In Search of Results
Incentive-Oriented Programs in Afghanistan, 2002-2021

Aman Farahi
Afghanistan Research Network

This series highlights the work and analysis of the Afghanistan Research Network (ARN), a project convened by LSE / PeaceRep, and the Civic Engagement Project (CEP). The network brings together over 20 Afghan researchers (and several non-Afghans) with diverse expertise and backgrounds investigating a range of issues. This project aims to support Afghan researchers who were recently forced to leave Afghanistan; to ensure expert and analytical provision; inform contextually-appropriate international policies and practices on Afghanistan; and to deepen understanding of evolving political, security, and economic dynamics.

PeaceRep: The Peace and Conflict Resolution Evidence Platform
School of Law, Old College, The University of Edinburgh
South Bridge, Edinburgh EH8 9YL
Tel. +44 (0)131 651 4566
Fax. +44 (0)131 650 2005
E-mail: info@peacerep.org
Website: PeaceRep.org
Twitter: @Peace_Rep_

The Afghanistan Research Network is supported by the Peace and Conflict Resolution Evidence Platform (PeaceRep), funded by the UK Foreign, Commonwealth & Development Office (FCDO) for the benefit of developing countries. The information and views set out in this publication are those of the authors. Nothing herein constitutes the views of FCDO. Any use of this work should acknowledge the authors and the Peace and Conflict Resolution Evidence Platform.

About the author:

Aman Farahi is a Washington, D.C.-based independent consultant. He worked as an Economist at the World Bank, Senior Economist at UNDP, and Director General for Policy in the Afghanistan Ministry of Finance.

Cover images: Getty Images. All images may be subject to copyright. ©2023

Design: Smith Design Agency

## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key Takeaways</td>
<td>01</td>
</tr>
<tr>
<td>Introduction</td>
<td>02</td>
</tr>
<tr>
<td>Conceptual Framework of Results-Based Programs</td>
<td>03</td>
</tr>
<tr>
<td>Incentive- and Results-Oriented Aid Programs in Afghanistan</td>
<td>05</td>
</tr>
<tr>
<td>Lessons</td>
<td>13</td>
</tr>
</tbody>
</table>
Key Takeaways

The use of incentives for results in official development assistance programs had mixed results in Afghanistan. Whereas general conditionalities, which lacked measurable indicators and direct monetized incentives, were largely ineffective, specific budget support results-based programs incentivized necessary reforms and development.

Four key lessons can be observed from the experience of incentive for results programs in Afghanistan:

- First, most aid programs were not results-based, meaning that most donors’ $70 billion USD ODA pledges continued to flow to Afghanistan despite poor results. This meant there was no real incentive to commit to reforms, and an opportunity was lost to achieve meaningful change.

- Second, incentive-based frameworks had hundreds of indicators, some of which did not contribute to the government’s self-reliance agenda and often usurped government priorities in favour of donor-selected and fragmented priorities. If indicators were selected more strategically, they could have helped lead to larger-scale systems reform.

- Third, incentive programs focused on specific activities and departments within the larger government system and did not impact the wider public sector. This means that indicators focused on specific government bills, such as improving the budget or specific sectoral ministries, but did not tackle the systemic risk of corruption endemic in the wider public sector.

- Finally, while results-based programs helped bring about important reforms in specific government areas, there was a missed opportunity to have an aggregate impact – meaning that the incentive programs were no more than the sum of their parts.
Introduction

Incentives for results are a good idea. Yet, getting incentives right and identifying results that matter has proven to be much more complicated, especially in the case of Afghanistan. From 2002-2020, Afghanistan received at least $70 billion USD in official development assistance (ODA) from OECD donors and $72 billion USD in military aid from the United States (Organisation for Economic Co-operation and Development; 2023, Brown University; 2021). A portion of the ODA was channelled through incentives mechanisms. Additionally, the International Monetary Fund (IMF), the World Bank, the European Union (EU), and the broader donor government compact were all designed with incentives and results at the heart of their programs. The aim was to encourage the Afghan government towards reforms.

