



# SYNTHESIS REVIEW OF DFAT- FUNDED MARKET SYSTEMS DEVELOPMENT INITIATIVES

Part II - annexes

## Abstract

This is a compendium of annexes for the report on the synthesis review of DFAT's Market Systems Development Initiatives

Bob Warner and Donna Loveridge

31 October 2019

## Contents

Abbreviations.....	4
Annex A: Review methodology and questions .....	6
Approach .....	6
Method.....	7
Review questions.....	8
Limitations and risks .....	8
Reporting and stakeholder engagement .....	10
Annex A - Appendix 1: Key features of MSD programs.....	11
Introduction/Background.....	11
Annex A - Appendix 2: List of persons consulted and interviewed .....	13
Annex B: The Government’s current aid priorities.....	15
Gender .....	15
Disability.....	16
Private sector engagement .....	16
Value for money .....	16
Policy engagement.....	17
Annex C: Changing market systems by addressing underlying causes and ensuring additionality .....	18
Addressing underlying causes .....	18
Additionality.....	18
Annex D: DFAT’s MSD Program Portfolio and Results .....	21
Countries .....	21
MSD programs and their characteristics .....	22
MSD Programs’ sectoral focus.....	24
Targeting of specific groups .....	24
Quantitative headline results .....	26
Examples of market systems changes.....	29
CAVAC I: Regulation of chemical pesticides in Cambodia .....	29
Small packaged silage in Pakistan .....	29
PRISMA: Influencing regulations in the Indonesian shallot sector .....	30
The VfM of MSD programs.....	31
Mid-term reviews’ and evaluations’ conclusions on value for money.....	31
Comparisons using DFAT’s internal assessments.....	32
Annex E: MSD Program M&E systems .....	34
Introduction.....	34

DFAT's and programs' information needs and uses .....	34
Information users and needs.....	35
Monitoring, evaluation, value for money assessments and reporting requirements .....	35
DFAT's M&E Standards.....	35
The DCED Standard for Results Measurement.....	36
Value for money requirements .....	37
The credibility of reported results .....	39
Outcomes focused .....	40
Quality assurance .....	41
Effectively used .....	44
Assessing the breadth and depth of outcomes .....	45
Headline indicators.....	46
Challenges with understanding headline results .....	46
Targets.....	48
Portfolio level .....	49
Market systems level .....	50
Defining and understanding systemic change.....	51
Assessing systemic change .....	54
Intervention level.....	55
Assessing sustainability of changes.....	57
Assessing value for money .....	59
VfM assessment purposes and approaches .....	60
Final observations - M&E for the MSD portfolio and programs .....	64
Addendum 1: Examples of different indicator definitions.....	67
Addendum 2: MSD program's approaches to assessing VfM.....	68
VfM assessment approaches.....	68
Other organisations' approaches to assessing VfM .....	70
Example 1: MADE Ghana .....	71
Example 2: Making Vegetable Markets Work .....	72
Example 3: Financial Sector Deepening Program, Kenya .....	72
Example 4: Samarth-Nepal Market Development Program.....	73
Alternative approaches to assessing VfM .....	74
Annex F: References and documents reviewed.....	75

## Abbreviations

AAER	Adopt-Adapt-Expand-Respond
ACIAR	Australian Centre for International Agricultural Research
ACR	Activity Completion Report
ADR	Aggregate Development Results
AFS	DFAT's Agricultural and Food Security Section
AIP	Aid Investment Plan
AIP-Rural	Australia-Indonesia Partnership for Rural Development Program
AMENCA	Australia Middle East NGO Cooperation Agreement Program
APG	Aid Programming Guide
AQC	Aid Quality Check
ARISA	Applied Research and Innovation Systems in Agriculture
AusABBA	Australia Balochistan Agribusiness Project
BEE	Business Enabling Environment
CAVAC	Cambodia Agricultural Value Chain
CoP	Community of Practice
CPIA	Country Policy and Institutional Assessment
DAC	Development Assistance Committee
DCED	Donor Committee for Enterprise Development
DFAT	Department of Foreign Affairs and Trade
DFID	Department for International Development
FAO	Food and Agriculture Organization
FWUC	Farm Water User Committee
GDP	Gross Domestic Product
GLAM	Global Learning for Adaptive Management Initiative
GREAT	Gender-Responsive Equitable Agriculture and Tourism Program
INGO	International Non-Government Organisation
LDP	Learning and Development Platform
LIFT	Livelihoods and Food Security Fund
LMIC	Lower Middle Income Country
M4P	Making Markets Work for the Poor
M&E	Monitoring and Evaluation
MDF	Market Development Facility
MELF	Monitoring, Evaluation and Learning Framework
MRM	Monitoring and Results Measurement
MSD	Market Systems Development
MTR	Mid-Term Review
NGO	Non-Government Organisation
ODA	Official Development Assistance
PHAMA Plus	Pacific Horticultural and Agricultural Market Access Plus Program
PPA	Partner Performance Assessment
PRISMA	Promoting Rural Incomes through Support for Markets in Agriculture Program
SAFIRA	Strengthening Access to Finance in Rural Agriculture
SBC	Social Behaviour Change
SDC	Swiss Agency for Development and Cooperation
SMART	Specific, Measurable, Achievable, Realistic and Timebound
TIRTA	Tertiary Irrigation Technical Assistance
TOMAK	To'os Ba Moris Diak Program
TOR	Terms of Reference

UMIC	Upper Middle Income Country
VfM	Value for Money
WEE	Women's Economic Empowerment

## Annex A: Review methodology and questions

This review of DFAT's MSD programs was commissioned to help staff involved in current MSD programs, or who are considering options for new or follow up investments where MSD approaches may be appropriate. It is also intended to inform a broader audience within and outside DFAT, including DFAT's implementing partners and the broader community of practice and other donors.

The **primary purpose** of this review is to assess the extent to which DFAT's MSD programs remain relevant, offer value for money (in terms of efficiency and effectiveness) and are meeting DFAT's expectations for impacts at the household and market system level. The review is also expected to identify improvements and innovations that programs are making, or could be making, to further strengthen the practical application of the MSD 'model' - particularly with respect to targeting and benefitting the poorest segments of the communities where programs operate, and to supporting DFAT's policy engagement agenda.

The **primary audience** is DFAT staff involved in current MSD programs, or considering options for new or follow-up investments where MSD approaches may be appropriate; while a secondary audience is DFAT's implementing partners, including managing contractors, NGOs, governments and private sector organisations; as well as the global MSD community of practice, including other donors.

Specific **objectives** for the review are to:

1. assess the alignment of DFAT's current MSD programs with Australia's aid and foreign policy objectives;
2. examine the effectiveness, efficiency, value for money and sustainability of Australian-funded MSD investments, and evaluate reported and likely impacts, including likely scale, and any evidence of 'systemic change';
3. evaluate whether MSD programs (a) explicitly target, and (b) succeed in benefitting the poor, women and other specific target groups (such as smallholder farmers);
4. Comment on whether the country context (for example, dynamic middle income versus 'thin markets' or conflict/fragility) significantly affects the suitability of the MSD approach and the level of impact that can be achieved.
5. Recommend any measures to further strengthen the development benefits from DFAT's MSD programs, including explicit linkage to other types of program.

The review's purpose is forward looking. It has both summative and formative elements, but the emphasis is on the formative, to ensure it provides useful inputs to DFAT staff considering options for new or follow-up investments using the MSD approach.

To achieve the objectives, the review also considered lessons from the broader universe of MSD programs, to complement the analyses of DFAT's portfolio, particularly with respect to questions about the applicability of the MSD approach in different contexts, the scope for adaptation, and how the challenges of results measurement can be addressed.

### Approach

The approach adopted by the review took into account that it:

- Is a synthesis review covering 11 MSD programs (current and past phases) and a broad set of questions within a relatively short time frame, and therefore takes a pragmatic approach.
- Must be useful for DFAT's needs that relate to future decisions about the use of MSD approaches.

- Incorporates consideration of several different concepts, for many of which there is no single definition.

As such, the review has had to present definitions that can be used to frame the analysis and inform conclusions, and to make these accessible to the key audiences. The key features of the approach are:

- The scope of and methods used for the review are tailored to the allocated time and resourcing. In particular, the review will draw primarily on existing DFAT documentation for each of the programs, principally investment designs, formative and summative reviews and evaluations, and annual aid quality checks. To a lesser extent, the review drew on the knowledge, experience and perceptions of key informants such as DFAT staff and people involved in the implementation, oversight and evaluation of MSD programs.
- The review adopted a 'utilisation focused approach' —based on the belief that a review should be judged by its usefulness to its intended users, including for AFS's purposes of drawing on it to prepare a policy paper for DFAT senior management. To increase the utility of the review, the review team consulted regularly with the DFAT staff who commissioned the review through inception to finalisation, and sought ways to engage other key individuals to the extent possible at key milestones to discuss preliminary findings. The Reference Group also made an important contribution to the utility of the review.
- The review examined questions concerning the usefulness and relevance of the MSD 'model' primarily as it relates to the effectiveness and impact of programs. The MSD model has been characterised in slightly different ways by practitioners, DFAT and other donors. The review therefore needed a strong foundation in the theories underpinning the MSD approach, bonded to an understanding of practical experience and contextual appreciation.
- The review was independent, since the consultants undertaking the review are independent of DFAT, and have no *direct* involvement in implementation of any of the programs in the growth portfolio.<sup>1</sup>

## Method

The review relied heavily on analysis of existing program documents and literature relating the MSD programs and approaches. This was complemented by consultations with the review managers in DFAT and selected interviews with key informants such as DFAT program managers (past and current), program implementing organisations, members of technical advisory groups, and members of the broader MSD community that have been involved in evaluating MSD programs and the overall MSD approach.

The team also tested ideas and findings with the Reference Group.

Data was collected and analysed in three stages: through a systematic review of program documentation to extract information related to review questions; collection of information through interviews of selected key informants, together with a workshop to discuss emerging findings summarised in an aide-memoire; and collation and analysis of data collected during the first two stages, including refinement of assessment frameworks. The review used this information to develop a summary characterisation of key features of MSD programs. While this framework was not considered to be sufficiently tested to 'rate' programs in DFAT's portfolio, it was used to

---

<sup>1</sup> The consultants are members of the Advisory Group for the Market Development Facility, one of the programs included in the review. This Group is independent of the Facility and reports to DFAT.

organise thinking about the programs and their alignment with 'good' MSD practice. The framework is presented in **Appendix A.1** to this Annex.

Key informants for the review included:

- **DFAT portfolio/program managers**: staff who were currently managing MSD programs, or who have managed them in the recent past, or have had responsibility for portfolios that have included MSD programs;
- **DFAT sectoral and technical advisors**: staff involved in an advisory and oversight role with respect to private sector engagement, agriculture and rural development, gender, and monitoring and evaluation;
- **Delivery team managers and members/partners**: managing contractors and NGOs involved in delivery of MSD programs; and
- **External advisors and informed third parties**: members of program technical advisory groups, experts from the global MSD community, and those involved in exploring the role and contribution of adaptive/learning programs, and how such programs monitor and report on results and use results for management, accountability and learning.

**Appendix A.2 to this annex** lists the key informants that were interviewed for the review.

The review team had a day of consultations with AFS review managers, the Reference Group, and other key DFAT staff (those with experience/interest in the work of the review) to discuss emerging findings presented in an aide-memoire. This workshop tested preliminary findings and interpretations with respect to the key questions. **Stage 3** consisted of collating the data from Stage 1 and 2 and analysing it in its entirety to answer the key questions. The draft reported was discussed at a further meeting with key DFAT staff.

#### Review questions

The terms of reference for the review laid out a clear set of questions to be addressed. Table 1 presents these questions, noting the specific review objective to which they seem to be relevant.

#### Limitations and risks

This review faced some considerable limitations, and risks that it might not be able to get sufficient information to draw defensible and well-substantiated conclusions with respect to all of the key questions. The main limitations are as follows.

- **Time and resources**: the rigour of the data gathering and analysis processes for this review was constrained by the time available. In particular, the number of programs (ten) and country contexts (fourteen) to be considered posed constraints on the depth of analysis that could be achieved through document review and interviews. While the synthesis review was not responsible for evaluating any of the programs directly, it had to make judgments of performance against DAC criteria, and to answer questions that involved considerable understanding of how programs go about particular tasks (especially results monitoring and reporting). The review was heavily dependent on the quality, transparency and depth of program reporting. To some extent, where data was not available in existing documentation, the review had to draw on the perspectives of key informants – which meant understanding the lenses through which informants view the MSD approach and individual programs. This caveat is particularly important given that MSD programs are by intention and necessity adaptive, and revise strategies based on program learning under implementation. This means that designs – and often reviews – may fail to capture the detail of analysis and thinking that must underpin such a program.



Table A.1 Review questions

Review criteria	Key Questions
Relevance	To what extent are current MSD approaches consistent with GoA's aid priorities? Is the additionality principle being applied to partnership interventions? Do investments support policy dialogue and help achieve policy influence? Have partner government's accepted MSD as a valid approach in the aid program? Are MSD programs creating effective links with other Australian (or other donor) aid programs?
Relevance Impact	Are programs explicitly attempting to reach specific target groups, especially the poor (and the poorest), women, smallholders, people with disability?
Effectiveness Impact Sustainability Systemic change	What evidence is available on achievement of sustainability and systemic change? Can conclusions be drawn about whether the MSD model is itself contributing to such achievements? Is there evidence that specific groups – the poor, women etc are benefitting? Do MRM systems and reporting frameworks provide credible and defensible data to support internal adaptive management and DFAT's accountability and communication needs? Does adherence to the DCED standard assure DFAT that M&E systems meet internal adaptive management and DFAT's accountability and communication needs?
Value for money	Do Australia's MSD programs represent value for money? What is the evidence on effectiveness and efficiency
Relevance	Does the portfolio of MSD programs indicate that MSD approaches are adaptable to very different country contexts? Does country context affect the suitability of the MSD approach and the level of impact that can be achieved?
Effectiveness	Have any significant challenges arisen in implementing Australia's MSD programs? Have they been addressed effectively?
Effectiveness	Can common elements be identified across the MSD portfolio contributing to relative success or failure? What conclusions can be drawn on the strengths of Australia's MSD investments relative to global or 'effective aid' benchmarks Is there a minimum level of resourcing that is needed to enable achievement of intended outcomes?
Effectiveness	Are there existing or potential innovations that might strengthen the practical application of the MSD model?

- **Judgments:** the review primarily involve rapid qualitative methods of inquiry (e.g. document review and analysis, and semi-structured qualitative interviews). When gathering and analysing the data, the review team needed to interpret the perspectives of the informants and the underlying purposes of specific pieces of documentation. (This recognises that some of the key documentary sources of information, such as Aid Quality Checks (AQC's) have limited space to present highly nuanced assessments, and have to select information to clearly underpin their judgments.
- **Access to information and informants:** The review did not visit any programs or activity sites, and it relied heavily on telephone interviews with key informants. The review may have been able to access all relevant documentation if it is not readily available on program websites or held by or accessible to Canberra based DFAT staff. This at times made it difficult to triangulate data from the initial document review and interviews, since the scope for more sophisticated validation of conclusions was extremely limited.
- **Conceptual clarity and benchmarking:** while there is an extensive literature on Market Systems Development (and its precursor, the Markets for the Poor (M4P) approach), the conceptualisation of the approach does not lend itself easily to benchmarking – within and across programs. This means, among other things, that it is not always easy to attribute characteristics and performance of programs to the *approach* or to other factors in the environment, program design or program implementation.

- It has also meant that the review has had to adopt a set of definitions and benchmarks to use in addressing some of the key evaluation questions, recognising that there may well be scope to context some of the choices implicit in these framings.

### Reporting and stakeholder engagement

Reporting from the review involved three products

- **An aide-memoire** presenting initial findings and insights based on the document review – due 9 August 2019. This was be discussed at a workshop/meeting with the review reference group
- **A draft review report** presenting a clear and cogent summary of findings, based on a balanced analysis of the available evidence and identification of relevant issues and recommendations to DFAT – delivered 23 September 2019, and discussed with key DFAT staff on 10 October 2019.
- **A final review report** which addressed comments received from DFAT on the draft report – due within 14 days of receiving final DFAT comments.

