved (especially in the field of mothers' and children's health); that some capacities were built

in the health sector, leading to better health practices; and that enrolment and retention rates increased in basic education for boys and girls.

In the livelihood sector, support for small infrastructure improved food security. Interventions resulted in more skills in farming or income-generating activities, leading to more resilience. These gains were mainly achieved ‘in the heads’ of people, in the form of skills, knowledge, and practices. Institutional gains, however, were rare. For projects at community level, factors which were identified as helpful for achieving sustainable results were the buy-in and ownership of communities and their leaders, as well as connectedness and collaboration with NGOs and with the government, since they have to ensure that the project continues after external support has ended. Occasionally, interventions also led to increased technical capacities in subnational government structures regarding agriculture practices.

While interventions aiming at gender equality were not effective, some results were obtained in improving the situation of rural women. Examples include rural literacy, increased access to health and education, and better livelihoods in women-specific activities within agriculture, such as mushroom farming and kitchen gardening. Also, VSLAs helped to improve women’s economic status, and interventions aimed at better educational outcomes for girls were often effective. In general, gender projects with an economic component, specifically targeting women’s economic situation, fared better than norm-changing, awareness-raising interventions, since they had broader acceptance among women and men and provided tangible and immediate benefits.

While capacity-building for the central government was usually not effective, there were some pockets of success, almost exclusively in very technical, apolitical fields at the subnational level, for example in education, rural development, and health.

Generalizing the lessons from these pockets of success, we find that smaller and localized projects performed better than larger, more complex projects. Also, projects with tangible results were often effective, such as building small infrastructure; providing services such as access to water and electricity; or providing new skills in farming, hygiene. Results were less strong where programmes aimed to be transformative in nature, either for capacity-building or to change cultural and social norms.

Are the gains made in these pockets of success sustainable? We have only limited evidence to answer this question, since many evaluation reports do not address issues of sustainability, and *ex post* evaluation is rare. From what we can tell, the prospect for sustainability is highest in projects that rely on people’s skill (such as better farming practices, better health practices, etc.); that involve relatively simple infrastructure (such as simple irrigation, water wells, etc.); that rely on relatively simple organizational structures which function without external partners (such as VSLAs). Projects which were assessed to be sustainable consisted of, for example, small infrastructure for food security, VSLAs, farming and gardening skills, better irrigation, or health training and practices.

The prospects for sustainability were lower for projects which required continuous external funding, strong partner capacities, or complex infrastructure. Such projects consisted, for example, of larger irrigation schemes, funding for the health sector, larger income-generating programmes, or cash transfer programmes.

These findings are not overly surprising. In such difficult contexts, one of the few hooks for sustainability is achieving success ‘in the heads’ of people, in the form of skills, knowledge, and practices. Institutional transformation is much harder to achieve.

11 Implications

The evidence collected through these three systematic reviews paints a clear picture: development aid to the most fragile states does not work as hoped. The most important and most disappointing finding is that development aid is not a suitable tool for addressing the core problems of fragile states. Aid does not improve governmental capacity, does not lead to better governance, and does not provide more stability.

This systematic review puts this inconvenient evidence into the spotlight, but it cannot provide silver bullets. Yet faced with this evidence, the aid community can no longer ignore the problem. What we need now is an honest discussion about a new aid strategy in fragile states, one that takes its starting point in the acknowledgement that aid is not an effective tool for making a fragile state more stable, more capable, and better governed. This finding is not entirely new. There is a solid academic literature that argues that external actors rarely succeed in strengthening institutions in fragile states (Barnett and Zürcher 2009; Bliesemann De Guevara 2010; Chowdhury 2009; de Waal 2015; Englebert and Tull 2008; Ottaway 2002), that fragile states are often trapped in fragility (Carment and Yiagadeesen 2019; Collier 2008; Pritchett et al. 2013), and that aid faces many challenges in fragile states (Gisselquist 2014). Yet, there is still a prevailing sense among many scholars and practitioners that the effectiveness of aid could somehow be increased by fine-tuning aid modalities, by better adapting aid to local contexts, by increasing aid flows, or by staying engaged for a longer time. This systematic review strongly suggests that this is not the case.

11.1 Taking opportunity costs seriously

Every aid dollar that is spent on an intervention which is not effective cannot be spent on an intervention which might have been effective. This is known as opportunity costs. The evidence shows that interventions in fields such as stabilization, good governance, capacity-building for the central government, or gender projects are usually not effective in fragile contexts. Continuing to allocate resources to sectors where the probability of success is low is not only ineffective, it is also ethically wrong because it binds resources which could otherwise have been used to improve the lives of people—for example by providing shelter, increasing food security, improving access to health, or teaching children how to read and write. Perhaps the most important implication of our systematic reviews is therefore that aid managers should always consider opportunity costs when making allocation decisions.