On-budget, incentive-based programs, which accounted for a small share of the ODA, had two important benefits. First, they provided the Afghan government with predictable financing, which paid for key programs, salaries, and development initiatives. Second, where incentive programs were results-based, they encouraged reform of key aspects of the public sector machinery, despite a highly volatile security situation and complex political economy environment.

With the power of hindsight, it is perhaps easy to group all interventions in Afghanistan over the last 20 years as a failure, as many evaluations have stated. The Special Inspector General for Afghanistan Reconstruction (SIGAR) “lessons learned” report and a meta-review of 148 donor evaluations identify a lack of security, insufficient understanding of the context, a volatile political economy, and weak capacity as constraints, which rendered aid to Afghanistan ineffective over the two decades (Special Inspector General for Afghanistan; 2021, Zürcher; 2020). Although broad evaluations can offer important insights, they focus on lessons at the aggregate level, ignoring specific program-level results, and critically lack analysis on the impact of implementation and funding modalities which contribute to program efficacy.

This report provides an overview of the results-based financing (RBF) initiatives in Afghanistan. It explores the development and execution of incentive and results-based programs and identifies key takeaways to enhance future programs in fragile and conflict-affected situations (FCS).
Conceptual Framework of Results-Based Programs

a. Incentives

In welfare economics, when incentives diverge between two parties, it can lead to inefficiency giving rise to adverse selection and moral hazard problems (Martens et al.; 2002). In the aid effectiveness field, adverse selection problems occur when the recipient country has weak governance and poor policies to use donor funds effectively (Martens et al.; 2002). Moral hazard issues arise when the recipient country changes reform behaviour after receiving aid, and the governing elite act in self-interest rather than maximizing the impact of aid funds to improve the welfare of society (Martens et al.; 2002).

b. Results-Based Financing

RBF is an approach in which focus shifts from inputs to results—effectively aligning stakeholder incentives around results—thus minimizing adverse selection and moral hazard problems (Pearson; 2011). RBFs can be tied to results in a range of ways. Some popular RBFs include output-based aid, social impact bonds, development impact bonds, a program for results, results-based climate finance, output-based disbursement, cash-on-delivery aid, Conditional Cash Transfers, and performance-based financing. The public, private, and NGO sectors can all be potential implementers of RBFs (Pearson; 2011).

RBF instruments vary depending on who is given incentives, ranging from individuals and households to national governments. For interventions to be considered "results-based", three key features must be present in the design. First, funding needs to be contingent on results. Second, the results need to be identified and agreed upon upfront. And third, an independent party verifies the delivery of the results before payments can be made (Pearson; 2011). In practice, implementation contexts and delivery capacity differ across countries, and an RBF approach that works in one context may not work in another.
c. Risks

Three broad risk categories need to be balanced between the funder and the recipient in implementing RBFs. The first risk category involves operational risks concerning the technical and financial ability of the implementer (The World Bank; 2018). If the implementer does not have enough financing to frontload the investment or the know-how to achieve the specified results, then RBF instruments will likely fail. In some cases, funders and implementers can share operational risks, with intermediate payments being made along the results chain rather than only at the end. Where sufficient capacity exists, the risk is completely transferred to the implementer (The World Bank; 2018).

The second risk category relates to performance issues where the implementer cannot achieve results efficiently because of cost overruns or other shortfalls. The last type of risk concerns variance in service demand, which can lead to underutilisation or overutilisation, making the results ineffective (The World Bank; 2018).

d. Prerequisites

RBFs require prerequisites, such as country ownership, a conducive policy environment, macroeconomic stability, and a minimum level of public financial management systems to succeed (Koeberle et al.; 2006). In addition, several context-specific pre-conditions also need to be considered to ensure success, such as suitable interventions, measurable results, data availability, the capacity of funders and recipients, system readiness, and stakeholder buy-in and alignment (Koeberle et al.; 2006). These pre-conditions ensure that the requisite infrastructure and skills are in place to work towards shared goals.