## Annex A - Appendix 1: Key features of MSD programs

### Introduction/Background

There are many existing lengthy guidelines on the MSD approach. There are variations in organisations' vision of market systems development and its objectives.

Our intention, firstly, was to see if we could boil-down lengthy guides to a one-page presentation of key features of MSD programs. This sounds easier than it was – we sent a first draft out to a small number of people for feedback, which confirmed again there are many nuanced perspectives. Although, the intention was not to re-invent the wheel, we may have done. The features captured in the one-pager are not unique to MSD programs. However, their distinction may only be that they are applied to economic systems, rather than health or education.

Our second intention was to use the one-pager to help understand DFAT's MSD program portfolio and the extent to which different features featured. This provided impossible – a lot more documentation and interviews would be needed. It has been fruitful in helping to think through the portfolio, which elements appear to feature more strongly than others; and how the features interact on the ground.

Below, we summarise the overarching TOC of MSD programs, a short description of market actors and facilitators.

#### *MSD programs' overarching theory of change*

<b>Problem:</b> Barriers constrain many poor and disadvantaged groups from participating and benefiting in economic markets
<b>Purpose:</b> To sustainably improve the well-being of large numbers of poor and disadvantaged groups
<b>Strategy:</b> To make markets more inclusive, by identifying root causes, working with and through market actors to develop and implement market-based, technically sound and politically feasible solutions to address constraints

**Who are the market actors?** Market actors are organisations and individuals that participate in market functions and/or set the rules. These include: governments; businesses – large, medium, small, micro; business / industry associations and research institutes;

**Who is the facilitator?** Donors contract 'prime' organisations, such as program management companies or NGOs, to deliver programs. We have called all these organisations facilitators, but there may be a chain of facilitators. For example:

- Prime's may set up a consortium with other organisations to implement the program. There will be a contractual relationship between the prime and each member of the consortia. Primes may sub-contract or sub-grant the implementation of some or all interventions or activities to other organisations e.g. local or international project management companies, consultancies or NGOs.
- Primes or their sub-contractor and sub-grantees sign agreements with market actors. Agreements may either provide a grant to the market actor for them to pay for services and products covered by the agreement directly or the prime, sub-contractors and sub-grantees may pay for services and/or products directly. At times, agreements may be between the donor and market actors, while the prime has responsibility for managing the agreement with the market actor.

The following table summarises the key features that appear to be evident in MSD programs.

<b>Donors</b>
<b>Managing adaptively</b>
<u>Adaptive design</u> : DFAT managers designs, procure, fund and evaluate programs in a way that enables programs to manage adaptively.
<u>Adaptive management</u> : DFAT managers manage programs (through appropriate structures, contracts, systems and processes) that enables facilitators' adaptive programming and delivery.
<b>Facilitating systems change</b>
<u>Systems thinking</u> : DFAT managers manage in a way that encourage programs to focus on changing market systems, the raison d'être of market systems development programs, to sustainably improve the well-being of large numbers of poor and disadvantaged groups.
<b>Facilitators</b>
<b>Analysing market systems</b>
Systems thinking: Programs recognise that market actors (e.g. businesses, government, farmers and their inter-relationships), functions (e.g. policies and regulations) and rules (formal and informal) are part of an interacting system. Programs also articulate the boundaries (e.g. sub-sector or cross-sector focused) of the market systems they are seeking to change.
<u>Market analysis</u> : Programs analyse markets, identifying constraints to large numbers of poor and disadvantaged persons benefiting from markets; analysing the economic and financial incentives for all actors; and the feasibility of the program addressing constraints and achieving sustainable and inclusive outcomes.
<u>Systems monitoring and evaluation</u> : Programs monitor and evaluate changes to market systems e.g. functions, rules, incentives, relationships, as a result of their support to market actors to address constraints. This includes analysing the extent to which market systems are becoming more inclusive in sustainable ways; and the effectiveness of the tactics used and partnerships formed.
<b>Systems thinking</b>
<u>Portfolio approach</u> : Programs thoughtfully use analysis to create and manage a portfolio of interventions, bounded by a clear set of articulated principles, to make markets more inclusive. The portfolio enables trial - error, success - failure, and mitigates risks e.g. execution, outcomes achieved, sustainability and external impacts.
<u>Tactics</u> : Informed by analysis (including market, economic and financial, political economy analysis) programs use a range of 'technically sound, politically feasible' tactics, relevant to the program's capabilities (resources, skills, influence).
<u>Systems action</u> : Programs pro-actively seek to move beyond individual interventions, which aim to directly benefit a limited number of market actors, implementing strategies so more market actors' benefit. Market actors include poor and disadvantaged groups.
<b>Working with and through market actors</b>
<u>Facilitating change</u> : Programs nearly always supports (e.g. knowledge, networks, funds) market actors to address constraints, versus directly delivering goods and services to end-beneficiaries. The program is not a market actor.
Assessments of additionality: Programs critically assess the additionality of their proposed support to market actors through the program cycle. That is, whether market actors would done it anyway, now or soon.
<u>Managing for fairness and transparency</u> : Programs assess the benefits and risks of providing a business with an advantage. Programs periodically assess how market changes impact on: favouring one market actor over another in a context of changing (diminishing) risks; and support to first movers of an inclusive business model is expected to lower the risks for the next followers.
<b>Managing adaptively</b>
<u>Adaptive management</u> : Prime organisations, and their consortium members, manage programs and provide enabling structures, systems and processes, people capabilities, and relationship management that enables programs to manage and deliver adaptively.
<u>Adaptive management</u> : Program teams use deliberate, regular and structured process of stepping back to reflect, conducting more in depth and focused analysis, and bringing in critical friends to help review and set new directions.
<u>Adaptive delivery</u> : Front line teams think on their feet, apply curiosity, evidence, emotional intelligence and instinct to learn (including sharing with others), adapt, and make decisions in their day to day work continually making best-guesses, then testing and correcting and employing 'everyday PEA'. Adapting may include changing, improving, stopping, starting, decreasing.

## Annex A - Appendix 2: List of persons consulted and interviewed

<b>Name</b>	<b>Position</b>	<b>Organisation</b>
Alwyn Chilver	Contractor Representative, MDF	Palladium
Aly Miehlebradt	Consultant	Consultant
Andrew Dollimore	Middle East Desk	DFAT
Annemarie Reerink	Gender Specialist	DFAT
Cass Grant	Manager, Strongim Bisnis and PHAMA Plus in Solomon Islands	DFAT, Solomon Islands
Chris Tinning	Chief Economist	DFAT
Corin Mitchell	Director, Asia Pacific	Adam Smith International
Curtis Slover	Agriculture Value Chain & Rural Finance Program Officer	LIFT
Dan Wood	Ex-Dili Post	DFAT
David Swete-Kelly	Design Leader GREAT and TOMAK Advisory Group, MDF 1	Consultant
Di Barr	Team Leader, PHAMA Plus	AECOM
Fiona Lynn	Director, AFS	DFAT
Frank Maiolo	Contractor Representative, PHAMA Plus	AECOM
Gary Ellem	Team Leader, CAVAC II	Cardno Emerging Markets
Goetz Ebbecke	Team Leader, PRISMA 1 and 2	Palladium
Harald Kreuscher	Senior Program Officer	LIFT
James Riturban	Portfolio Manager – Sri Lanka, Pakistan and the Middle East	Oxfam Australia
Joe Manteit	MDF Program Manager	Agricultural Development and Food Security Section, DFAT
John Farghar	Design – MTR – AMENCA3	Consultant
Julie Delforce	Senior Sector Specialist	Agricultural Development and Food Security Section, DFAT
Khaled Khan	PRISMA	Palladium
Kirsten Hawke	Design	DFAT
Laura Baines	Program Officer	CARE
Hasan Abdeljabbar	AMENCA Project Manager	CARE
Anan Kittaneh	AMENCA	CARE
Luke Arnold	Deputy Head of Mission	DFAT
Ma'ake Komailevuka	Program Manager	DFAT, Suva
Mark Palu	Director	DFAT
Matt Harding	Director, Office of the Pacific	DFAT
Mike Albu	Director, BEAM Exchange	DCED
Naomi Cook	Previously Program Manager for PRISMA, Indonesia	DFAT
Neal Forster		ODE, DFAT
Paul Keogh	Team Leader, MDF	Palladium
Peter Roggekamp	Team Leader, 3i (Previous Team Leader CAVAC I)	Palladium (previously, Cardno Emerging Markets)
Peter Wilson	Assistant Director, former MDF Program Manager	DFAT

Phil Harman	Team Leader, GREAT	CowaterSogema International Inc
Phil Montgomery	Team Leader, Strongim Bisnis	Adam Smith International
R. Sivasuthan	Senior Program Manager	DFAT, Colombo
Richard Holloway	TOMAK Team Leader	Adam Smith International
Rob Hitchens	PRISMA Advisory Group / MDF Market Advisor	Consultant
Sally-Ann Thomas	Performance Specialist	DFAT
Simon Cann-Evans	Contracting and Aid Management	DFAT
Sophie McKinnon	First Secretary Development Cooperation, Suva	DFAT
Steve Hogg	CAVAC Strategic Review Panel	Consultant
Tim Gill	Sector Specialist	Agricultural Development and Food Security Section, DFAT
Tristan Armstrong	Sector Specialist	Agricultural Development and Food Security Section, (AFS) DFAT

## Annex B: The Government's current aid priorities

The 2017 Foreign Policy White Paper presents the most recent articulation of the Government's objectives and priorities for aid. It confirms the contention in the 2014 statement on aid policy (DFAT 2014a) that Australia's development assistance is focused in the Indo-Pacific Region and promotes the national interest by contributing to sustainable growth and poverty reduction, and is an important aspect of Australia's contribution to global prosperity and stability (Australian Government, 2017).

The White Paper highlights the need to work with partners to encourage private sector-led growth, poverty reduction and inclusion. It acknowledges that growth alone will not guarantee prosperity and stability, and expresses an intention to focus on poverty reduction and promoting the participation in the economy of the most disadvantaged, including women and girls and people living with disabilities. The Paper also presents four tests to be applied when allocating development assistance, focusing on the national interest, promotion of inclusive growth and poverty reduction, value addition and leveraging partner funding, and delivery of results and value for money.

The firm linkage of development assistance to the national interest has meant that over time, and as the Government's perception of Australia's strategic and commercial interests has changed, expectations of the aid program have changed too. Some MSD programs started implementation before the 2014 statement, and have had to evolve over time to better align with the priorities presented in that document. But all programs are adjusting to evolving expectations and perceptions of their potential contribution.

DFAT's strategy for Australia's aid investments in agriculture, fisheries and water seeks to increase incomes of **poor people** (DFAT, 2015c). Its strategy for Australia's aid investments in private sector development highlights support for specific markets and sectors that provide a greater potential to reduce poverty and increase incomes of the poor; and promoting business activities that bring the poor into the market as producers, consumers and employees (DFAT 2015a).

Some areas where priorities have strengthened, and where more guidance has been provided include gender, private sector engagement, value for money and policy engagement. The current status of guidance on these issues, where most relevant to the MSD portfolio is summarised below.

### Gender

DFAT's 2016 Gender equality and women's empowerment strategy established gender equality and women's empowerment for foreign policy, economic diplomacy and the development program, and identified three priorities for all of the work of the Department: enhancing women's voice in decision making, leadership and peace building; promoting women's economic empowerment; and ending violence against women and girls (DFAT 2016).

The strategy established a specific performance target – that 80 per cent of investments, regardless of their primary objectives, must also effectively address gender issues in their implementation.

Among other things, this has led to programs being required to develop and implement program specific gender strategies, and to an assessment of performance on gender becoming an important part of program reviews and annual Aid Quality Checks.

## Disability

DFAT's strategy for strengthening disability-inclusive development in Australia's aid program emphasises: enhancing participation and empowerment of people with disabilities, as contributors, leaders and decision makers in community, government and the private sector; reducing poverty among people with disabilities; and improving equality for people with disabilities in all areas of public life, including service provision, education and employment.

## Private sector engagement

The 2014 statement on aid policy established private sector development as one of the two pillars of Australia's development policy. The 2015 Strategy for Australia's investments in private sector development provided guidance on how programs could improve the growth and inclusiveness of the private sector in the countries where aid is provided (DFAT 2015a). It proposed that aid would focus on: building better business and investment environments, supporting growth in specific markets; and maximizing the development impact of individual businesses.

Alongside the intention to work on the growth and inclusiveness of the private sector, DFAT has also articulated an operational framework for working with the private sector to promote growth, reduce poverty and transition from aid relationships to economic partnerships with developing countries (DFAT, 2019). This framework articulates objectives and opportunities for various modes of working with Australian and developing country private sectors, and also works through some of the risks that might be encountered.

Amongst other things, the framework identified principles to ensure that engaging with the private sector in this way did not create market distortions:

- **Additionality:** Public funds will not finance activities that a commercial enterprise of a financier would have funded without intervention;
- **Fairness:** Collaboration with DFAT will not provide one business with an unfair advantage over its competitors; and
- **Transparency:** Information shared with the private sector should be publicly available and funding opportunities should be shared with all relevant stakeholders. Transparent processes will guide all decisions to engage with the private sector.

Given that the framework cites examples from DFAT's MSD portfolio as showing how the aid program works with the private sector, it is clear that these principles are meant to apply to MSD programs.

## Value for money

DFAT set benchmarks for value for money (VfM) in its 2014 performance framework for the aid program (DFAT 2014b), namely that the Department would deliver high standards of VfM in at least 85 per cent of aid investments, and that investments not delivering VfM would be required to improve within 1 year or be cancelled.

DFAT has published a set of eight principles to guide decision making and maximise the impact of investments. The principles are: cost consciousness; encouraging competition; evidence based decision making; proportionality; performance and risk management; results focus; experimentation and innovation; and accountability and transparency.

While these principles indicate how DFAT will *pursue* VfM, they do not provide much guidance on how the VfM for an individual investment is to be assessed. As discussed in chapter 3 in the main



report, this has proved to be something of a challenge for stakeholders and managers of MSD programs.

### Policy engagement

The 2014 statement on aid policy laid out the rationale for Australia's aid program to support developing country partners to improve the policy and institutional frameworks that provide the foundations for economic growth, private investment, trade and the delivery of public goods and services. And the Foreign Policy White Paper made it clear that aid would be deployed to 'support better state capability to improve governance and reduce poverty and inequality. Across country programs, DFAT is increasingly expressing an intention to deliver policy knowledge and advice as part of a shift away from a traditional donor-recipient aid model towards economic partnerships, and to do this in a way that builds on and understanding of the political economy of change.

In the last couple of years, country programs have expressed an increasing expectation that MSD programs would support this intent, and program designs are being developed to build in this expectation. There is also an apparent shift to require programs working in the market and private sector space to focus increasingly on competitiveness. This shift also intersects with elements of DFAT's economic and commercial diplomacy agenda, particularly in the agenda's pillars related to business facilitation and advocacy for the global rules based trade system.

## Annex C: Changing market systems by addressing underlying causes and ensuring additionality

### Addressing underlying causes

Market Systems Development (MSD) programs aim to bring about changes in the underlying determinants of the performance of market systems to improve their functionality and inclusivity.

They do this using a range of instruments which are chosen after detailed analysis - or ongoing experimentation - because they are considered to address or respond to a fundamental market problem, such as a market failure, state failure or an adverse norm or belief – or a particular manifestation of such a problem. In other words, they are intended to come to grips with problems with the institutions or rules of the game that underpin market exchange. Programs have at their disposal a range of instruments, and make judgments as to which of these instruments is likely to offer the most effective and efficient way of tackling the problem or a particular manifestation, if they can in fact address the problem or manifestation

As a report prepared for DFAT on how the Market Development Facility (DFAT 2018a) approaches systemic change points out:

*If a country's market system does not work well, sustainable inclusive growth is very hard to achieve. But the functionality of markets depends very much on the presence and impact of formal and informal 'rules of the game'. Pro-market 'rules of the game' are needed to address problems of trust and information, to make it possible to make transactions and contracts that span family, tribal, ethnic, religious and geographic boundaries, and that can operate over time. Without such rules, the scope for economic activity is limited, risk-taking is discouraged and investment – crucial for growth - will be constrained. And if the prevailing norms do not allow for economic participation of all people, regardless of gender, age, religion or ethnicity, growth will not be inclusive, nor will it make best use of the nation's human resources.*

### Additionality

MSD programs often work with a selection of businesses to introduce innovations, practices and relationships that are new in a particular market context. The aim is not to assist in the development of those businesses. Rather it is a means of addressing underlying reasons why commercially viable inclusive opportunities are not being sought out and implemented as a result of the everyday operations of market forces. MSD programs aim to shift understanding of the commercial viability of more inclusive ways of doing business, or of adopting innovations that significantly benefit poor or excluded people. MSD programs do not offer long-term subsidisation to affect a long-term shift in the commercial profitability of operating in a more inclusive way.