Related to this is the need for more cost-effectiveness analyses in those sectors which have a reasonably high probability of results, for example education, resilience in rural areas, or health. With very few exceptions (all in health), the reviewed evaluations do not provide cost–benefit analyses. However, if we take opportunity costs seriously (as we should!), then we also should invest in more cost–benefit analysis of aid interventions.

11.2 Reconsidering ‘relevance’

Not once did the 322 evaluation reports reviewed identify a programme that was seen as ‘not relevant’, even though most programmes were not very effective. But this is because donors assess ‘relevance’ according to the needs of a country: when needs are endless, as in Afghanistan, Mali, and South Sudan, then every intervention is relevant. We doubt that such a definition is useful. Instead, we think ‘relevance’ should be assessed considering both needs and the probability of having a positive impact, given the conditions under which a programme is implemented. Adopting such a definition would lead us to classify programmes which are based on wrong

assumptions about the prospects for success as not relevant. That would help to minimize well-meaning but ill-defined programmes which are bound to fail.

11.3 Improving the quality and intellectual honesty of evaluations

Having read hundreds of project-level evaluations, we think that the overall quality of these evaluations can and must be improved. Often, baseline and endline data were not available, indicators were unrealistic, and in-field monitoring was lacking. In some instances third-party monitoring was used, but it is unclear to what extent it was effective. Another important field for improvement is the quality of surveys, which were often not appropriate for their purpose. Organizations which use surveys as part of their evaluations should take the extra step and ensure that their survey design is fit for purpose, for example by having it peer reviewed by scholars before implementation.

Furthermore, all project evaluations, and especially qualitative ones, should include an intellectual engagement with the counterfactual. It is true that it is not possible to measure what would have happened without an intervention. But it is always possible to conduct informed speculation about the counterfactual, and about factors besides the intervention which might have contributed to the observed outcome. This would make evaluations intellectually more interesting, more honest, and therefore more valuable.

11.4 More systematic review in different fragility contexts

We need more systematic reviews on aid at the country level. To the best of our knowledge, the three systematic summaries in this report are the first such reviews. Donors should embrace this instrument. Compared with the amount of money that donors spend in fragile states, systematic reviews are inexpensive to conduct, yet they provide much value.

In order to corroborate the findings presented in this report, we need more systematic reviews on aid in highly fragile states. But we also need systematic reviews on aid in less-fragile states in order to investigate whether aid effectiveness increases once states are less fragile. In our sample, we see that aid effectiveness in South Sudan—the most fragile country of the three—was very low across all sectors. Effectiveness appears to have been slightly higher in Mali—the least fragile country of the three—and especially so before the breakdown of the democratic order in 2012. In order to corroborate the intuition that aid effectiveness increases when fragility is reduced, we need systematic reviews on countries at different levels of fragility. By systematically varying the fragility score, we can learn about whether and how aid effectiveness across different sectors is correlated with initial fragility.

11.5 Engaging with contextual factors, specifying guiding assumptions

As well as evaluations from project level up to systematic reviews, project designs need to more systematically engage with the scope conditions and contexts which enable or prevent effectiveness. The guiding assumptions on which a project, a programme, or a country-level strategy is based should be critically examined against the backdrop of an analysis of the context, local power constellations, and the political economy. As mentioned above, there is a distinct political economy in very fragile states where elites have little incentive to strengthen formal state institutions but instead nurture their patronage networks. In such a context, many of the assumptions guiding development interventions may no longer hold. A good understanding of the context is therefore essential for designing and evaluating interventions and their theories of change. In defining such guiding assumptions, donors are well advised to consult the existing scholarly literature. For example, reading the evaluation reports on South Sudan, we were surprised

to see how little reference is made to existing analysis of the political economy in South Sudan, or in general in the Horn of Africa. The same holds true for Afghanistan. In both cases, the large academic literature on the political economy of these countries is mostly ignored. The guiding assumptions which appear to inform many aid strategies in South Sudan (and in fragile states in general) often ignore readily available insights from scholarly research about the inner workings of fragile states. There appears to be a disconnect between existing knowledge about how such fragile states ‘really work’ and the assumptions on which many interventions are based. This is not a problem of lack of knowledge; it is one of knowledge transfer.