The strength of RBFs lies in focusing on results, shifting away from ex-ante conditionalities to partnerships, enhancing ownership and medium-term focus, selecting programs, and scalability when successful. RBFs also promote accountability and transparency around results (Koeberle et al.; 2006). The weaknesses of RBFs include fiduciary risks if systems are weak, limited capacity building, trade-offs between monetised and un-monetised priorities, and aid volatility if performance risks are high (Koeberle et al.; 2006).
Incentive- and Results-Oriented Programs in Afghanistan

a. Overview

Afghanistan was among the early adopters of incentive-oriented reform programs, despite a fragile context. Such programs aimed to provide the government with much-needed funds while tying the grants to completing reform activities categorized under state-building, strengthening the country's public financial management systems, and improving service delivery. The government, on its part, preferred such programs to show its willingness to implement reforms in exchange for development aid.

Four large programs in Afghanistan fit the characterisation of incentivised programs. These included budget support operations that had a policy conditionality linked to disbursement, such as the IMF’s programs, the World Bank-managed Afghanistan Reconstruction Trust Fund (ARTF), the European Union-financed State and Resilience Building Contract (SRBC), and the overall donor government framework of mutual accountability, which provided the overarching mechanism for aid.

IMF programs, the World Bank, and the EU’s budget support operations and donor-government compacts are not new to fragile and conflict situations; such interventions have matured and are standardised. The World Bank’s budget support operations were the longest running, disbursing more than $4 billion USD to the government during its operations (Haque & Nassif; 2021). The EU’s SRBC program was part of a broader engagement, providing budget support to FCS. The IMF’s program laid the groundwork for the rest of the aid by creating macroeconomic stability and strengthening the banking and financial systems. Lastly, the mutual accountability framework between donors and the Afghan government coordinated donor-wide conditionalities so the Afghan government could continue to receive aid.

These three programs were similar in design and had two key intervention rationales. First, they helped support the core budget directly (which financed key services, such as teachers’ salaries) by providing discretionary resources to the government. Second, these programs supported structural and public financial management reforms so the government could move towards longer-term fiscal stability.
In contrast to these three programs, Afghanistan's donors also instituted the Donor-Government Mutual Accountability Framework, which was 'results-inspired', or attempted to mimic results-based financing (RBF) instruments without technical rigor.

The RBF and results-inspired programs identified results ex-ante, although there was a proliferation of indicators rather than a cohesive identification of selective targets. Funding was contingent on results for RBF programs but not for the results-inspired program. The results were not solely outputs or outcomes; sometimes, activities further up in the results chain also qualified as potential payments triggers. Only the EU budget support program used an independent party to collect information and verify the results, which is a gold standard for objective evaluations. The remainder of the programs used in-house verification or none.

A summary of the design of these four programs is presented in Table 1. The following sections provide a background of the Mutual Accountability Framework and the World Bank budget support operation, elaborating on each intervention's design features.
### Table 1
*Summary of Incentive-Oriented Programs in Afghanistan*

<table>
<thead>
<tr>
<th>Intervention Rationale</th>
<th>IMF</th>
<th>EU</th>
<th>World Bank</th>
<th>Donor-Government Mutual Accountability Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rationale</td>
<td>Macro economic, Banking, Financial Stability</td>
<td>Public Policy, Public Financial Management</td>
<td>Structural Reforms, Public Financial Management</td>
<td>Political, Human Rights, Aid Effectiveness, and Development</td>
</tr>
<tr>
<td>Upfront Agreement on Results</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Number of Results per year</td>
<td>7-8 Quantitative targets, 7-8 Structural reforms</td>
<td>20 plus</td>
<td>10-13</td>
<td>Total: 24 16 Governmental, 8 Donors with sub-indicators</td>
</tr>
<tr>
<td>Funding linked to Results</td>
<td>Yes</td>
<td>All fixed tranches must be met for eligibility and variable for funds</td>
<td>Yes</td>
<td>Directly: None. Indirectly: Highest-Level compact to facilitate aid flow to Afghanistan</td>
</tr>
<tr>
<td>Verification and Reporting</td>
<td>Government submits evidence to IMF</td>
<td>Third-Party collects evidence on behalf of EU</td>
<td>Government submits evidence to World Bank</td>
<td>Self-reported progress: Donors and Government. UN coordinates</td>
</tr>
<tr>
<td>Frequency of Payment</td>
<td>IMF</td>
<td>EU</td>
<td>World Bank</td>
<td>Donor-Government Mutual Accountability Framework</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------</td>
<td>---------</td>
<td>------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Quarterly</td>
<td>Biannual</td>
<td>Quarterly</td>
<td>No payments. Pledging conference every four years</td>
</tr>
<tr>
<td>Annual Funds</td>
<td>US$ 40 million ECF</td>
<td>EUR 100 million</td>
<td>US$ 300-400 million</td>
<td>Donor pledging conferences: Basis for progress reports</td>
</tr>
<tr>
<td>Responsible Government Unit</td>
<td>Ministry of Finance</td>
<td>Ministry of Finance</td>
<td>Ministry of Finance</td>
<td>Ministry of Finance</td>
</tr>
<tr>
<td>Third-Party / Independent Verification</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Note. Author’s compilation.
b. Mutual Accountability Frameworks Between Donors and the Afghan Government