While MSD programs are not concerned with developing businesses per se, the frameworks that donors are developing for engaging with the private sector, and in particular for providing grants and concessional finance to for-profit businesses, are generally applicable to MSD programs.<sup>2</sup>

Such frameworks typically spell out conditions for using public funds to subsidise the activities of private or commercially driven entities. These conditions derive from conventional public finance policy, and ideally require that the intervention should:

---

<sup>2</sup> Examples of such frameworks are the DCED paper on additionality (Heinrich, 2014) and DFID's paper on provision of grants and concessional financing to private firms (DFID, 2015)

- have a clear development and economic rationale: it should be addressing a market or institutional failure, or have a clear equity or distributional objective;
- demonstrate additionality – without the intervention the outcomes would not occur, or would be substantially delayed;
- bring about sustainable changes – the continuation of development outcomes should not be dependent on continued donor support, from whatever source; and
- target the underlying cause of poor market performance as directly as possible and minimise market distortions.

The conditions will also indicate that the reputational and fiduciary risks to the donor need to be managed, and that the for-profit entity should be able to manage the risks of working with a donor program (as well as the commercial risk associated with the activity).

DFAT's 2019 operational framework for private sector engagement in the aid program (DFAT, 2019) contains most of these elements, but on its own provides little guidance for stakeholders as to how to implement them. Providing examples of the types of market or institutional failure that would warrant subsidisation, and the kinds of market distortions that programs should be wary of creating would be helpful to MSD programs to determine the most appropriate instrument to use to address an identified problem.

As a general principle, where a market, policy or institutional failure is impeding efficient and equitable market functioning, it is preferable to implement a market wide solution. More limited interventions may have a role, depending on the nature of the problem being tackled, the lack or inaccessibility of market wide options, or where such interventions help get a better understanding of a problem and enable better-informed choices about potential solutions.

Limited interventions might be warranted where the markets for information do not allow businesses to properly assess the risks or costs of innovations with uncertain outcomes (the 'first mover failure'), or where there are development externalities or missing markets. But such business-specific subsidies have to be chosen with care, recognising that a market wide intervention is likely to be preferable in principle, and, as with all considerations of externalities, the benefits of pursuing the externality have to be weighed up against the costs associated with the instrument chosen to pursue it.

Most MSD programs operate in markets subject to a wide range of distortions, and it is often difficult to be confident that an intervention is unambiguously improving efficiency and equity. Therefore frameworks to guide program decisions and that place a strong emphasis on additionality are important.

- Additionality may be financial or non-financial: where co-investors are not able to mobilise finance provided by the intervention from another source, or where the intervention adds value beyond the provision of finance. In either case, programs need to be able to demonstrate that the inputs provided by the intervention would not otherwise be accessible

Largely because of the way in which programs interpret compliance requirements related to donor financial management and procurement legislation, regulation and guidelines, MSD programs are inclined to work with larger, well-capitalised businesses that can meet due diligence requirements. This may increase the market power of already dominant businesses, particularly in some the small island economies with shallow markets and limited private sector maturity and depth. In large and small countries commercial and political power are often closely inter-twined. In small economies with shallow markets the prospect of other businesses copying an innovation, and hence reducing

the initial market advantage of the recipient business, may be quite limited. Further, the strategies that businesses have adopted to deal with very adverse physical and business environment conditions, and that households have implemented over long periods of time to deal with uncertainty and limited market access, may well involve quasi-monopolistic - or monopsonist - behaviours which programs will not erode in the short to medium term. In these circumstances, programs will have to make a judgment as to whether the gains in inclusivity and market functionality outweigh the costs of possibly entrenching market dominance.

The literature on MSD programming and private sector engagement provides useful discussions of these challenges and offers suggestions on ways of managing them. The Donor Committee for Enterprise Development (Tewes-Gradi, Blomberg and Scholl, 2018) presents some practical suggestions for minimising risks of creating market distortions, recognising that incentives within donor agencies may not always align with the demands of minimising these risks. The paper points out that the program managers have an incentive to 'produce compelling narratives of individual companies and their contributions' which may create pressures to support companies for longer than absolutely needed, or to place less emphasis on efforts to encourage crowding in or other ways of mitigating the potential to distort markets.

The MSD approach explicitly recognises some of these risks. It focuses on changing market systems, not private sector development, and in some of the prevalent conceptualisations of systemic change emphasises the role of copying and crowding in to spread innovations that improve the inclusivity of market functioning<sup>3</sup>.

It is important for programs to be very clear about the objective being pursued, and that the intervention targets that problem as directly as possible. Without clear targeting, there is a risk of waste (that is not making most efficient use of the resources within a program or intervention), or of creating market distortions by reinforcing or creating market power at the expense of other players or of consumers. Programs need to develop frameworks for working through these situations, achieving clarity with respect to the underlying reasons why markets operate the way that they do, and spelling out reasons for choosing particular intervention instruments and the risks associated with their use. For this to happen, DFAT needs to provide stronger guidance, and spell out its expectations as to how programs should address these challenges. It is becoming increasingly clear that in most of the countries where DFAT's MSD programs operate, programs will have to work with a broader range of partners and instruments besides co-investments, including collaboration with governments and other types of programs that address pervasive market distortions and weaknesses in the institutional underpinnings of market transactions.<sup>4</sup>

---

<sup>3</sup> In fact, constraints to the diffusion of information may well be one of the areas of 'underlying causes' of poor market performance that MSD programs have to tackle. In shallow markets in fragmented economies, programs may have to find strategies to facilitate diffusion of information about successful innovations. And in larger markets, they may have to find a way around the disincentives created by poorly designed and implemented government extension and input provision programs.

<sup>4</sup> The recent ex-post evaluation of the Samarth-Nepal Market Development Programme recommended that MSD programs working in fragmented economies with shallow markets should engage public sector stakeholders at an early stage in the program, identify entry points in the enabling environment and actively collaborate with existing donor and public-sector initiatives dealing with broader enabling environment issues (e-Pact, 2019).

## Annex D: DFAT's MSD Program Portfolio and Results

This annex presents information on various aspects of DFAT's MSD program portfolio, including: country and sector coverage, targeted beneficiaries, program 'headline' quantitative results, examples of market systems change and value for money assessment judgements.

### Countries

The programs covered by this review are currently implemented in fourteen countries. These countries span a wide range of economic, political and social characteristics, including:

- The world's fourth largest country by population (Indonesia) as well as some of the world's smallest, such as Tonga, Vanuatu and Fiji;
- Lower middle income and upper middle-income countries;
- Countries assessed to have relatively strong and weak government capabilities, including some countries that are considered to be fragile states (e.g. Myanmar, Palestinian Territories);
- Countries assessed to have reasonable and quite poor business enabling environments.
- The largest economy, Indonesia, (measured by Gross Domestic Product, GDP) is 2500 times larger than the smallest, Tonga.

Table D.1 presents some key characteristics of these countries. It should also be noted that some of these countries are highly fragmented collections of islands with high costs of internal and external connectivity with rural sectors containing high numbers of quasi-subsistence farming families.

*Table D1: Characteristics of countries covered by the portfolio*

Country	Population	GDP	Income status	CPIA Rating	Doing Business Ranking	World Bank Fragile state?
	million	US\$ billion (2017)	2019	2017	1 to 190	
Cambodia	16.5	22.2	LMIC	3.36	138	
Fiji	0.9	5.1	UMIC	-	101	
Indonesia	270.7	1,015.5	LMIC	-	73	
Myanmar	54.3	69.3	LMIC	3.00	171	Yes
Palestinian Territories	5.2	14.5	LMIC	-	116	Yes
Pakistan	204.6	304.1	LMIC	3.24	136	
Papua New Guinea	8.6	21.1	LMIC	2.96	108	Yes
Samoa	0.2	0.9	UMIC	4	90	
Solomon Islands	0.6	1.3	LMIC	2.93	115	Yes
Sri Lanka	21	87.2	UMIC	-	100	
Timor-Leste	1.4	3.0	LMIC	2.85	178	Yes
Tonga	0.1	0.4	UMIC	3.49	91	
Vanuatu	0.3	0.9	LMIC	3.3	94	
Viet Nam	97.4	223.9	LMIC	3.69 <sup>a</sup>	69	

Notes GDP – Gross Domestic Product, LMIC – Lower Middle Income Country, UMIC – Upper Middle Income Country, CPIA – Country Policy and Institutional Rating (World Bank)

## MSD programs and their characteristics

Of the 11 programs included in the review<sup>5</sup>:

- The earliest commenced in 2009 and the most recent in 2019;
- They range in value from \$5m to over \$120m;
- Most are single country programs;
- Most combine interventions using an MSD approach with interventions based on non-MSD approaches, such as direct delivery;
- Most are managed by managing contractors although NGOs play significant roles as implementers (either as lead, consortium members or grantees) across about half of the portfolio;
- Most programs focus on Asia;
- Most focus on primarily on smallholder agriculture, although some are working in manufacturing, tourism and services; and
- Most focus on improving the incomes of target groups and women's economic empowerment, but some address other aspects of well-being such as nutrition.

Some programs can safely be located in the mainstream of MSD, implementing what could be called a 'classical' approach. Others are hybrids with elements using an MSD approach alongside elements using other approaches such as direct delivery or working with governments on policy and administration. Some hybrids are structured with components that are expected to complement each other in pursuing common outcomes. Others hybrids combine relatively unrelated components working to a common high level purpose in a common sector or location, under a single contractual umbrella.

**Table D.2** lists the programs reviewed, with an initial characterisation of which 'class' of MSD program – classical or hybrid – they seem to fall into.

### *Classical*

- The Market Development Facility (MDF) and Promoting Rural Income through Support for Markets in Agriculture (PRISMA) Phase 2 seem to be the most 'mainstream' of the programs in the portfolio. Their approach clearly builds on the framework used in pioneering MSD programs implemented by development agencies such as the Department for International Development (DFID) and the Swiss Agency for Development and Cooperation (SDC).
- The Australia-Indonesia Partnership Rural (AIP- Rural) was intended to follow an MSD approach across its four components. However, it appears only PRISMA really operated as a classical MSD program, with the other components eventually either experimenting to deliver a proof of concept (Tertiary Irrigation Technical Assistance – TIRTA and Strengthening Access to Finance in Rural Agriculture – SAFIRA) or using a different approach (Applied Research and Innovation Systems in Agriculture - ARISA).
- PRISMA 2 continues the work of the PRISMA component of AIP-Rural, but is expected to engage more on policy issues than its predecessor.

### *Hybrids*

- Pacific Horticulture and Agricultural Market Access Plus (PHAMA Plus) is explicitly a blend of two components, one aiming to address policy and administrative hurdles to agricultural export, and

---

<sup>5</sup> The terms of reference for the review also included the Investing in Infrastructure (3i) program, but closer investigation indicates that this is a direct delivery program, albeit working with and through the private sector using output-based financing. It has been dropped from the scope of the review.

the other using an MSD component aiming to help smallholders take advantage of the opportunities created by formal market access.

- AMENCA3 is a blended program, combining a component with some elements of an MSD approach with work to help improve the capacity of smallholders to be able to take advantage of emerging market opportunities. Other parts of the program use a direct delivery approach.
- Gender Responsive Equitable Agriculture and Tourism (GREAT) is a blended program, structured around three components, each with its own objective. One component focuses on partnering with businesses on inclusive business initiatives while the other two components partner with NGOs and government agencies for empowering local women and improving sector governance and policy.

Table D2: DFAT's portfolio of programs implementing an MSD approach

Name	Country	Period	Value (\$ million)	Type	Implementer
AIP-Rural (includes PRISMA I)	Indonesia	2010-19	112.0	Classical	MC plus AGA
AMENCA3	Palestinian Territories	2014-21	28.4	Hybrid	INGOs plus MC
CAVAC I	Cambodia	2010-15	46.0	Hybrid	MC
CAVAC II	Cambodia	2016-21	93.5	Hybrid	MC
GREAT	Viet Nam	2017-21	33.7	Hybrid	MC
LIFT	Myanmar	2009-24	42.3 <sup>6</sup>	Fund	UN
MDF I & II <sup>a</sup>	Fiji, PNG, Timor-Leste, Sri Lanka,	2011-22	121.8	Classical	MC
PHAMA Plus	Fiji, PNG, Solomon Islands, Vanuatu,	2018-22	32.0	Hybrid	MC
PRISMA 2	Indonesia	2019-23	95.0	Classical	MC
Strongim Bisnis	Solomon Islands	2016-20	58.2	Hybrid	MC
TOMAK	Timor-Leste	2016-21	25.0	Hybrid	MC plus INGO

<sup>a</sup> MDF is a multi-country program under a single contract

Notes: INGO – International Non-Government Organisation, MC – Managing Contractor, AGA – Australian Government Agency.

- Cambodia Agricultural Value Chain program (CAVAC) I and II are described by some interviewees as facilities, since they incorporate three relatively distinct sub-programs, of which a smaller one utilises an MSD approach.
- Strongim Bisnis, while seemingly designed as a mainstream program, appears to combine MSD elements with features more common to a business enabling environment program, with elements of a facility capable of responding to DFAT requests to work on specific issues.
- To'os ba Moris Di'ak (or Farming for Prosperity / TOMAK) comprises two components: food security and nutrition and market systems development. The MSD component promotes commercial agriculture by nudging subsistence agriculture towards commercial, and includes working with businesses, research institutes and government.

<sup>6</sup> This represents DFAT's contribution to a total of USD505m (end of 2018)

## Others

- The Livelihoods and Food Security Fund (LIFT) is really a multi-donor fund with some of its grants to businesses and NGOs under the thematic area 'inclusive market development and the private sector' using an MSD approach to integrate smallholder farmers and businesses to 'step up' into commercial agriculture value chains, principally using contract-farmer models. LIFT also has some financial inclusion projects that use an MSD approach.

## MSD Programs' sectoral focus

Given the context and drivers for DFAT's adoption of the MSD approach, it is not surprising that all programs operate in the agriculture and rural development space. However, some have no sectoral boundaries placed on their operation, while others include one or more sectors outside of agriculture. And some programs have specific locational areas of focus - see Table D.3.

Table D.3 Sectoral and geographic boundaries in DFAT's MSD programs – at design stage

Program	Sectoral boundaries of MSD component	Other components or outcome areas	Geographic boundaries
AIP-Rural (PRISMA I)	Agriculture, agricultural finance, irrigation, agricultural innovation	None	5 provinces in Eastern Indonesia: East Java, West and East Nusa Tenggara, West Papua and Papua
AMENCA3	Agriculture, horticulture	Infrastructure, extension services	No boundaries
CAVAC I	Agribusiness	Irrigation, Business Enabling Environment	3 provinces: Kampot, Takeo and Kampong Thom <sup>a</sup>
CAVAC II	Agribusiness	Irrigation	No boundaries
GREAT	Agriculture, Tourism	Women's agency, governance and policy	2 provinces in north west Viet Nam – Lào Cai and Sơn La
LIFT	Agriculture	Finance, migration, nutrition, civil society and four geographic programs	No boundaries
MDF I and II	None <sup>b</sup>	None	Five countries, no internal boundaries
PHAMA Plus	Agriculture, horticulture and cultural products	Export market access	Four countries, no internal boundaries
PRISMA 2	Agriculture, horticulture, livestock and aquaculture	Business enabling environment	AIP-Rural provinces plus Central Java
Strongim Bisnis	<b>Design</b> - Cash crops, tourism, and inter-island shipping, <b>Currently</b> cocoa, coffee, tourism,	<b>Design</b> - business practices, market development, risks and resilience, and women's participation in the economy <b>Currently</b> - business enabling environment and women's agency	No boundaries
TOMAK	Commercial agriculture	Nutrition Sensitive Agriculture, nutrition and food security policy and behaviour change	3 Provinces Bobonaro, Baucau and Viqueque

<sup>a</sup> subsequently expanded to all of Cambodia

<sup>b</sup> in Sri Lanka later focus targeted on tourism and related activities to conform with DFAT country strategy

## Targeting of specific groups

While all programs now identify poor people as the main beneficiary group for the operations, this was not always explicit in original designs. And early programs had limited targeting of women's economic empowerment or agency. This is changing – see Table D.4.