11.6 Documenting and periodically re-evaluating key assumptions

Organizations should make it a habit to re-evaluate their key assumptions regularly and at short intervals. To give an example: a recurrent theme in the evaluation reports on Afghanistan is that donors time and again overestimated the capacity and political will of the government and the prospect for peace and therefore designed overambitious programmes. A regular re-evaluation of guiding assumptions would have prevented aid policies up until the collapse of Afghanistan from being based on the incorrect assumption that there was enough state capacity to handle the aid flows, and that the elites were interested in state-building.

Furthermore, donors should document in writing their contextual analysis and their main assumptions, for example in a written country strategy, which can then serve as the basis for a candid discussion about how the context and the political economy influence the prospect for effective aid, and as a tool for a regular assessment of whether guiding assumptions are still valid.

11.7 Accepting that stabilization does not work

One of our key findings is that stabilization in such contexts does not work. This is a blow to the international community, which over the last two decades has spent billions of aid money in highly insecure regions in the hope that it would lead to less violence and more stability. A recent wave of academic work on the impact of aid on violence has now amassed convincing empirical evidence that this hope is futile, and our results from Afghanistan, South Sudan, and Mali confirm this (Findley 2018; Iyengar Plumb et al. 2017; SIGAR 2018; Zürcher 2017, 2020). Aid might even increase violence, when injected into highly insecure regions where violence is a reality and insurgents still retain some capacity. It is beyond the scope of this report to review the main explanations for this in detail. The short version is this: at the national level, aid is not a strong-enough lever to fix the many woes that plague governments in fragile states, such as lack of capacity, predatory elites, and a lack of a political settlement among rival fractions. And at the micro level, aid is not a strong-enough incentive to convince the local population that they should side with the government and oppose insurgents. In addition, in regions where insurgent groups are present, aid will trigger a strategic reaction: insurgents will either shut down aid projects because they fear that aid may increase the co-operation between local communities and the government, or they will try to regulate aid flows in order to ‘tax’ it and to increase their prestige and legitimacy among local communities as enablers of it. These strategic responses will lead to more immediate violence (since shutting down aid projects or ‘taxing’ requires violence or the threat of violence) and increase the capabilities of insurgents for future violence (because of the profits made from aid).

These findings should prompt donors to prioritize a development logic over a stabilization logic. Providing resilience and welfare gains to an impoverished population in a conflict situation, while avoiding political backlash, is no small achievement and a worthy goal.

11.8 Paying attention to ‘doing no harm’

Surprisingly, doing no harm is rarely mentioned in the reviewed evaluation reports. We think that donors and evaluators should engage much more with the issue of doing no harm, since the risk of aid doing harm in such contexts is very high. For example, power-brokers deny or grant access for humanitarian aid, thereby using aid as a means to reward or penalize certain groups, which will exacerbate conflict. Furthermore, aid is often stolen or ‘taxed’, and profits can be reinvested in the organization of violence. Competition for aid can fuel conflict between groups, and aid projects can increase resource competition. Donors should make much more effort to identify and mitigate the ways in which aid can be misused by parties to conflict, and evaluators should make a do-no-harm assessment mandatory for their evaluations in such contexts.

An important lesson from these systematic reviews is that aid has a fair chance of being effective in highly fragile contexts only when programmes are modest, small rather than large; do not assume unrealistic partner capacity; are designed with an awareness of the context in mind; do not spend aid money too fast; do not spend aid money in highly insecure regions controlled by insurgents; and, most importantly, avoid aid sectors where aid is most probably not effective. Such principles amount to a new paradigm: ‘grand visions out, local tangible small gains in’.

One consequence of such an approach would be that development aid becomes more modest and slower. The reviewed evaluation reports strongly suggest that there would be benefits. In general, smaller projects performed better than larger, more-complex projects. Also, projects aimed at direct results were often reasonably effective, such as those building small infrastructure; providing services such as access to water, electricity, and education; or improving livelihoods. Results were less strong where programmes aimed to be transformative in nature, either for capacity-building or to change cultural and social norms. Projects would also become more localized. Localization requires that donors have a better understanding of the situation on the ground in different subnational contexts or even communities. This would enable donors to fine-tune their interventions to the local context. For example, some locations may have good-enough governance and good-enough security that interventions which require co-operation with local governments can be effective, whereas in other locations the risk of aid being misused by local power-brokers is too high. Ideally, donors would then choose approaches and modalities which are appropriate to the given context, thereby maximizing effectiveness and minimizing the risk of doing harm.