In 2004, after setting the political foundations for a new state, the newly formed Afghan government and the international community entered a new era of partnership, focusing on establishing a functioning and self-reliant government in Afghanistan (Fields & Ahmed; 2011). This spirit was best captured by the Afghanistan Compact of 2006, which enshrined the principles of the donor-government partnership and a set of aspirational results (Fields & Ahmed; 2011). While the compact was not classic results-based financing, it established results-based partnerships and continuing international aid to Afghanistan through donor-pledging conferences every four years, depending on progress towards meeting the donor-government compact.

The compact focused on security, governance, the rule of law, human rights, and economic and social development (Fields & Ahmed; 2011). It outlined a set of targets in these areas but stopped short of directly linking monetary incentives. Since its inception in 2006, there have been five major versions of the international community's compact with the Afghan government, including the Tokyo Mutual Accountability Framework 2012, Smart Mutual Accountability Framework 2016, Geneva Mutual Accountability Framework (GMAF) 2018, and GMAF 2020. The various frameworks were precursors to pledging conferences. However, since they compiled preferences from many donors and no funds were directly linked to performance, the results were among the weakest, even when well-implemented.

Under its last fully-implemented iteration of GMAF 2018, the government was required to highlight priorities, increase revenues, and report performance on benchmarks. In exchange, donors would provide predictable aid, improve on-budget support, build capacity, strengthen public financial management (PFM), and provide information on aid flows and how effectively off-budget aid was spent. A review of GMAF in its final months of implementation in 2020 noted several shortcomings, including weak design and relevance of indicators; an overstretch of strategic focus to donor priorities rather than a unified government development vision; and data, reporting, and measurement issues (AnA Consultancy; 2020). The review also highlighted that only seventeen out of 110 actions were completed. As summarised in Table 2, at least 67 indicators—more than half of the portfolio—could not be met for various reasons, such as differences in interpretation of the benchmarks and exogenous issues, the onset of COVID-19, and the worsening security situation (summarised in Table 2).
Some indicators that were considered “failed”—including reforms on land management, fiscal regimes for extractives, asset declarations, and e-governance—are challenging even for middle-income countries to tackle. It is not surprising that the government of Afghanistan considered these reforms failures, as insecurity was increasing and government attention was diverted towards maintaining security and advancing peace with the Taliban.

Table 2
GMAF 2020 End-of-Year Progress

<table>
<thead>
<tr>
<th>Status</th>
<th>Number of Indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed, confirmed, and uncontested</td>
<td>17</td>
</tr>
<tr>
<td>Completed, but contested/ambiguous</td>
<td>11</td>
</tr>
<tr>
<td>Pending but uncontested</td>
<td>15</td>
</tr>
<tr>
<td>Pending, contested, and ambiguous</td>
<td>9</td>
</tr>
<tr>
<td>Incomplete or failed</td>
<td>11</td>
</tr>
<tr>
<td>Delayed</td>
<td>25</td>
</tr>
<tr>
<td>Performance impaired by insecurity</td>
<td>5</td>
</tr>
<tr>
<td>Performance impaired by COVID-19 pandemic</td>
<td>17</td>
</tr>
<tr>
<td><strong>Total Indicators</strong></td>
<td><strong>110</strong></td>
</tr>
</tbody>
</table>