Table D.4 Targeting in DFAT's MSD programs

Program	In design	In operation
AIP-Rural (PRISMA I)	Targeting poor smallholder farmers No gender targeting	Sector and intervention selection shaped by expected impacts on poor and women. Reporting on results by household income, identifying poor and very poor beneficiaries Moving from do no harm to prioritizing activities that would benefit women Gender disaggregated reporting developed over time.
AMENCA3	No specific targeting of poor people. Empowerment of women and youth a specified outcome.	No specific targeting of poor people. Empowerment of women and youth a specified outcome.
CAVAC I	Implicit targeting of the poor by sector choice No gender targeting	Program worked with farmers with means and capacity to achieve a marketable surplus – not the poorest Do no harm approach to gender No clear targeting on or reporting by gender
CAVAC II	Implicit targeting of the poor by sector and locational choice No gender targeting	Program worked with farmers with means and capacity to achieve a marketable surplus – not the poorest WEE and Gender strategy developed, but limited implementation No targeting on or reporting by gender
GREAT	Empowering local women a specific objective, and there is a strong focus on women in the other two objectives	Empowering local women a specific objective, and there is a strong focus on women in the other two objectives
LIFT	Target poor women and men	Target poor women and men
MDF I and II	Mandate to increase employment and incomes of poor men and women	Sectoral choice and diagnostic work focused on opportunities for poor men and women to benefit. Phase I developed facility wide WEE framework and strategy. Gender disaggregated reporting, targeting of engagement areas where women will benefit, and women-focused interventions
PHAMA Plus	No explicit targeting of poor people. WEE mainstreamed	Gender disaggregated indicators, analysis of the number of women benefit and how they benefit. Disaggregation of beneficiaries to identify those below a poverty line.
PRISMA 2	60 per cent of smallholder households benefitting from the program are to be poor (below a specified poverty line). Women to be specifically targeted	60 per cent of smallholder households benefitting from the program are to be poor (below a specified poverty line). Women to be specifically targeted
Strongim Bisnis	Poor people as target beneficiaries Explicit outcome related to women and youth's economic opportunity, with specified performance measures WEE specified as a sector of engagement	Activities targeting women's agency
TOMAK	Gender and social inclusion as a cross-cutting theme	Gender targets in expected development outcomes, value chain selection expected to demonstrate opportunity for women's economic engagement

Across AQC scores on the gender criteria, MSD programs score above non-MSD programs in the same set of countries (4.75 compared to 4.26 – Table D.3).<sup>7</sup> And over the period 2015-2018, the gender ratings of MSD programs have averaged 4 or above (6 is the highest rating while 1 is the lowest).

<sup>7</sup> AQCs between 2015 – 2018 covering a range of programs. There is not an AQC for all years for all programs.

Table D.3: Average AQC scores by year of the gender equality criteria across the MSD portfolio

AQC year	Average gender equality
<b>2015</b>	5.00
<b>2016</b>	4.00
<b>2017</b>	4.83
<b>2018</b>	4.86
<b>Average</b>	<b>4.75</b>

Table D.4: Average AQC scores of the gender equality criteria by program

Program	Average Gender AQCScore
AIP-Rural	5.00
AMENCA 3	4.50
Cambodia Agricultural Value Chain II	4.00
Market Development Facility - Fiji	5.00
Market Development Facility - Pakistan	5.00
Market Development Facility - Sri Lanka	3.00
Market Development Facility - Timor-Leste	5.00
TOMAK	5.50

### Quantitative headline results

The tables below summarise the results related information largely drawn from AQCs and program progress reports. Consistent program data is not available across all data fields as programs may use different indicators of success or data is not included in the reports reviewed. Program commencement periods and implementation contexts mean it is difficult to compare performance directly across programs. Additionally, programs may use different definitions for the same indicators, e.g. investment leveraged. AIP-Rural and Strongim Bisnis report some data by household rather than individual beneficiaries.

Table D3 incorporates measures of efficiency – that is the percentage of expenditure and implementation. However, explanations of whether or not this is on track or for variations may not be included in AQCs. Some MDF information is not available for individual country programs as the MDF program is reported as part of a wider economic or inclusive growth aid investment.

Table D5 summarises commonly used DFAT ADR results.

Table D3: Program progress – AQC 2019<sup>8</sup>

Program	Program Duration	Country	Budget (AUD, m)	Expenditure to date (AUD m)	Expenditure as %	Implementation as %
<b>TOMAK</b>	2015-21	Timor-Leste	28	13	46.0	63.9
<b>MDF II</b>	2017-22	Timor-Leste	15.8	9.5	60.0	65.7
<b>MDF II</b>	2017-22	Fiji	29	14.8	50.9	67.9
<b>MDF II</b>	2017-22	Pakistan	18.7	4.3	22.7	35.4
<b>MDF II</b>	2017-22	Sri Lanka	Not available	Not available	Not available	Not available

<sup>8</sup> This information has been extracted from AQCs. Definitions for the percentage of implementation and expenditure are not known.

<b>MDF II</b>	2017-22	PNG	Not available	Not available	Not available	Not available
<b>MDF II – Central</b>	2017-22	Canberra	15.7	3.2	20.3	31.9
<b>Strongim Bisnis</b>	2017-20	Solomon Is.	up to 14			
<b>CAVAC II</b>	2016-21	Cambodia	86.9	58.6	67.4	67.6
<b>AIP Rural</b>		Indonesia	112.0	99.0	89.0	100.0
<b>PRISMA 2</b>	2019-23	Indonesia	95.0	Not available	Not available	Not available
<b>AMENCA 3</b>	2016-21	Palestine	26	15.6	64.8	59.8

Table D4: 2018 Annual Results (ADR): Innovation; poor beneficiaries, including women; production and exports

Program Details			Timeframe		Innovation		Poor beneficiaries		Value additional production (USD)	Value of exports facilitated (AUD)
Name	Country	Sector	Year commenced	Results up to	No. poor who adopt Innovative practices	No. Poor women adopt Innovative practices	No. poor with increased incomes	Poor women with increased incomes		
TOMAK	Timor-Leste	Agri & Livestock: Component 2	2016	Dec 18	Not Collected					
MDF II <sup>9</sup>	Fiji	Agriculture, Tourism, out-sourced services	2011	Dec 18	8,169	4,085	9,685	4,825	608,000	0.430 m
	Timor-Leste	Agribusiness, tourism and manufacturing	2012	Dec 18	2,984	1,573	3,885	2,231	343,000	0.542 m
	Pakistan	Agriculture	2013	Dec 18	30,780	15,120	33,203	15,427	4,997,000	15.413 m
	Sri Lanka	Tourism & Authentic Sri Lankan Goods	2015	Dec 18	25,648	2,134	26,329	2,653	608,000	3.774 m
	PNG	Agribusiness	2015	Dec 18	796	207	859	232	0	0
CAVAC I <sup>10</sup>	Cambodia	Fertilizer, Pesticides	2011	Sep 15	Not collected					
CAVAC II <sup>11</sup>	Cambodia	Agribusiness	2016	Jun 18	Not collected					
AMENCA 3 <sup>12</sup>	Palestine	Agriculture	2016	Dec 18	7,646*	Not collected	20,691*	6,461*	10.35 m	7.32 m
Strongim Bisnis	Solomon Islands	Tourism, cocoa and coconut	2017	Dec 18	84	58	9 (Households)	n/a	A\$ 67,905	None
AIP-Rural PRISMA 1 <sup>13</sup>	Indonesia	Agribusiness and Agri-Finance	2013	Dec 18			345,000 (Households)			

<sup>9</sup> MDF Annual Report, 2018 (February, 2019)

<sup>10</sup> ODE, December 2017 Cambodia Agricultural Value Chain Phase One evaluation. P.37. Agribusiness covers pesticide and fertilizer (DFAT, 2017b)

<sup>11</sup> CAVAC II, Six Monthly Report 1 January — 30 June 2018, (July 2018)

<sup>12</sup> AMENCA 3 2018 Annual Report

<sup>13</sup> PRISMA did not record data against ADR indicators related to increased income for poor men and women etc after 2017. This data comes from the AQC 2019.

## Examples of market systems changes

Outlined below are some examples of market systems changes achieved on longer-running MSD programs.

### CAVAC I: Regulation of chemical pesticides in Cambodia

#### **Market problem:**

- In 2010, pesticide use in Cambodia was unregulated. Farmers were not following the government's integrated pest management approach and were instead using chemical pesticides. Farmers were buying products of dubious quality and using them indiscriminately and unsafely. Rice farmers were losing up to half their crop yields to pests.
- The government introduced a new law for pesticides but did not develop supporting regulations. This meant pesticide retailers had no means of complying. There were no credible sources of information for retailer or farmers on chemical pest control, such as national pesticide guidelines and textbooks.

#### **Solution:**

- First CAVAC attempted to engage private sector companies (i.e. chemical pesticide distributors) to influence them to improve their products and services. However, without a change in the market system, there was little incentive for these companies to change.
- Second CAVAC engaged with a number of key market actors across government, retailers and farmers to increase understanding of how to apply chemical pesticides correctly and safely. CAVAC worked with the Ministry of Agriculture, Forestry Fisheries to develop a handbook on the preparation of pesticides; worked with provincial government departments to improve the quality of training to pesticide retailers; and worked with pesticide companies to support their own training of retailers and farmers.

#### **Systemic change:**

- Policy change: Cambodia became one of the first developing countries to standardise labelling for the safe and effective use of chemical pesticides.
- By the end of 2017, 168,717 households had changed how they used chemical pesticides, to use safely and correctly.

### Small packaged silage in Pakistan <sup>14</sup>

#### **Market problem:**

- Small dairy farmers in Pakistan faced a shortage of nutritious fodder, especially in extreme summers and winters. During these months, farmers are forced to feed their animals' substandard fodder which resulted in low milk yields and reduced income.
- Silage, a fermented fodder, has the benefit of being storable and its high nutritional value makes it an effective input throughout the year. However, historically silage has been produced and sold in 1,000 and 300 kilogram bales, a size that is neither affordable nor practical for smallholder farmers.

#### **Solution**

- Develop and test a small-scale silage business model (selling 60 kg bales), with customised distribution channels, among rural entrepreneurs.

#### **Systemic change**

---

<sup>14</sup> MDF II Annual Report (2018)

- Crowding-in of private and public sector actors:
  - 98 small-scale silage businesses operating throughout the country, reaching over 28,000 farmers.
  - Punjab Livestock and Dairy Development Board and other key market providing 60kg bales for small farmers.
  - State Bank of Pakistan included silage machinery in its agriculture machinery portfolio, enabling small-scale silage entrepreneurs to obtain access to low-interest loans.

PRISMA: Influencing regulations in the Indonesian shallot sector<sup>15</sup>

**Market problem:**

- More than 120,000 farmers cultivate shallots, an important ingredient in Indonesian cuisine, but productivity and quality is low. This is because farmers purchase low quality shallot bulbs from traders or split the shallots that they harvest to re-planting (which also leads to low quality). Farmers do not grow shallots from seeds because this is technically difficult and the failure rate is high.

**Solution:**

- Work with private sector nurseries in Lombok and Sumbawa to develop a higher-quality shallot planting bulb from imported seeds. PRISMA worked with a company called EWINDO, which ran private nurseries and imported high-quality seeds. The high-quality bulbs produce a 50% increase in productivity.
- Work with the government to train private nurseries in Central Java to grow the seeds into bulbs to sell to farmers.
- Work with local and national government authorities to change regulations to enable the uptake of the new bulbs.

**Systemic changes:**

- The government agriculture authorities developed a new certification process to certify the new bulbs.
- The Ministry of Trade lifted restrictions on the import of higher volumes of shallot seeds giving EWINDO and another shallot seed importer permission to import as much seed as required to meet the increased demand
- EWINDO's sales of shallot seeds have now grown from 300kg to 2,000kg, and there are more than eighty nurseries in Indonesia serving 10,000 farmers with improved planting bulbs.

---

<sup>15</sup> PRISMA II investment design document (2017)

## The VfM of MSD programs

### Mid-term reviews' and evaluations' conclusions on value for money

Table F.2 below summarises conclusions made by MTR and evaluation teams on value for money.

*Table F.2: Summary of VfM judgements from MTR and evaluation reports*

Program	Report	Value for money judgement
<b>AIP-Rural</b>	MTR (2016)	AIP-Rural monitors three value for money indicators: investment leverage, social return on investment, and investment per farmer. Early evidence suggests that PRISMA delivers value for money. Several PRISMA interventions already generate more net income for farmers than the direct cost of the intervention (e.g. Shallot Quality Seeds and early flowering mango). Even with the total program costs accounted for, AIP-Rural provides value for money when benchmarked against alternative rural development approaches such as the graduation model (citing one study). The report also notes that intervention staff costs are budgeted and reported to DFAT as management rather than as an attributable intervention cost. Personnel and management costs are routinely monitored, and factored into portfolio reviews of interventions, but should be reported to DFAT as an integral cost of each intervention.
<b>AMENCA 3</b>	MTR (2019)	While the overall budget was high when compared to similar MSD projects in similar sized economies, current expenditure levels are well aligned with expectations for MSD projects, impacts are being achieved and therefore the 'approach is on track to deliver excellent value for money'. Lower than expected expenditure to date should not be interpreted as poor performance.
<b>CAVAC I</b>	MTR (2012)	Projected benefit-to-cost ratio of seven to one using a ten-year time horizon from the start of CAVAC Phase one in 2010 and a 6% discount rate.
	Evaluation (2017)	The number of households assisted through CAVAC and the likely (although unmeasured) changes in incomes of these households represents a positive return on DFAT investment. CAVAC allocated two-thirds of their activity/intervention budget to technical delivery, which is in line with DFAT's good practice benchmarks. In the agribusiness component, CAVAC leveraged 66 cents in private sector financing for every dollar invested by DFAT. CAVAC could not calculate returns on investment in agribusiness or irrigation components because they did not have information on increases in income and yields. However, analysis of the outreach figures collected by CAVAC, along with interview data, does provide sufficient evidence to support the conclusion that a range of factors, in addition to development returns, played a role in shaping the CAVAC investment mix. At least half CAVAC (Irrigation) interventions yield net social return (direct intervention costs to projected income change in farmer households).
<b>CAVAC II</b>	Irrigation Impact Study (2019)	Irrigation component only based on an analysis of three schemes: The analysis, computed over a 30-year period and using constant 2013 prices, generated an internal rate of return of 17.5 per cent for Boeung Leas, 10.3 per cent for Chamlong Chrey and 41.6 per cent for Wat Thmey. Ten per cent was considered a 'pass' for the financial feasibility of economic projects (e.g. usually used for infrastructure) and four to six per cent was the proposed threshold for social projects. Qualitative data was also collected from beneficiaries to understand other benefits that were not as easily quantified, and their significance.
<b>MDF I</b>	MTR (2012)	The benefits expected were expected to meet or exceed the benefits of the program estimated at design. The total additional income expected to be generated by the end of Phase 1 amounted to AUD 94m which the review team felt compared very favourably with the AUD 48m that will have been invested by that time.

Programs not listed in the table have not had a MTR or evaluation or the report did not provide an assessment of VfM. CAVAC, a hybrid MSD program, conducted an impact study of its irrigation component, which did not use a MSD approach.

#### Comparisons using DFAT's internal assessments

Partner Performance Assessments (PPAs) were not examined as part of this review<sup>16</sup>. AQC's do not have a specific section or rating for value for money although DFAT investment managers sometimes include the VfM indicators reported by programs. However, given the multi-dimensional nature of value for money, there is some value in comparing AQC scores for MSD and non-MSD programs. Table F.8 reports the results of this comparison. (Only countries where MSD programs are being implemented have been included in the comparison). This analysis provides some initial insights, which should be further examined in a more considered analysis.