As we have argued, the grand visions for aiding fragile states have mostly failed, and that failure has been costly, not only because of the resources spent but also because of the opportunity costs of not having used resources on programmes which might have improved the lives of beneficiaries. In this respect, the failure of grand visions has very tangible consequences for the populations in fragile and conflict-affected situations.

Switching ‘from grand visions to tangible local gains’ may be a bitter pill to swallow for many aid agencies. While it is true that the allocation of aid is predominantly driven by domestic political considerations, bureaucratic inertia, ideology, sunk costs, and the vested interests of various stakeholders, it is also true that most aid practitioners are driven by a genuine desire to be ‘transformational’—that is, to use aid as a lever to transform societies so that that structural impediments to development are removed. Moving to this new paradigm would mean that both old routines (‘doing things because we have always done them’) and old visions (‘we can be transformational’) need to be replaced. It would also mean that aid allocation would largely avoid sectors where needs are large but the probability of success is low (such as stability, good governance, capacity). Instead, some aid would go to sectors where aid can be effective but results are probably not sustainable without prolonged financing from donors (for example health and

food security). Donors would therefore need to accept that some of the more effective interventions are not sustainable and are prone to long-term dependency. But such permanent band-aids may still be a better option than either neglect or grand but ineffective transformational aspirations.

References

- Barnett, M., and C. Zürcher (2009). 'The Peacebuilder's Contract: How External Statebuilding Reinforces Weak Statehood'. In R. Paris and T.D. Sisk (eds), *The Dilemmas of Statebuilding: Confronting the Contradictions of Postwar Peace Operations*. London and New York: Routledge.
- Bliesemann De Guevara, B. (2010). 'Introduction: The Limits of Statebuilding and the Analysis of State-Formation'. *Journal of Intervention and Statebuilding*, 4(2): 111–28. <https://doi.org/10.1080/17502970903533652>
- Carment, D., and S. Yiagadeesen (2019). *Exiting the Fragility Trap: Rethinking Our Approach to the World's Most Fragile States*. Athens, OH: Ohio University Press. <https://doi.org/10.2307/j.ctv224tvcv>
- Chowdhury, A. (2017). *The Myth of International Order: Why Weak States Persist and Alternatives to the State Fade Away*. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780190686710.001.0001>
- Collier, P. (2008). *The Bottom Billion: Why the Poorest Countries Are Failing and What Can Be Done About It*. Oxford: Oxford University Press.
- De Waal, A. (2015). *The Real Politics of the Horn of Africa: Money, War and the Business of Power*. Cambridge: Polity Press.
- Englebert, P., and D.M. Tull (2008). 'Postconflict Reconstruction in Africa: Flawed Ideas about Failed States'. *International Security*, 32(4): 106–39. <https://doi.org/10.1162/isec.2008.32.4.106>
- Findley, M.G. (2018). 'Does Foreign Aid Build Peace?'. *Annual Review of Political Science*, 21: 359–84. Available at: www.annualreviews.org/doi/full/10.1146/annurev-polisci-041916-015516 (accessed 2 November 2022).
- Gisselquist, R.M (2014). 'Aid and Institution-Building in Fragile States: What Do We Know? What Can Comparative Analysis Add?'. *The ANNALS of the American Academy of Political and Social Science*, 656(1): 6–21. <https://doi.org/10.1177/0002716214546991>
- Iyengar Plumb, R., J.N. Shapiro, and S. Hegarty (2017). *Lessons Learned from Stabilization Initiatives in Afghanistan: A Systematic Review of Existing Research*. Santa Monica, CA: RAND Corporation. Available at: www.rand.org/pubs/working_papers/WR1191.html (accessed 2 November 2022).
- Noyes, J., A. Booth, M. Cargo, K. Flemming, A. Harden, J. Harris, R. Garside, K. Hannes, T. Pantoja, and J. Thomas (2022). 'Qualitative Evidence'. In J.P.T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M.J. Page, and V.A. Welch (eds), *Cochrane Handbook for Systematic Reviews of Interventions*, version 6.3 (updated February 2022). London: Cochrane. Available at: <https://training.cochrane.org/handbook/current/chapter-21> (accessed 2 November 2022).
- Ottaway, M. (2002). 'Rebuilding State Institutions in Collapsed States'. *Development and Change*, 33(5): 1001–23. <https://doi.org/10.1111/1467-7660.t01-1-00258>
- Pritchett, L., M. Woolcock, and M. Andrews (2013). 'Looking Like a State: Techniques of Persistent Failure in State Capability for Implementation'. *The Journal of Development Studies*, 49(1): 1–18. <https://doi.org/10.1080/00220388.2012.709614>
- Sexton, R. (2016). 'Aid as a Tool against Insurgency: Evidence from Contested and Controlled Territory in Afghanistan'. *American Political Science Review*, 110(4): 731–49. <https://doi.org/10.1017/S0003055416000356>