c. World Bank Budget Support Operation in Afghanistan

Budget support operations is a subset of results-based aid where the funder incentivises national governments to achieve shared development objectives. Budget support operations give discretion over aid to the host countries and help fund structural and economic reforms to reduce poverty and spur economic growth (Koeberle et al.; 2006). Budget support operations also involve dialogues between the funder and recipient on reforms and results, assessing progress, and transferring funds when the results are verified (Pearson, M.; 2011).

The World Bank-administered Afghanistan Reconstruction Trust Fund (ARTF) was established in 2002 and disbursed $12 billion USD by 2020 (Haque & Nassif; 2021). It consisted of budget support programs and an investment project support window. The rationale for the budget support program was to support recurrent costs to improve overall service delivery, strengthen state legitimacy, and achieve progress toward fiscal sustainability (Haque & Nassif; 2021).

Under the budget support program, structural and public finance reforms were developed with RBF-type conditionalities to assist the government in implementing reforms. Ten benchmarks were formulated each year. Funding of approximately $30-40 million USD was tied to each reform benchmark, which provided government incentives for the timely completion of criteria given the cash flow needs of the treasury. If the actions were completed within the timeline, one hundred percent of the funds were disbursed. The incentive funds would be discounted to zero after a year of delay. The World Bank’s technical teams helped devise some benchmarks and provided ongoing technical assistance to help the government meet the criteria. Once the requirements were completed, the government would submit confirmation to the World Bank, whose staff would then confirm in-house if the benchmarks were completed. An assessment decision to disburse funds would be made each quarter to help provide the government with cash throughout the year.
The World Bank’s budget support operations performed better than the mutual accountability framework. Over ten years, 23 benchmarks remained unmet within their timelines, and only two ultimately failed—compared to eleven failed actions under a two-year mutual accountability program from 2018 to 2020 (Haque & Nassif; 2021). The two never-completed benchmarks belonged to a series of customs administration reforms, and successive Afghan governments were notorious for corruption in this area. The World Bank budget support operations disbursed over $4 billion USD for 53 actions achieved on time, seven actions completed with delay, 21 achieved after the program, and two never achieved (Haque & Nassif; 2021).

The success of the World Bank budget support operations can be credited to selectivity in reforms, technical assistance provision to help the government achieve those reforms, and an explicit linkage of policy actions to incentive payments.

However, not all aspects of the World Bank budget support operations were successful. The ARTF recurrent and capital cost window evaluation found the operations and maintenance facility was "a clear failure" (Haque & Nassif; 2021). Furthermore, they note the risk of "paper-based" reforms, which are superficial even if implemented well (Haque & Nassif; 2021). In addition, the eligibility criteria—which indicates if government expenditure was conducted according to proper procurement and documentation processes—kept declining over time. In other words, while individual benchmarks were being completed, the overall system was heading in the wrong direction.
Lessons

a. Strengthen Focus on Results at the Aggregate Level

While the World Bank, European Union, and International Monetary Fund programs had an incentivised accountability mechanism where funds were disbursed after results completion, the overall donor-government mutual accountability compacts did not have results-based payments. The mutual accountability frameworks did not capitalise on the large amount of money committed through pledging conferences to encourage donors and government to achieve meaningful reforms. From 2002 to 2020, seven donor-pledging conferences were held, with donors pledging over $70 billion USD in official development assistance. Separate security sector pledging conferences were also held to grant Afghanistan security sector funds.

An opportunity was lost to tie at least half of the funds to meaningful results. A stalemate on the battlefield with the Taliban, successive fraudulent presidential elections, and a significant slow-down of development results intensified the adverse selection and moral hazard problems around aid. However, the message to the Afghan government and political elite at this point was that funds would continue to flow in regardless of the results.

b. Select Results Strategically to Advance Country Priorities

Most of the incentive-based frameworks in Afghanistan had hundreds of sub-indicators. It was not clear if performing a reform series had any quantifiable benefits apart from a checklist of regulations formulated or laws approved.