While there is little variation across program types, a comparison of AQC scores shows how market systems development programs rate, on average, in comparison to other interventions. Six is the highest possible rating.

Table E.3: Average program AQC efficiency scores<sup>17</sup>

Program type	AQC Year					Ranking across all years
	2015	2016	2017	2018	All years	
Multi-lateral	4.40	4.36	4.64	4.83	4.54	1
MSD	4.50	4.00	4.29	4.63	4.40	3
Non-MSD	4.43	4.26	4.23	4.46	4.34	4
Budget Support	5.00	2.00			3.50	6
Scholarship	4.29	4.20	5.00	5.00	4.50	2
Covers more than one program type (including MSD)	4.00	3.00	4.00	4.00	3.83	5
All program types	4.42	4.23	4.30	4.53	4.37	

<sup>16</sup> The PPA process (and ratings matrix) asks DFAT managers to assess the extent to which the implementer 'Operates in accordance with DFAT's Value for Money Principles'. However, the desired behaviours at each performance level – satisfactory to unsatisfactory – do not explicitly describe what the principles look like operationalised and rather focus on the economy aspect of VfM.

<sup>17</sup> Only includes programs in countries where MSD programs are being implemented. Some AQC's are prepared for entire aid investment plans and it is not possible to extract only the MSD component. For instance, MDF Fiji, Sri Lanka and Pakistan are reported as part of DFAT's wider economic opportunities or inclusive growth portfolio.



Table E.4: Average program AQC effectiveness scores<sup>18</sup>

Program type	AQC Year					Ranking across all years
	2015	2016	2017	2018	All years	
Multi-lateral	4.40	4.50	3.91	4.08	4.25	4
MSD	4.50	4.67	4.43	4.50	4.50	2
Non-MSD	4.55	4.39	4.32	4.44	4.42	3
Budget Support	5.00	2.00			3.50	6
Scholarship	4.86	4.20	5.00	4.67	4.67	1
Covers more than one program type (including MSD)	4.00	4.00	3.50	4.00	3.83	5
All program types	4.55	4.38	4.28	4.40	4.40	

<sup>18</sup> Only includes programs in countries where MSD programs are being implemented. Additionally, some AQCs are prepared for entire aid investment plans and it is not possible to extract only the MSD component.

## Annex E: MSD Program M&E systems

### Introduction

This synthesis review provides some insights into the current state of results, monitoring, evaluation and progress reporting. These insights have been gathered largely through reviewing program documents such as progress reports; external advisory group, mid-term review and evaluation reports, and DCED Standard audit reports; as well as M&E manuals and plans. The array of documents reviewed varies across programs since programs are at different stages of implementation. The nature of this review also means it is not possible to examine any program M&E system in detail.

Section 2 summarises DFAT's and programs' information needs and uses, including reporting requirements highlighted in DFAT's M&E Standards. As most MSD programs in the portfolio use the Donor Committee for Enterprise Development Results Measurement Standard, an overview of the Standard is provided along with a brief assessment about how it meets DFAT's information needs.

Section 3 provides findings from a rapid review of whether the MSD Programs M&E systems provide credible and defensible information, utilising key characteristics of effective M&E systems drawn from DFAT's evaluation of investment monitoring systems (DFAT, 2018b).

Section 4 examines the breadth and depth of information reported from the headline indicators to intervention level, highlighting where programs are meeting information needs and where there are gaps. Section 5 highlights the lack of evidence on the sustainability of results while Section 6 provides some final observations on M&E for MSD programs and emphasises areas needing further discussion, including strategic priorities and resourcing.

The term monitoring and evaluation used in this Annex refers to process and outcome monitoring, internal evaluative activities, and external reviews and evaluations. This captures the range of activities under DFAT's M&E Standards. It includes the DCED Standard for Results Measurement's concepts of monitoring and results measurement (MRM).

### DFAT's and programs' information needs and uses

#### Key points:

- DFAT and programs have a range of information needs. DFAT needs headline indicator information to meet some of its accountability obligations within the Australian Government. DFAT also needs a range of additional information to meet a variety of accountability, learning and communication needs. Programs need information for management and learning. Some of DFAT's requirements, such as those relating to value for money assessments, are not sufficiently clear.
- Program progress reports are often required to meet a diverse range of information objectives. Often reports do not cover areas of central importance for an MSD approach such as clearly analysing how market constraints have changed, the effectiveness of interventions in addressing constraints, and the significance of changes achieved for market actors, especially vulnerable groups.
- Meetings between programs and DFAT Posts are better able to meet some of DFAT's information needs.

## Information users and needs

The main elements of DFAT's and MSD programs needs are summarised in Box E.1.

Box E1: Users information needs

**DFAT's accountability needs:** recording how funding and other resources are being used, how policies and safeguards are being complied with, and how risks are being managed.

**DFAT's performance management needs:** periodic synthesized information related to progress in implementation including against outcomes, *including systemic change, at a sector, country and portfolio level (relevant to the nature of the program).*

**DFAT's learning needs:** periodic synthesised information indicating what is working, what is not working, and significant changes in the context that might affect results.

**DFAT's communication needs:** regularly short, clearly written engaging human-interest stories appropriate for Ministerial briefings and social media communications; synthesised targeted information suitable for one-page cables; and occasional longer and more detailed communication products, such as case studies and video presentations.

**Implementers' performance management needs:** timely progress in implementation against outcomes, *including systemic change, at an intervention, sector and portfolio level.*<sup>19</sup>

**Implementers' learning needs:** timely information indicating what is working, what is not working, and significant changes in the context that might affect results.

## Monitoring, evaluation, value for money assessments and reporting requirements

### DFAT's M&E Standards

DFAT's needs are captured in a variety of places - the DFAT M&E Standards, Aid Quality Check template and Partner Performance Assessment Framework (see DFAT 2018d). For the purposes of this review, program progress reports have been the main documentary source to assess the degree to which systems meet DFAT's information needs. The range of documents available across programs varied considerably: naturally more documents were available for longer-running programs. DFAT's M&E Standards include standards for investment progress reports (Box E.2).

The DFAT's M&E Standards (DFAT, 2017c) 'are not meant to represent a checklist of rules that must be adhered to' but rather provide guidance for DFAT and implementers as to what a quality product ought to look like. The Standards provide 'guidance on where flexibility is appropriate and how to handle this without compromising on rigour. The two key messages around flexibility are a) the degree of rigour should be proportionate to the importance of the decisions being made; and b) the basis of findings must be transparent.' Examples of flexibility include the ability to change end of program outcomes and flexibility to choose appropriate evaluation theory, methodology and methods.

---

<sup>19</sup> This is wider than performance monitoring which is usually equated with quantitative data relating to outputs.

### Investment progress reports

*3.11. For claims of achievement or barriers to achievement, credible supportive information is provided.*

For claims of achievement (both in terms of achievement of outcomes, and quality of outputs or deliverables), or the explanation of barriers to achievement, credible supportive evidence is provided. The basis by which the claim is made is articulated. There is not an overemphasis on using examples to demonstrate achievement, rather the emphasis is on how we know that these examples have been achieved or outputs are of sufficient quality (the basis of the claim).

*3.4. An assessment of the adequacy of progress toward sustained end-of-program outcomes is described*

Where relevant, a distinction is made between end-of-program outcomes and those outcomes that are expected to be sustained beyond the life of the investment. A firm judgement of the adequacy of progress toward these outcomes is described. This also includes outcomes relating to cross-cutting issues such as gender equality, disability and social inclusion. A simple list of immediate and/or intermediate outcomes or achievements is not sufficient. A judgement of the adequacy of this progress must be explicit, and explained with reference to appropriate criteria, standards and evidence. Where progress has been particularly good, or particularly poor, a full exploration of the supportive and inhibiting factors that account for this situation is provided. Particularly where problems or challenges have been identified, there is a full analysis of the implications of the situation on the likely achievement of end-of-program outcomes. It is clear from the report the extent to which DFAT and other stakeholders will need to track the relevant issues, including whether there is a need to reassess the suitability of end-of-program outcomes.

### The DCED Standard for Results Measurement

Many of the MSD programs in the portfolio use the Donor Committee for Enterprise Development (DCED) Results Measurement Standard. The DCED Standard was developed in 2009, three years before DFAT's Department-wide evaluation guidance was published. The DCED:

*promotes a pragmatic approach to results measurement. It calls on programs to measure results to a level that is complex enough to be credible, yet simple enough to be practical. In the words of John Maynard Keynes, "it is better to be roughly right than precisely wrong". The key test of the DCED Standard is whether the approach taken by the program would convince a reasonable but sceptical observer. (DCED, 2018)*

Since the DCED Standard's inception, DFAT has supported its adoption and most program investment designs emphasise that the DCED Standard provides a solid foundation on which the M&E system should be built<sup>20</sup>. The designs of more recent programs, such as PHAMA Plus and TOMAK reference the DCED Standard and DFAT's M&E Standards, and as such these programs have expanded their M&E system beyond the DCED Standard. For instance, PHAMA Plus, Strongim Bisnis and TOMAK M&E Plans include evaluation questions, which is a feature of DFAT's M&E Standards but not the DCED Standard.

### Comparison of DFAT's M&E Standards and the DCED Standard

Applying the DCED Standard brings a level of rigor and meets some of DFAT's needs e.g. aggregated high-level indicators that are attributable to the program. The DCED Standard provides one mechanism that can help programs manage adaptively. The Standard also includes some elements not captured in DFAT's M&E Standards, but which are considered good practice.

---

<sup>20</sup> This includes: AMENCA 3, CAVAC, GREAT, MDF Phase I and II, PHAMA Plus, , PRISMA

For instance, the DCED Standard requires senior staff being familiar with sector and intervention-level results chains (or theories of change) and stipulates that the results chain should be reviewed at least annually (Dart, 2016).

DFAT's M&E Standards take an evaluative approach structured around key questions. Using the DCED Standard can help programs assess the questions relating to relevance (to some extent), effectiveness, impact and sustainability. However, it does not seek to assess efficiency, nor cost-effectiveness. There are perceptions that the DCED Standard more strongly supports the monitoring of a small set of quantitative indicators – outreach, increased income and jobs. Since strategies like policy engagement are unlikely to contribute directly to these indicators, some see the Standard as less applicable for these types of interventions. Programs may give less attention to M&E of these types of interventions compared to those that directly contribute to outreach and increased income.

### **DCED Standard audits**

DFAT also wants to understand if the results reported by programs are credible. The DCED Standard audit is a systems audit, which can in part help DFAT understand the robustness of programs M&E systems. The audit reports, beyond the audit score, can be valuable tools for programs and DFAT. DCED Standard audits, however, are not a verification of results see Box E.3). There are, perhaps, some misunderstandings of what the Standard is, and therefore what an audit audits and what an audit score means<sup>21</sup>. The Standard's bar is 'set' at a level that is already happening in the field of several private sector development programs. It not ahead of the 'leading' programs in the field. The Standard is periodically updated to reflect changes in the field. The Standard provides a set of guidelines for users to apply to a variety of different situations and contexts. There are some perceptions that the Standard is rigid but that may result from how the guidelines are interpreted, rather than from inherent rigidity.

#### *Box E.3: DCED Standard for Results Measurement audit*

Audits are an audit of the system in practice and the extent to which the system in practice complies with each of the control points that comprise the Standard.

Audits are not a verification of results reported. Audit reports highlight strengths and weaknesses regarding the quality of a 'pragmatic approach' to data collection and analysis and therefore provide insights into the credibility of results reported.

Generally, depending on the size of the program, two auditors spend an intensive week with a program. This is in addition to preliminary work. Across the MSD programs, the audit reports present the 'deepest dives' into program M&E systems that have been available for this review.

A strength of the DCED Standard is that it is grounded in practice, supporting people in the field to address their key challenges with regards to M&E. Assessing systemic change is one of the most challenging areas covered by the Standard and is one of the lowest scoring areas in audits. This challenge is discussed more in Section 4.

### **Value for money requirements**

The Australian Government Public Governance, Performance and Accountability Act (2013) provides a foundation for DFAT's focus on value for money. The purpose of the Act is to:

---

<sup>21</sup> MDF II Design peer review.

- establish a coherent system of governance and accountability across Commonwealth entities; and
- establish a performance framework across Commonwealth entities; and
- require the Commonwealth and Commonwealth entities to:
  - meet high standards of governance, performance and accountability; and
  - provide meaningful information to the Parliament and the public; and
  - use and manage public resources properly; and
  - work cooperatively with others to achieve common objectives, where practicable; and
  - require Commonwealth companies to meet high standards of governance, performance and accountability.

DFAT's 2014 paper *Making Performance Count: enhancing the accountability and effectiveness of Australian aid* articulates how DFAT will operationalize the Act, and includes a specific focus on ensuring value-for-money. It states the aid program will:

*Deliver high standards of value-for-money in at least 85 per cent of aid investments. Where standards are not met and improvements are not achieved within a year, investments will be cancelled. Investments not delivering value-for-money will be required to be improved within one year or be cancelled. This target will require that from 2014–15, at least 85 per cent of investments being implemented satisfactorily meet effectiveness and efficiency standards reflecting new value for-money principles. This will build on and strengthen existing systems which require the performance of investments to be assessed annually. We will also ensure that value-for-money is achieved in the deployment and remuneration of advisers.*

DFAT has also developed value for money principles, reflecting the requirements of the PGPA Act<sup>22</sup>. See Table E.1.

*Value for money is a key consideration in decision-making for all aspects of the aid program. DFAT's Value for Money Principles seek to ensure the effective, efficient, economical and ethical management of Australian aid, in a way that advances Australia's national interests and achieves the Government's policy commitments.*

The APG also notes that all aid investment plans are to identify a set of performance benchmarks for the life of an Aid Investment Plan (AIP), that is key planned program results and significant milestones and include measures of improved operational effectiveness or efficiency.

DFAT investment managers 'must ensure that responsibility for agreement management is clearly assigned to a nominated officer'. Agreement management includes 'ensuring deliverables meet the required standard, are produced within the agreed timeline and achieve value for money'. Further guidance emphasises that managing contracts includes ensuring deliverables meet the required standard, are produced within the agreed timeframe and achieve value for money'.

The APG does not provide detailed guidance on how investment managers are expected to manage for or assess value for money. DFAT investment managers assess managing contractors and NGOs' performance as part of the partner performance assessment (PPA) tool and through the annual aid quality checks (AQC). PPAs include a section on value for money<sup>23</sup>. The AQC includes an assessment of efficiency and effectiveness. Generally, MSD program implementers do

---

<sup>22</sup> Aid Programming Guide (APG) (DFAT, 2018c)

<sup>23</sup> desired behaviours at each performance level – satisfactory to unsatisfactory – do not explicitly describe what the principles look like operationalised and rather focus on the economy aspect of VfM.

not receive strong signals on the importance of value for money assessments from DFAT investment managers.

Table E.1: DFAT's value for money principles

Criteria	Principles	Brief description
<b>Economy</b>	Cost consciousness	To seek reasonable opportunities to reduce costs at every level of operation, with due consideration to its impact on efficiency and effectiveness. It does not mean always choosing the lowest cost option.
	Encouraging competition	A culture of contestability and competition of ideas and alternative solutions to development problems; use competitive processes when selecting partners.
<b>Efficiency</b>	Evidence based decision making	Requires systematic, structured and rational approaches to decision making framed around logical arguments supported by analysis
	Proportionality	Organisational systems need to be proportional to capacity and need to manage for results and/or deliver better outcomes and maximise efficiencies.
<b>Effectiveness</b>	Performance and Risk Management	The performance of interventions must be continuously reviewed to assess the extent to which they are meeting their objectives and delivering maximum impact. Robust risk management systems, which define risk appetite, are critical.
	Results Focus	Clearly identified objectives and targets are crucial to a strong results orientation. Innovation and adaptability, based on clear and logical evidence, is also important. Decision makers need to balance anticipated outcomes and benefits with the potential for increased risk and manage these accordingly
	Experimentation and innovation	Creative and innovative approaches to program design and delivery are required. Trialling experimental and innovative mechanisms where there is reasonable grounds to expect better outcomes overall. Appetite to trial new ways of delivering calibrating with risk tolerance.
<b>Ethics</b>	Accountability and transparency	Accountability and transparency can strengthen responsibility for results and continuous improvement of delivery and operational processes. Helps to create incentives. Donors must hold delivery partners accountable and demand transparency at all levels to facilitate honest dialogue about impact.