- SIGAR (Special Inspector General for Afghanistan Reconstruction) (2018). *Stabilization: Lessons from the U.S. Experience in Afghanistan*. Available at: www.sigar.mil/interactive-reports/stabilization/index.html (accessed 2 November 2022).
- Zürcher, C. (2017). 'What Do We (Not) Know about Development Aid and Violence? A Systematic Review'. *World Development*, 98: 506–22. <https://doi.org/10.1016/j.worlddev.2017.05.013>
- Zürcher, C. (2020). 'The Impact of Development Aid on Organised Violence'. 3IE Working Paper 37. New Delhi: International Initiative for Impact Evaluation (3IE). Available at: <https://www.3ieimpact.org/sites/default/files/2020-08/WP37-Systematic-Review-Aid-Violence.pdf> (accessed 2 November 2022).

Appendix: Data sources for the search

Databases	Bilateral donors	Multilateral and international organizations	Repositories of impact evaluations in international development*	Major developmental NGOs
Academic Search Complete	US/USAID (Development Experience Clearinghouse)	African Development Bank (AfDB)*	3ie RIDIE (Registry for International Development Impact Evaluations)	ACTED*
AfricaBib.org*	UK (Foreign, Commonwealth and Development Office, formerly DfID)	African Union*	3ie Development Evidence Portal	ActionAid*
Cairn*	Canada (Global Affairs Canada/GAC)	European Bank for Reconstruction and Development (EBRD)	AgEcon	Aga Khan Development Network
EconLit	Australia (Department of Foreign Affairs and Trade/DFAT)	European Investment Bank*	AGRIS (International System for Agricultural Science and Technology)	CARE International
Érudit*	New Zealand (Ministry of Foreign Affairs and Trade/MFAT)	European Bank for Reconstruction and Development (EBRD)	BREAD (Bureau for Research and Economic Analysis of Development)	Catholic Relief Services
GenderWatch*	Germany (KfW, German Agency for International Cooperation/GIZ, and Federal Ministry for Economic Cooperation and Development/BMZ)	European Investment Bank	Center for Effective Global Action (CEGA)	Danish Refugee Council*
Global Health*	France (Agence française de développement AFD)	European Union	CGIAR: Consultative Group on International Agricultural Research	IRC (International Rescue Committee)
International Pol. Sci. Abstracts	Italy (Italian Agency for Development Cooperation (AICS))	International Bank for Reconstruction and Development (IBRD; part of the World Bank Group)	Deval (German Institute for Development Evaluation)	Médecins sans frontières
MEDLINE*	Sweden (SIDA)	International Fund for Agricultural Development (IFAD)	GEF (Global Environmental Facility)	Mercy Corps
PAIS Index	Norway (NORAD)	International Monetary Fund (IMF)	Global Facility for Disaster Reduction and Recovery	Oxfam International
Pascal (up to 2015)*	Denmark (Danida)	International Organization for Migration (IOM)	ICNL Research Centre	Plan International*
RePEc (Research Papers in Economics)/IDEAS*	Finland (Finnida)	UNMAS, United Nations Mine Action Service	IFPRI (International Food Policy Research Institute)	Samuel Hall*
Web of Science	Belgium (Enabel)	UN	Independent Development Evaluation, AfDB	Save the Children*

Worldwide Pol. Sci. Abstracts	Netherlands (Ministry of Foreign Affairs)	United Nations Children's Fund (UNICEF)	J-Poverty Action Lab (J-PAL)	Welthungerhilfe
	Switzerland (DEZA)	United Nations Conference on Trade and Development (UNCTAD)	Millennium Challenge Corporation	World Vision
	Japan (JICA)	UNDP, Evaluation Resource Center	RePEC IDEAS	HALO Trust*
	China (China International Development Cooperation Agency/CIDCA)*	United Nations Environment Programme (UNEP)		Oxfam Novib*
		United Nations High Commissioner for Refugees (UNHCR) United Nations Industrial Development Organization (UNIDO) World Bank Group (esp. World Bank e-library, CAS Completion Report Review, Country Performance Portfolio Review, IEG Evaluations, Impact Evaluation) World Food Programme (WFP), Evaluation Library* World Health Organization (WHO)* UN Women, GATE System OECD DReC Asian Development Bank**		Save the Children NL*

Note: * not searched for Afghanistan; ** searched only for Afghanistan.

Source: author's construction based on own work.