The GMAF, for example, had too many disparate indicators that were difficult to implement and did not directly contribute to Afghan self-reliance, even when implemented. The authors of the GMAF review note that GMAF had “usurped the role of Afghanistan National Peace and Development Framework as the central document guiding development and development cooperation in Afghanistan” (AnA Consultancy; 2020). Selecting indicators that do not advance the country’s development vision directly contradicts the aid effectiveness principles of “ownership” and “alignment” that the GMAF was expected to enforce.
c. Move From a Program-Centric to a System-Centric Approach

While the IMF, EU, and World Bank instruments were well-designed to ring fence program expenditures, they failed to reduce corruption throughout the system. The World Bank had hired a third-party fiduciary agent to ensure that funded expenditure was spent as intended, rather than on improving the fiduciary health of the Afghan system overall. This point is also acknowledged in the ARTF evaluation and further, that “an opportunity may have been lost” (Haque & Nassif; 2021). This action carefully selected a clean deck of government bills to fund, without taking on the systemic risk of corruption endemic to the entire public sector.

For example, from 2002 to 2018, expenditure’s “eligibility ratio” remained at around seventy percent, trending downwards after 2015 (Haque & Nassif; 2021). The eligibility ratio determined how much of the spending was conducted following existing regulations, World Bank procurement guidelines, and other internal controls and standards. In other words, seventy percent of the total expenditure had followed the process, had supporting documentation, and was spent as intended. The other thirty percent, “ineligible expenditure”, lacked documentation or deviated from the procurement procedure and other internal controls. While the ineligible expenditure does not directly imply corruption, it clearly showed fiduciary weaknesses and signalled that the budget was vulnerable to corruption (Haque & Nassif; 2021).
d. Getting the Right Quality Indicators and Assessing the Overall Impact of Results

The results of these aid programs were, ultimately, no more than the sum of their parts. Over the two decades, the government met hundreds of indicators with limited aggregate impact, despite a net spend of $75 billion USD in ODA.

However, looking at the aggregate impact over time, it appears the reforms were costly and had little effect. For instance, the Country Policy and Institutional Assessments (CPIA) of the World Bank, which assesses each country's quality of institutions and current policies, best summarises the aggregate impact. The structural policies cluster and transparency, accountability, and anti-corruption cluster scores for Afghanistan in 2020 were precisely the same as for 2007, at 2.3 and 2.0, respectively. Despite the billions of dollars that were invested, improvements in other clusters were marginal. Similarly, several other third-party assessments, such as Public Expenditure and Financial Accountability (PEFA), World Governance Indicators, and Transparency International Corruption Perception Survey, noted low outcomes. Insecurity, no doubt, was a constraint. However, it is also apparent that policymakers did not understand the context and the transmission mechanism for institutional change required for defining the right quality indicators.
## Table 3

**Afghanistan Country Policy and Institutional Assessment (CPIA)**

<table>
<thead>
<tr>
<th>CPIA Clusters (score: 1=low to 6=high)</th>
<th>2007</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public sector management and institutions cluster average</td>
<td>2.2</td>
<td>2.6</td>
</tr>
<tr>
<td>Structural policies cluster average</td>
<td>2.3</td>
<td>2.3</td>
</tr>
<tr>
<td>Transparency, accountability, and corruption in the public sector rating</td>
<td>2.0</td>
<td>2.0</td>
</tr>
<tr>
<td>Policies for social inclusion/equity cluster average</td>
<td>2.3</td>
<td>2.7</td>
</tr>
</tbody>
</table>

Note. Source: The World Bank development data indicators.
References


About Us


Consortium members include: Conciliation Resources, Centre for Trust, Peace and Social Relations (CTPSR) at Coventry University, Dialectiq, Edinburgh Law School, International IDEA, LSE Conflict and Civicness Research Group, LSE Middle East Centre, Queens University Belfast, University of St Andrews, University of Stirling, and the World Peace Foundation at Tufts University.

PeaceRep is funded by the Foreign, Commonwealth and Development Office (FCDO), UK.