## The credibility of reported results

- DFAT is concerned that reported results are credible.
- MSD program M&E systems are outcome focused, quality assured, and information is used.
- MTRs, evaluations and DCED Standard audits find reported results are credible. There are some variations of quality in pockets of programs.

DFAT's evaluation of investment monitoring systems (DFAT 2018b)<sup>24</sup> identifies the most important features that distinguished higher-quality monitoring systems<sup>25</sup>. The evaluation examined some MSD programs and highlighted elements of better practice from them in the report (Table E2). These elements were used in a rapid assessment of the M&E systems of DFAT's

<sup>24</sup> The report included several recommendations to improve investment monitoring systems. DFAT identified several actions in response to these recommendations.

<sup>25</sup> MSD programs may refer to systems by different names. This review will use the term monitoring and evaluation (M&E) that covers monitoring, evaluation, learning, results measurement.



MSD programs. (It is important to note that M&E systems are not static and as such may have changed following the findings drawn on from available documents.)

Table E.2: Important features of higher-quality monitoring systems

<b>Outcomes focused</b>
There is clarity in what the investment, it's individual interventions, needs to achieve and how it is expected to effect change.
Outcomes, intended and actual, are distinguished by the different groups who experience them (e.g. men, women, poor, very poor, disabled etc.).
Feedback is generated early enough to inform changes in implementation
Feedback is used to inform any needed adjustments in the investment or interventions design, including the theory of change.
<b>Quality assured</b>
Monitoring standards are applied to assure high-quality data and reporting.
Contestability and independent review are built into the system and culture.
Technical quality assurance of monitoring systems and products are undertaken.
<b>Effectively used</b>
A mix of qualitative and quantitative data is collected and analysed.
Reporting is functional, and tailored to the needs of its audiences.
External partners and stakeholders are supported to engage with the data.

### Outcomes focused

#### **There is clarity in what the investment, it's individual interventions, needs to achieve and how it is expected to effect change**

Most programs have an overarching theory of change as part of the design or of strategies or plans developed following a period of analysis. At a high-level these are clear. Documentation reviewed differs in the level of detail about what will be done to achieve change. Generally, more detailed and clearer theories of change are developed when sector or sub-sector strategies are developed and/or partnerships agreed. Theories of change for program components using a direct delivery approach are likely to be more detailed and clearer from the design stage. This variation reflects the different implementation approaches.

Most of the program documents reviewed present only a high-level overview of how and why the umbrella theory of change of how the program is expected to trigger change within the specified context. Most programs, e.g. TOMAK, have conducted extensive research that has informed strategy and intervention development - this information may be captured in intervention-level theories of change or results chains, their assumptions and sector or sub-sector level strategies, which were not reviewed. MDF, for instance, also publishes discussion papers that provided further insights into the theory of change underpinning specific interventions.<sup>26</sup> For those programs that have conducted a DCED Standard audit, articulating results chains scores highly as meeting the Standard.

#### **Outcomes, intended and actual, are distinguished by the different groups who experience them (e.g. men, women, poor, very poor, disabled etc.)**

Generally, programs identify different beneficiary target groups. For instance, MDF and Strongim Bisnis intervention results chains will map out how women and/or youth will benefit. Programs disaggregate data according to selected target groups. For instance:

- Women / men: AMENCA 3, MDF, Strongim Bisnis, TOMAK

<sup>26</sup> For example, see the Small-bale Corn Silage Experience in Bear and Bekkers, 2018.



- Poor and very poor: AIP-Rural
- Youth: AMENCA 3<sup>27</sup>, Strongim Bisnis.

**Feedback is generated early enough to inform changes in implementation and is used to inform any needed adjustments in the investment or interventions design, including the theory of change**

Examples include:

- Some reviews and audits of the DCED Standard have suggested that systems supporting adaptive Programs conforming to the DCED Standard are expected to review and revise, where necessary, intervention-level results chains at least annually. Some reviews and audits of the application of the DCED Standard have suggested that systems support adaptive management quite well (for example, CAVAC).
- The CAVAC II MTR (2019) noted the monitoring of individual component activities is of high quality and meets the day-to-day management needs of activity implementation.
- Some LIFT project reports highlight mixed results e.g. the report for Yoma Bank highlights lessons and how adjustments to the intervention were made, while another project evaluation found no evidence that value for money assessments informed decision making.
- AIP-Rural's portfolio review process uses quantitative and qualitative indicators (outreach, benefit, rationale, value for money, quality of deal, quality of collaboration, systemic change potential, poverty, gender and environment) to rank interventions to support decision making regarding dropping, changing or improving interventions. (MDF Phase II uses a similar process.)

## Quality assurance

**Monitoring standards are applied to assure high-quality data and reporting.**

As noted earlier, programs are guided by the DFAT M&E Standards and/or the Donor Committee on Enterprise Development (DCED) Standard for Results Measurement. DFAT until recently assessed programs M&E systems annually as part of the AQC process. On average DFAT's MSD program M&E systems scored 4.69 in the AQCs over the period 2015-18 (with 6 being the highest possible score), higher than non-MSD programs'<sup>28</sup> score of 4.24 (Table E.3).

*Table E.3: Average AQC scores for M&E over 2015-2018*

Program	Average M&E AQC Score
AIP-Rural	5.00
AMENCA 3	4.50
Cambodia Agricultural Value Chain II	4.00
Market Development Facility - Fiji	4.75
Market Development Facility - Pakistan	5.00

<sup>27</sup> Initially AMENCA 3 defined youth by an age band that was not relevant to the implementation context and has since amended the definition. The MTR report (2019: 47) notes 'Overall the Australian definition of youth (16-24 years) is not helpful in the Palestinian context given the social and cultural mores. The restriction has locked youth interventions into peripheral activities such as vocational training for university students. Expanding the definition to include men and women up to 30 years of age (or even 35 years as is the case with Australian government funded programs in Myanmar) would enable AMENCA 3 to apply its core skills (agricultural market linkages) to young Palestinians, a critical group for future economic development.'

<sup>28</sup> Does not include multilateral, scholarships, budget support investments, or those investments that include MSD and non-MSD programs.

Market Development Facility - Sri Lanka	5.00
Market Development Facility - Timor-Leste	4.75
TOMAK	4.38
<b>Overall average</b>	<b>4.55</b>

### **Contestability and independent review are built into the system and culture**

Contestability and independent review come in several forms:

- External advisory groups: CAVAC II, MDF I and II, AIP-Rural, PRISMA 2 and TOMAK have such groups, some of which examined M&E quality as well as results and information reported.
- DFAT-led mid-term reviews and evaluations: Mid-term reviews have been conducted for AIP-Rural, CAVAC I and II and MDF I.
- Program-led evaluations: mid and end of term evaluation of LIFT projects appear standard practice. CAVAC II recently completed an external evaluation of its irrigation component. PHAMA Plus, PRISMA 2 and Strongim Bisnis plan to conduct selected evaluations<sup>29</sup>.
- DCED Standard audits: AIP-Rural, CAVAC I, MDF I Sri Lanka, MDF I Timor-Leste, MDF II Fiji, MDF II Pakistan have all been subject to these audits. PRISMA 2 notes it will complete another audit of its M&E system.

Overall, MTRs and evaluations find that data reported is credible. Observations from MTRs and evaluations include:

- AIP-Rural's Strategic Review Panel concluded that the results can be attributed to the program with considerable confidence.
- The CAVAC I end of program evaluation described the actual outreach figures of 214,550 by September 2015 for CAVAC's agribusiness interventions as plausible. However, in verifying outcomes related to the irrigation component the evaluation team felt that CAVAC recorded data was at the upper end of yields achieved; found fewer households connected to water during focus groups than outlined in CAVAC reporting; for some schemes, farmers are not actually paying for water.
- The CAVAC II MTR (2019) noted that while the program has conducted a high-quality evaluation of the irrigation component it has yet to 'develop a suitably rigorous methodology for measuring income increases in MSD interventions'. Following the development and costing of a defensible methodology, the MTR recommends DFAT and CAVAC discuss the likely cost-effectiveness, or alternatives such as less defensible and cheaper methodologies or using resources to ensure assessments of other outcomes are good quality.
- The MDF I MTR (2012) noted the program's approach to ex-ante and ex-post assessments of attributable results was overly conservative since:
  - projections were made for up to two years after an intervention ended (as guided by the DCED Standard that suggests sustainable results should only be claimed for two years since after that time it is expected many other factors will come into play) or to the end of Phase 1, if sooner given that the program may not be monitoring beyond that.
  - possible replication by other market actors or ripple effects from partnerships were not estimated ex-ante only ex-post
  - only effects that can be directly attributed to MDF's partnerships are counted; effects where MDF is upgrading the business as a whole (e.g. through book-keeping expertise) are not considered.

---

<sup>29</sup> PHAMA Plus MRM Plan, page 46; Strongim Bisnis Monitoring and Results Measurement Plan and Manual, page 36; PRISMA 2 Investment Design Document, page 39.

- The MTR recommended MDF include wider market affects, ensuring assumptions are explicit, to better communicate MDF's potential. A risk, however, is the projections will not be met and programs will be sanctioned. While MDF's interventions differ from CAVAC's irrigation interventions, MDF's two-year time contrasts markedly to CAVAC's irrigation scheme study that used a 30-year timeframe to estimate costs and benefits.

DCED Standard audits have confirmed that programs nearly always meet or exceed the compliance criteria related to attributable results. In coordinating audit processes, the DCED Secretariat has found programs find assessing the attribution of changes and assessing market systems changes are the most challenging components. Table E.4 compares DFAT's audited programs with a benchmark of other audited programs.<sup>30</sup>

Table E.4: DCED Standard for Results Measurement – Selected audit components<sup>31</sup> compared to benchmark

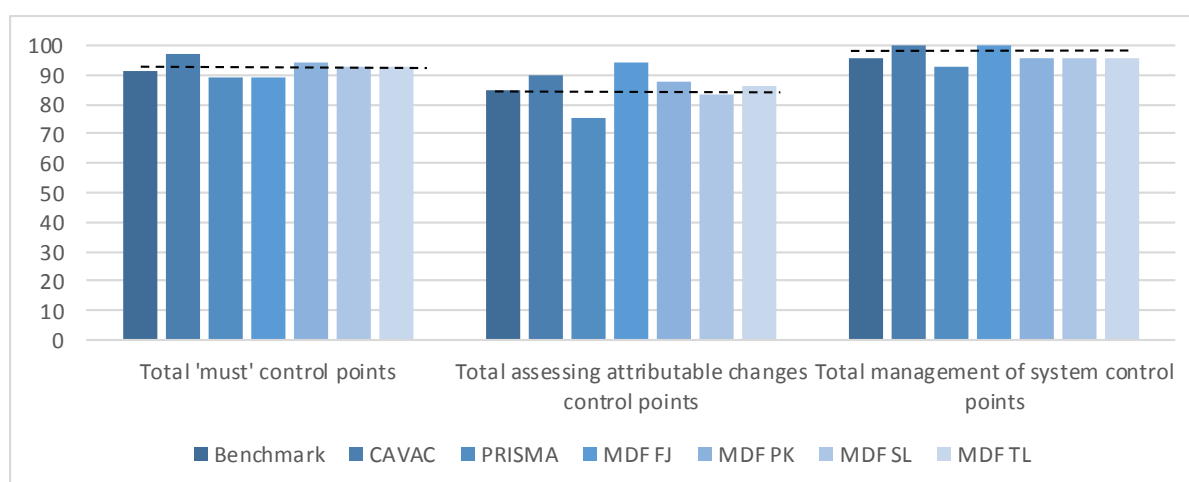
	Year audited	Total 'must' control points %	Assessment of attributable changes %	Management of system %
<b>Benchmark</b>		91	85	96
<b>AIP-Rural PRISMA</b>	2016	89	75	93
<b>CAVAC I</b>	2013	97	90	100
<b>MDF II FJ</b>	2018	89	94	100
<b>MDF II PK</b>	2018	94	88	96
<b>MDF I SL</b>	2017	93	83	96
<b>MDF I TL</b>	2017	93	86	96

Four of the six programs are higher than the benchmark for 'must' control points, two are higher for the assessment of attributable changes; while five are higher for management of the M&E system (Figure E.1). In many cases, programs have not operationalized processes to assess systemic change and therefore these were not audited. Programs are likely to have adjusted systems since audits, so these results may not reflect current practices.

<sup>30</sup> Only those published on the DCED website

<sup>31</sup> Since the first DCED Standard audit of Katalyst in 2013, a further 20 audit reports have been published. Some reports remain unpublished because organisations have not agreed to make them public, usually due to audits determining that systems did not meet the Standard. The benchmark was constructed by calculating the average score of nine other programs that published audit reports on the DCED website between 2013 and 2019. Two early audits of Katalyst in 2011 and 2013 were not included as reports did not include all the information necessary for comparison. While there have been some changes to the Standard over time, the reports generated at different times are broadly comparable. The latest version of the Standard combined two components – measuring change and assessing attribution. To make a comparison between earlier and more recent scores, the scores from these two components were combined.

Figure E.1: DCED Standard for Results Measurement – Selected audit components compared to benchmark



### Technical quality assurance of monitoring systems and products is undertaken

Program M&E manuals reviewed<sup>32</sup> include a variety of internal quality assurance processes such as.

- Quality checks embedded into various processes;
- Mitigating risks of double-counting impact;
- Internal and external quality assurance reviews of the system;
- Annual verification of a selected number of indicators<sup>33</sup>.

The most recent update of the DCED Standard (2017) added an internal review of the M&E system into the component relating to the management of the system. Some programs have conducted internal reviews of M&E systems. For instance, MDF reflected on lessons in developing a consistent and functional M&E system across five countries and finding a balance between anecdotal and rigorous, overly prescriptive methods. The PRISMA Progress Report (August 2018) noted a consultant was due to carry out a review of its M&E systems in August 2018.

### Effectively used

#### A mix of qualitative and quantitative data is collected and analysed.

Programs collect a range of quantitative and qualitative data, with most annual reports, such as AIP-Rural<sup>34</sup> and AMENCA 3<sup>35</sup>, more focused on quantitative information. Where qualitative information is provided, this is usually descriptive. Some programs report the frequency of achievement of qualitative indicators of success e.g. TOMAK's Monitoring, Evaluation and Learning Framework (MELF) includes 'Proportion of women of reproductive age that report having greater decision-making power and satisfaction in regards to household decision making, especially household food production, consumption and related household expenditure'.

MDF annual reports present qualitative information in the form of vignettes of positive changes. TOMAK and AMENCA 3 also present similar stories. Some vignettes include quantitative data.

<sup>32</sup> AIP-Rural, MDF, PHAMA Plus, Strongim Bisnis.

<sup>33</sup> Not all processes are mentioned by all programs.

<sup>34</sup> DCED Standard audit report (2016)

<sup>35</sup> AMENCA 3 MTR report

TOMAK MELF includes some other qualitative indicators, which are not indicators in the traditional sense of being SMART, (that is specific, measurable, achievable, realistic and timebound) but rather focus on collecting stories of positive change. For instance, 'Instances of implementing partners promoting improved nutrition practices in line with the SBC (Social Behaviour Change) strategy' and 'Instances of improved systems delivery (management/coordination of nutrition-related activities)'. These indicators do not assess the extent of the behaviour changes across target populations.

It is likely that programs are collecting a lot of qualitative data through informal processes (e.g. meetings with partners) and formal processes (semi-structured interviews, focus group discussions). From the documents reviewed, it is hard to determine the extent to which qualitative information is analysed.

### **Reporting is functional and tailored to the needs of its audiences.**

Programs produce six monthly or annual progress reports, with outcome-related results usually reported annually. Several annual reports present annual aggregated quantitative results complemented by qualitative stories or vignettes. It is often difficult however to understand how portfolios, or components of portfolios, are performing and what challenges and lessons are arising.

AQC's highlight some concerns by DFAT that programs are not meeting its information needs. The AMENCA 3 MTR report concludes there are numerous 'hidden' outcomes that are not reported suggesting it did not meet some audiences' needs. The CAVAC I evaluation notes that the M&E system was intended to produce information that could easily be incorporated into publications, but found many items of information difficult to digest e.g. explanations of the regression analyses used to estimate the yield impacts.

### **External partners and stakeholders are supported to engage with the data.**

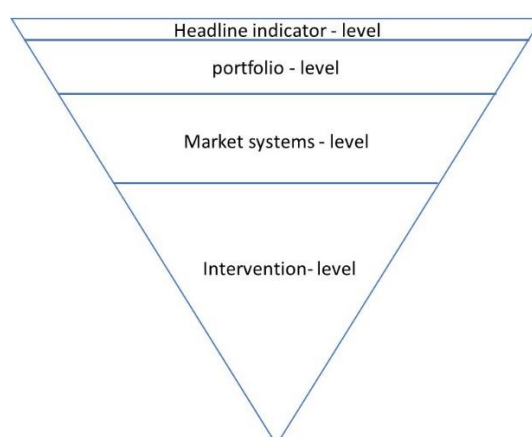
There is insufficient information in reports to determine the extent to which programs may be engaging external partners and stakeholders in data and analysis.

## **Assessing the breadth and depth of outcomes**

Ideally, programs should report on both the breadth and depth of desired and actual change. Figure E.2 presents a framework for differentiating between information breadth and depth from intervention to headline indicators for a typical MSD program. This section examines the type of results reported by programs at each level, and highlights different approaches to communicating results.

Overall programs seem to be doing better at meeting DFAT's needs for reporting results against headline indicators and programs' own needs for detailed information at the intervention level.

*Figure E.2: Breadth and depth of information across MSD programs*



## Headline indicators

### Key points:

- Results against headline indicators are useful in understanding a little about the breadth of results for individual programs.
- Some indicators are common across most programs.
- Aggregated numerical results, including averages, does not aid understanding of variations of performance within programs.
- Cross-program comparisons are difficult, if not impossible, since programs report a range of indicators and even when there are similarities programs may use different definitions for the same or similar indicators e.g. private sector investment leveraged.
- There are mixed perspectives on the role of targets for incentivising program performance.

**Annex D** provides tables of summarised headline results for the longer running programs – AIP-Rural, AMENCA 3, CAVAC I and II, and MDF. Other programs, such as PRISMA 2, PHAMA Plus, Strongim Bisnis and TOMAK are at earlier stages of implementation and are mainly reporting outputs.

### Challenges with understanding headline results

All programs report results progress and outcomes achieved annually, and some report more frequently <sup>36</sup>. Outcomes are normally characterised by quantitative indicators e.g. number of adults or households benefiting; percentage of female or poor beneficiaries; total increased net income for beneficiaries. All programs also report against selected DFAT's Aggregate Development Results (ADR). The ADR are reported as annual results, not cumulative. Examples include.

- |        |                                                                              |
|--------|------------------------------------------------------------------------------|
| ■ ADR1 | Number of poor women and men who adopt innovative agricultural practices     |
| ■ ADR2 | Number of poor women and men with increased incomes                          |
| ■ ADR3 | Number of poor women and men who increase their access to financial services |
| ■ ADR4 | Value of additional agricultural and fisheries production [USD]              |
| ■ ADR5 | Value of private sector investment leveraged [AUD]                           |

It is difficult to assess differences between planned (at the design stage or agreed later) and actual achievements in relation to headline indicators. Original targets, budgets and expenditure

<sup>36</sup> Reporting periods differ across programs. Strongim Bisnis reports the most frequently, while other programs report six monthly (e.g TOMAK) and others (e.g. AMENCA 3, LIFT, MDF II) annually.

variance are not often included in progress reports<sup>37</sup>. There are a range of practices across the portfolio, of which some are highlighted below.

- AMENCA 3 provides clear tables on aggregated actual quantitative results against targets, cumulatively and annually. The tables include all ADR indicators except ADR3 (access to finance). The results data is not disaggregated by intervention or implementing organisation, meaning variations in the performance across the portfolio are not reported.<sup>38</sup>
- MDF II provides clear tables on cumulative and aggregated actual achievements against targets, disaggregated by country. MDF II also provides graphs to illustrate the projected and actual trajectory of results over time (starting from phase I). MDF II provides expenditure data for the entire program, which is not disaggregated by country. It is not possible to compare expenditure to beneficiary outreach and increased income since much of the increased income is the result of expenditure under MDF I. Separate tables are provided to report on ADR1, 2, 4 and 5.
- The AIP-Rural Activity Completion Report<sup>39</sup> reported expected results at the end of December 2018 but not actual results. In previous years, PRISMA reported actual results in comparison with the results projected/targeted.<sup>40</sup> The comparison was presented as absolute numbers or percentages above or below the target e.g. 19,734 households reached above target; 244% above target for net income increased. The information presented is not always consistent across reports. ADR indicators are not included in progress reports<sup>41</sup>.

Some quantitative results include indicators of depth as well breadth. For example, AIP-Rural PRISMA and PRISMA 2 measure net income increases, with a target of a 30% increase.<sup>42</sup> However, reports often do not clearly explain the significance of the change for beneficiaries.

There is also a range of specific challenges related to how programs define and measure indicators and present analysis. Across programs, the same indicators are defined differently. For example, programs include and exclude different contributions when counting the value of private sector investment leveraged e.g. PRISMA includes amounts loaned to intermediary service providers and farmers, AMENCA includes opportunity costs to farmers who provide land for demonstration plots. DFAT's guidance on measuring ADR indicators allows programs some flexibility about what is counted. MDF (Box E.4) also identified lessons related to the specificity and clarity of indicators, which effects measurement.

Programs rarely analyse the relationship between headline indicators. AIP-Rural's Strategic Review Panel<sup>43</sup> noted that the averages summarised in the Activity Completion Report masked a wide range of performance. CAVAC's sector monitoring group (2011) observed that 'the

---

37 It is also unclear how total investment is tracked, particularly when additional funding is sourced within DFAT. For instance, the MDF I was allocated additional monies from the Gender Equality Fund managed directly from DFAT Canberra.

38 Each implementing organisation provides its own report of aggregated results (not intervention level results) which is then used to prepare a combined AMENCA 3 program report.

39 AIP-Rural ACR, October 2018

40 PRISMA refers these to 'planned' comparing them to actuals. It is not clear if these are the same 'numbers' that were included in the AIP-Rural design or agreed at the commencement of the program.

41 AIP Rural also recorded increased income by household while the ADR indicators for income relate to poor men and women.

42 PRISMA 2 investment design document.

43 PRISMA SRP Aide Memoire, November 2018



indicators [CAVAC's four core indicators] are not sufficiently comprehensive to enable stakeholders including DFAT<sup>44</sup> to understand the impact of the program'.<sup>45</sup>

*Box E.4: Lessons from MDF I's internal review of its M&E system*

MDF I conducted an internal review of its M&E system to determine how well it served the program's purpose. This is good practice. The review found that changes in indicator definitions across the results chain acted like 'mini-hypotheses', defining how one change (or several them) were expected to lead to the next change. Over time, the indicator definitions were increasingly simplified and shortened but this also meant they became less precise and therefore less useful in measuring the outcomes they were expected to measure. To address this, MDF has developed a more comprehensive set of indicators, which provide greater specificity and consistency of key indicators.

## Targets

Numerical targets for quantitative indicators are used in different ways across programs.

In line with the DCED Standard, some programs (such as MDF and Strongim Bisnis) establish 'projections' once intervention activities are agreed with businesses or other organisations. The Standard considers 'targets'<sup>46</sup> to be the quantity of outcomes set at the design stage<sup>47</sup>, which may or may not be reflected in the DFAT contract with implementing organisations. Projections are simply revised targets based on actual interventions, which are themselves grounded in market analysis and a more thorough consideration of the breadth or depth of outcomes. Projections are periodically updated based on the program's experience and/or development of new intervention activities to provide a picture of the expected trajectory of quantitative results. Result projections can be useful to manage donor expectations, although the difference between targets, projections and actuals has caused some confusion among DFAT staff new to MSD programs. When targets are decreased, some stakeholders perceive that implementing organisations are not trying hard enough to reach the target. However, targets may also be adjusted upwards, as AMENCA 3's MTR recommended for the program.

GREAT uses the targets contained in the program design. Following the inception period, market analysis and development of interventions with businesses, NGOs and government, the program sought to adjust the targets downwards to better reflect what it felt was feasible. DFAT preferred to maintain the design-targets as an incentive for the program. While PRISMA also uses the Standard, it uses 'hard' or unchanging targets suggested by DFAT. There are differing perceptions regarding the advantages of hard targets on PRISMA. Some perceive the targets to be an incentive that positively drives program performance. Others perceive that targets have led to PRISMA being overly focused on numbers at the expense of sustainable systemic changes. The AMENCA 3 MTR observed that the program's focus on high-level accountability (reporting against 20 indicators) had caused the implementing NGOs to give much more emphasis to quantitative outcomes than necessary, and at a cost to learning. It noted that the indicators can:

- distort market selection and engagement by pushing implementing organisations to focus on interventions that deliver short-term quantitative impacts; and

---

<sup>44</sup> Then AusAID

<sup>45</sup> CAVAC I evaluation

<sup>46</sup> MDF uses the term estimates instead of targets

<sup>47</sup> Not all program designs have targets included



- de-prioritise some of the more fundamental investments needed to establish critical market functions or to improve inclusion.

The type and trajectory of results against headline indicators for MSD programs may differ from direct delivery programs. Where direct delivery programs may be able to achieve notable quantitative results in a short amount of time, e.g. number of men and women with access to safe water, the quantitative results normally attached to MSD programs, that is increased income and jobs created, are expected to take longer to achieve. In the initial phases of MSD programs, considerable effort is spent on conducting market analysis and develop relevant partnerships. The differences in results trajectories reflect the different approaches. Some DFAT staff feel the MSD timeframe does not fit well with DFAT's current organisational systems such as annual results reporting, which may create expectations that headline results will be achieved every year.

### Portfolio level

- Overall, MSD programs find it challenging to present portfolio level performance, finding a balance between results against headline aggregated indicators and information on changes at the mid and lower levels of theories of change, that satisfies DFAT.
- MSD programs generally use a variety of approaches to report more in-depth information about portfolios. This includes: selected case stories; using ratings, e.g. red, amber, green ratings, at intervention level; reporting against additional, largely quantitative, indicators.
- There is often little analysis of portfolio performance and variation within programs or sectors. The use of aggregated results and averages masks a wide range of variations.
- Short case stories or vignettes are commonly used to communicate the achievement of intended positive outcomes. There appear to be fewer cases that illustrate failures, lessons learned from these and adaptations made even though failure is a key aspect of MSD programs.

Beyond headline indicators it might be expected that progress reports provide a portfolio-perspective of progress. Program progress reports provide some portfolio level information e.g. expenditure against budget, summary of the efficiency and effectiveness of management operations.

Programs use a range of strategies to present performance across their portfolios: 1) using ratings to indicate progress or performance; 2) supplementing headline indicators with short positive case stories; and 3) reporting a wider range of quantitative indicators. Examples of each are provided below.

### Ratings approaches:

- TOMAK's progress reports include an annex that lists all its interventions with an assessment of whether that have been completed or are on track, delayed, being reassessed or cancelled, along with brief comments.
- AIP-Rural progress reports summarise perceptions of intervention quality (assessed according to various criteria) and whether they are being pushed or dropped.

### Additional quantitative indicators:

AMENCA 3, TOMAK, MDF and GREAT monitor a range of additional quantitative indicators. MDF uses aggregated intermediate or 'leading' indicators, such as the value of additional market transactions, since increased income and jobs are viewed as 'lagging' (longer-term) indicators of success. TOMAK and GREAT also link indicators to key evaluation questions.

The CAVAC II MTR noted that progress reports should include more aggregated, and less ‘granular’, data at lower levels of the theory of change, that is outcomes relating to behaviour change of the private sector and FWUC (Farm Water User Committee) functionality.

This approach has not always met DFAT’s needs. For instance, the AMENCA 3 MTR recommended less breadth and more depth to bring to the surface the numerous ‘hidden results’ in the program.

### **Short positive case stories:**

Headline indicators are often supplemented by short descriptive case stories to illustrate the human dimensions of results achieved. These tend to be read as more anecdotal since supporting evidence is often not provided in progress reports<sup>48</sup>. The lack of evidence does not necessarily mean that the evidence does not exist. The evidence may be in other documents, such as in-depth case studies, and not included in progress reports due to requests for progress reports to be succinct and/or within page limits.

Mostly, the examples highlighted in progress reports are positive cases. While failure is a normal part of MSD programs, it is more difficult to find detailed examples of where efforts did not work and why. TOMAK’s experience with its learning and development platform (LDP) highlights this challenge. It notes ‘While it has been challenging to get development partners that are more used to operating in a competitive environment to open up on their problems as well as successes, indications are that the LDP approach is working and contributing significantly to the intended learning culture. The cross-visits, in particular, are providing a great opportunity for partners to directly observe, question and learn from each other’s field programs, and are attracting a high level of interest from participants.’<sup>49</sup> There may be a range of reasons why it would be difficult for MSD programs to report failure. Programs may be able to be incentivised through more emphasis on learning than accountability, and broader appreciation that working out what works for whom in what situations takes time. Positive stories likely receive positive responses from DFAT as it meets their communication needs – while full or partial failures provide important learning opportunities. Further, if programs do not have ‘failures’ it may be reasonable to ask how well they are pursuing additionality and pushing boundaries.

### **Market systems level**

- Long-running MSD programs have recorded examples of systemic change in very different market contexts.
- Defining systemic change is a challenge. Most programs refer to systemic change as change in the underlying causes of market system performance that leads to a better-functioning, more pro-poor market system.
- Most programs also use the Adopt-Adapt-Expand-Respond (AAER) framework to conceptualise change pathways, with varying degrees of success. The framework has yet to be refined to increase its relevance to different situations, such as shallow markets or different types of impact pathways that underpin, for example, advocacy efforts. This is a potential opportunity as a learning focus for future MSD programs.
- Measuring systemic change is a second, related, challenge. Largely, programs are not systematically analysing how constraints (functions, rules, relationships) have changed and what effect this has had on the inclusivity of market systems. Some programs use quantitative

---

<sup>48</sup> As noted earlier, DFAT M&E Standards highlight that credible supporting evidence is provided.

<sup>49</sup> TOMAK (2019). Six Month Progress Report (July–December 2018)

indicators relating to beneficiary impact as a proxy for systemic change. This risks undermining the MSD approach by skipping over the assessment of systemic change because it is difficult, messy and complex. Qualitative case studies are also used by programs to illustrate examples of systemic change.

Describing and communicating systemic change seems to be one of the more problematic elements of the MSD approach. This section examines the challenges from two perspectives: how programs understand systemic change and how they assess such change.

#### Defining and understanding systemic change

As DFAT's MSD portfolio matures, DFAT's expectations of seeing examples of systemic change understandably also increases. A peer review of the MDF II design noted that the request for tender to manage phase II 'will need to emphasise the importance of an even stronger focus on generating systemic change over the 5-10 year lifespan of MDF II'. DFAT programs use the DCED's definition of systemic change:

*Change in the underlying causes of market system performance that leads to a better-functioning, more pro-poor market system (DCED, 2017).*

Many MSD programs struggle to provide a consistent and coherent conceptualisation of the 'consequential' changes or market system change that provides the *raison-d'être* of MSD. Commentators (Cunningham and Jenal, 2016; Lomax, 2019) have noted that despite more than a decade of MSD programs there is not precise definition of what systemic change is, while the use of the attributes of scale, sustainability, and resilience to define systemic change means that practitioners focus on, and measure these, rather than focusing on what makes change 'systemic'.

Programs' conceptualisation of systemic change influences where they focus, what gets assessed and how. Describing systemic change, or changes in market systems, seems to be one of the more problematic elements of the MSD approach. Few programs report on the underlying causes of market system performance and describe what has changed.

This section first summarises how MSD programs conceptualise systemic change and secondly how they measure it.

#### Conceptualisations of systemic change in use

Most programs (including CAVAC I and II, AIP-Rural and PRISMA II and Strongim Bisnis) have adopted the widely used Adopt-Adapt-Expand-Respond (AAER) framework (see Box E.5), which is founded on generalised theories regarding the diffusion of innovations.

**MDF I** argued that pathways to change may look very different in different contexts and pathways to systemic change may be difficult to anticipate given constraints may be founded on 'intractable issues of power, culture and social norms'. This insight is borne out of MDF I's experience implementing the MSD approach in shallow and fragmented markets where 'crowding in' was unlikely to occur because (see Box E.6):

- transaction costs are very high relative to transaction values in shallow, small and highly fragmented markets.
- the scale of impacts might be very limited.
- there might be multiple constraints – calling for multiple innovations – bearing on the problem of making market systems work better and be more inclusive.
- systemic changes might be achieved that lead to no copying or crowding in, since one business alone might be enough to fill a missing support function.

**Adopt:** A market player(s) has successfully adopted a behaviour/practice change to the ultimate benefit of the poor producer/worker/consumer, recognises the value of continuing with these changes irrespective of program inputs, and has accordingly made plans to invest in upholding these changes and cover any associated recurrent costs.

**Adapt:** The market player(s) that adopted the behaviour/practice changes pioneered during the pilot has made qualitative and/or quantitative investments that allow them to continue with or augment changed practices, without program support. These actions, independent of the program, constitute an 'acid test' for whether pro-poor outcomes will sustain at any level.

**Expand:** A number of market players similar to those that pioneered the pro-poor behaviour/practice changes have adopted comparable changes - either pure copies or variants on the original innovation - that are upheld without program support.

**Respond:** The emergence and continued presence of the pro-poor changes have incited market players in supporting systems to react to the new market reality by re-organising, assuming new/improved roles, developing their own offers, or moving to take advantage of any opportunities that have been created. The response enables pro-poor behaviour/practice changes to develop further, or evolve, and indicates a new capability within the system, suggesting that it can and wants to support pro-poor solutions to emerge and grow.

**Strongim Bisnis** is planning to review results across the whole results chain. This will include outlining the key market systems changes against the results chains and examining any divergence from the original strategy, sustainability of change, breadth and depth of changes within the market system, challenges and lessons learned.

MDF I learning on pursuing systemic change in different markets:

'A key lesson learned from MDF's systemic change approach under Phase 1 was that large markets, such as Pakistan and Sri Lanka, could rely more on 'crowding in' to achieve systemic change, as market players could see what is working and replicate it. MDF found that most 'off-the-shelf' systemic change frameworks were based around this 'crowding in', and were not applicable to the thin markets in which MDF operated in PNG, Fiji and Timor-Leste. Therefore, the team developed new tools for these markets, which could be adapted to a variety of contexts, including a broader focus on sector wide opportunities, greater flexibility and investment in infrastructure and in-house training before facilitating business engagement.'

Specifically, in relation to the "R" ("respond") of the AAER framework, the program will monitor "respond" behaviour from non-competing actors (such as government or markets that are linked with the target sector), to identify broad shifts in the overall market to accommodate and support the changes stimulated by the program. "Respond" normally occurs in the supporting

<sup>50</sup> Developed in 2011 and part of the Springfield Centre training since 2012. Nippard, Hitchins and Elliott (2014), argue the AAER does not represent chronological degrees of systemic change. Change pathways may be A+A+E+R but also could follow other pathways such as A+R. but there have been debates about whether: 'adopt' alone is sufficient evidence of pro-poor changes being 'systemic'; the A and A is systemic when the MSD attributes of sustainability and scale attributes are clearer at the E and R elements.

<sup>51</sup> MDF I Activity Completion Report

functions and rules of the market and demonstrates a systemic adjustment in the market that will underpin sustainability of the changes, resilience of the system, and scalability.

**PHAMA Plus** defined systemic change as a gradual institutionalisation (including social normalisation) of changed behaviours across a market system, including household producers, private market players, public agencies, industry bodies and other industry stakeholders. The program sees systemic change as change in the market institutions – or rules of the game – that govern how markets operate. Formal rules executed by government agencies are key but informal rules that influence how market players and other industry stakeholders behave are also important. PHAMA Plus<sup>52</sup> notes that system change can also be a change within specific private or public market players or in relationships among market players. Market system change for PHAMA Plus then leads to improved sector performance (e.g. productivity, sales, cost reduction, efficiency, and competitiveness of producers and enterprises) as a result of the application of innovations and increased sales through the new and existing market access pathways.

Some other programs do not provide an identifiable conceptualisation and/or it is not clear how systemic change is assessed as part of their monitoring, evaluation and learning systems. For instance:

- TOMAK's guiding strategy notes TOMAK will 'place a systemic 'lens' on its impact logic for the program to achieve scaled, sustainable changes in TOMAK's target communities. As a result, monitoring efforts will focus on changes in market systems and systems surrounding nutrition sensitive agriculture and social behaviour change communication, as well as the ultimate impact for beneficiaries. However, it is not clear from the questions and indicators in the MREL framework how systemic change or changes in market systems will be assessed.
- GREAT describes expected systemic changes as outcomes such as improved market interaction with women-led households, groups or cooperatives. Specifically, the design highlights income increases and women's roles in decision making:
  - More women - particularly younger women and those from ethnic minority groups - will have greater access to skills and jobs (beyond casual and part time work) and greater income.
  - Women led or co-managed households, groups or cooperatives increase rapidly and make more money as vendors, suppliers, and processors increase their local sourcing, their reach, their value adding, and/or the diversity of product sought, because they can see the benefits of working more closely with women suppliers.

#### *Other conceptualisations of market systems change*

As noted earlier, programs' definitions of systemic change conflate the desired characteristics of the impact that systemic change is hoped to lead to with the change itself.<sup>53</sup> For instance, increased small holder farmers income is viewed as being the same as systems change.<sup>54</sup>

Systemic change often seems to be viewed as synonymous with scale of change. The trajectory of results – outreach and increased income – presented by some MSD programs, such as AIP-Rural and MDF, seem to support this. However, it also seems possible that there are alternative models of or pathways to impact that contribute to market systems change (see Figure E.3 below).

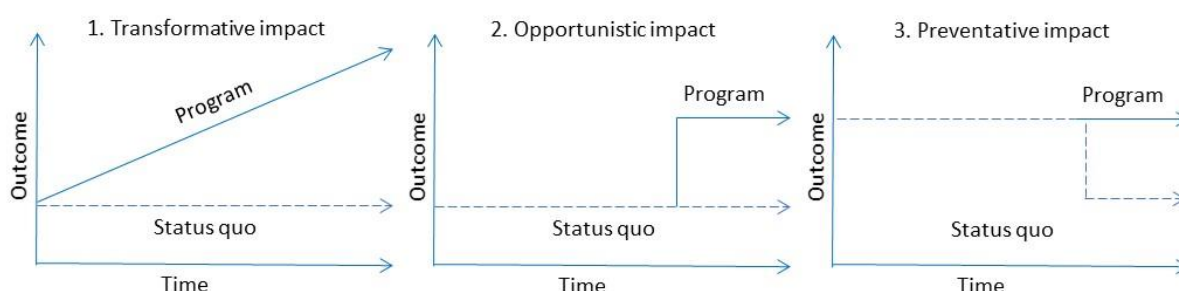
---

<sup>52</sup> PHAMA Plus MRM Plan

<sup>53</sup> JenaI and Cunningham (2016)

<sup>54</sup> Programs' measurement of increased income may be reliable but it is not a valid measure of systems change.

Figure E.3: Models of impact<sup>55</sup>



For instance, MDF supported the establishment of a private sector tissue culture business in Fiji, with a rapid propagation capacity that can significantly speed up the reestablishment process for crops such as papaya and banana after storms. For this intervention the preventative impact model may be a better reflection of the systemic change than a transformative impact model. The intervention may not generate large numbers of increased income but rather may prevent at least some decreased income, enabling farmers to more quickly re-establish crops. Similarly, Davis (Davis, 2016) notes that barriers to scale often lie outside the business, but reform to rules and regulations does not always lead to scale.<sup>56</sup>

The examples of systemic change from CAVAC, MDF and PRISMA suggest that these changes occurred in a non-linear fashion. These different non-linear types of change might be happening more often within programs but are not being 'seen' as systemic change and/or reported—this is a missed opportunity for learning.

#### Assessing systemic change

DCED Standard audits confirm that assessing systemic change remains a challenge for MSD programs. After the challenge of conceptualising systemic change, the second challenge is how to monitor and identify systemic change, when such changes are, by definition, beyond a program's direct control and influence.

Two of the longest-running DFAT MSD programs have approached the assessment of systemic change as follows:

- The approach adopted by the PRISMA component of AIP-Rural<sup>57</sup> is founded on the AAER framework. It measures systemic change by trying to identify the number of new actors that fell into the "respond" category. This was done by disaggregating outreach numbers, differentiating between those directly reached compared with indirectly reached, the latter taken as an indicator of systemic change. On this basis, by the end of 2018, PRISMA assessed that 36.7% of total outreach had been achieved through systemic change. This approach

<sup>55</sup> Ruedy (2018).

<sup>56</sup> This is one lesson. Another seven are: Lesson 1: Getting to Scale takes time; Lesson 2: Only a few interventions are likely to get to scale; Lesson 3: Early on, check the business case; Lesson 4: The quickest route to scale is through just one 'big actor', but it's risky; Lesson 5: Do not take the demonstration effect for granted; Lesson 6: If the demonstration effect fails, to get to scale programs may need to directly support second movers (often as or more intensively as the first movers); Lesson 7: Creating a viable fee-based service market is harder than it sounds. If the expected change does not materialise, it suggests the underlying cause may not have been addressed.

<sup>57</sup> AIP-Rural Activity Completion Report (2018)



provides some information on the breadth of systemic change, but not the depth,<sup>58</sup> and could be conflating the impact of the change with the systems change itself. Reporting does not make clear what ‘underlying causes’ of market system performance have been addressed – which poses questions about sustainability of the claimed systemic change.

- MDF I expanded on the AAER framework and developed a process for mapping pathways to systemic change, focusing on six parameters for systemic change (autonomy, sustainability, resilience, inclusiveness, scale and women’s economic empowerment). MDF found the approach difficult to operationalise for capturing systemic changes. Rather, they used case studies to capture depth once they were confident systemic change had surfaced.

The examination of MSD program reports conducted for this review suggests that part of the challenge in assessing systemic change is that program’s focus on the constraints to be addressed, seems to evaporate once they have been identified and interventions have been developed. Few program reports provide analysis of how the targeted constraint or underlying cause of market under-performance has changed over the course of the intervention. Rather, reports focus on measuring and reporting beneficiary level increased income. The MDF Advisory Group and AMENCA 3 MTR have commented on the depth of market intelligence program team members have on constraints and other issues – and on the fact that reporting typically fails to reflect this intelligence.

This challenge is not unique to DFAT’s MSD programs. Koh et al (Koh et al, 2014) highlight (Figure E.4) that funders and intermediaries often skip over evaluating systemic change because it is messy, complex and difficult. This therefore represents a ‘missing middle’ in the M&E of MSD programs. One program offered the following analogy: the program has a street-level view, we are down on the street, working with and around pedestrians, potholes, cars etc. The program doesn’t have a higher-level view of what is happening at the map-level.

## Intervention level

- There are numerous examples of where programs are using intervention level information to manage interventions.
- Programs have in-depth information at the intervention level, but this is often not captured in progress reports. Programs often discuss intervention level progress and challenges with DFAT Posts at regular meetings. These may not be documented discussions to provide a record for accountability and learning, and to share with other stakeholders (e.g. new DFAT program staff).
- Given that DFAT generally prefers succinct reports, it is impractical to include in-depth information on the progress of every intervention in regular progress reports. However, programs need to find ways to increase DFAT’s understanding of intervention level progress and results.

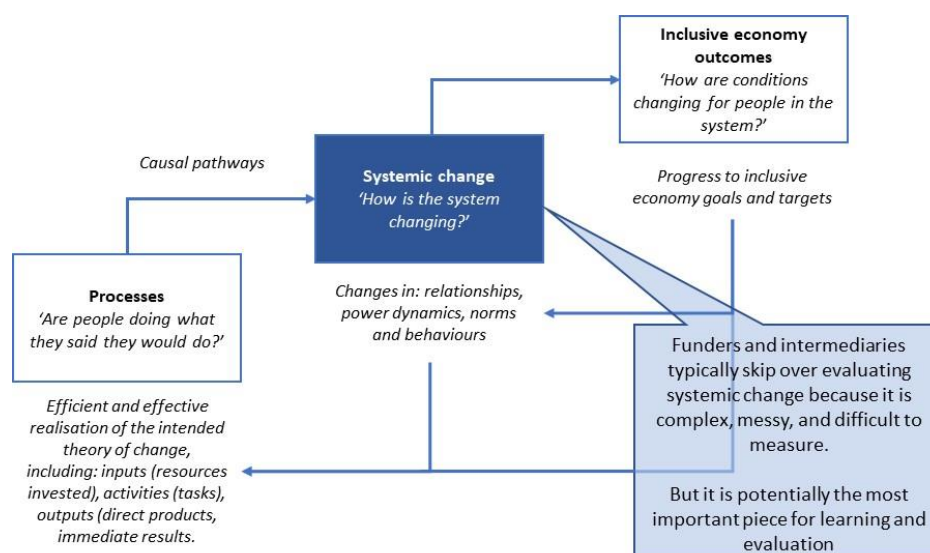
*Programs using the DCED Standard link quantitative indicators and qualitative questions about how and why change happened to intervention-level theories of change (results chains). This draws on a range of information collected and made use of by program teams in their day to day work and periodic portfolio, sector and intervention review processes<sup>59</sup>.*

---

<sup>58</sup> The ACR also notes PRISMA developed a tool (adapted AAER and ‘will and skill’ matrix) to assess and promote behaviour changes in partners, although findings and analysis are not outlined in the report. The reports includes a list of interventions, each with a systemic change score, but does not provide an explanation of the scores.

<sup>59</sup> The DCED Standard notes results chains should be updated at least annually. MDF I and II and AIP-Rural and PRISMA 2 have six monthly review processes.

Figure E.4: Elements of an adaptive monitoring and evaluation system for inclusive markets<sup>60</sup>



Various program reports and documents illustrate that programs use information generated through their M&E systems. For example:

- The AIP-Rural DCED Standard audit noted ‘The program team regularly uses information on results to improve interventions, make decisions on terminating or expanding interventions, develop sub-sector strategies and shape the program portfolio’.
- The CAVAC II MTR noted that monitoring information was used by the team.
- The AQC for MDF Pakistan provides a specific example e.g. Data from the M&E system has been central in guiding MDF's engagement across the meat, silage and dairy value chains. Regular impact studies and data collection allows for managers to evolve partnership models to make them more pro-poor. For example, MDF's initial approach to work with meat fattening farms (Oasis) was not delivering as expected as it mainly benefitted relatively large farmers and lacked incentives to reach out to small farmers in a fragmented meat supply chain. The data captured through the system allowed managers to look for alternative model (Al-Saffah) that was more inclusive and provided 360-degree solution to small livestock owners. Similarly, monitoring silage partnerships allowed MDF to identify opportunities to support women in livestock and remove information barriers.<sup>61</sup>

Some intervention-level, and market-level, information is communicated to DFAT Posts during regular meetings. DFAT appreciates MSD programs ‘market intelligence’. Beyond these meetings, it appears there is less documented analysis of the effectiveness of specific interventions. Some programs document intervention-level case studies, although as noted previously these tend towards positive examples of change. LIFT, however, conducts evaluations of at least some of its projects, e.g. Making Vegetable Markets Work, which provide more comprehensive documentation of progress, achievements, challenges and lessons. Strongim Bisnis plans to conduct some intervention-level evaluations utilising data and analysis from internal monitoring activities and impact assessments to examine the relevance, efficiency, effectiveness and

<sup>60</sup> Koh, H et al (2016).

<sup>61</sup> AQC, 2019; DFAT email correspondence



sustainability of interventions. These will normally occur following an intervention lifecycle, including broader systemic change, if it occurs.<sup>62</sup>

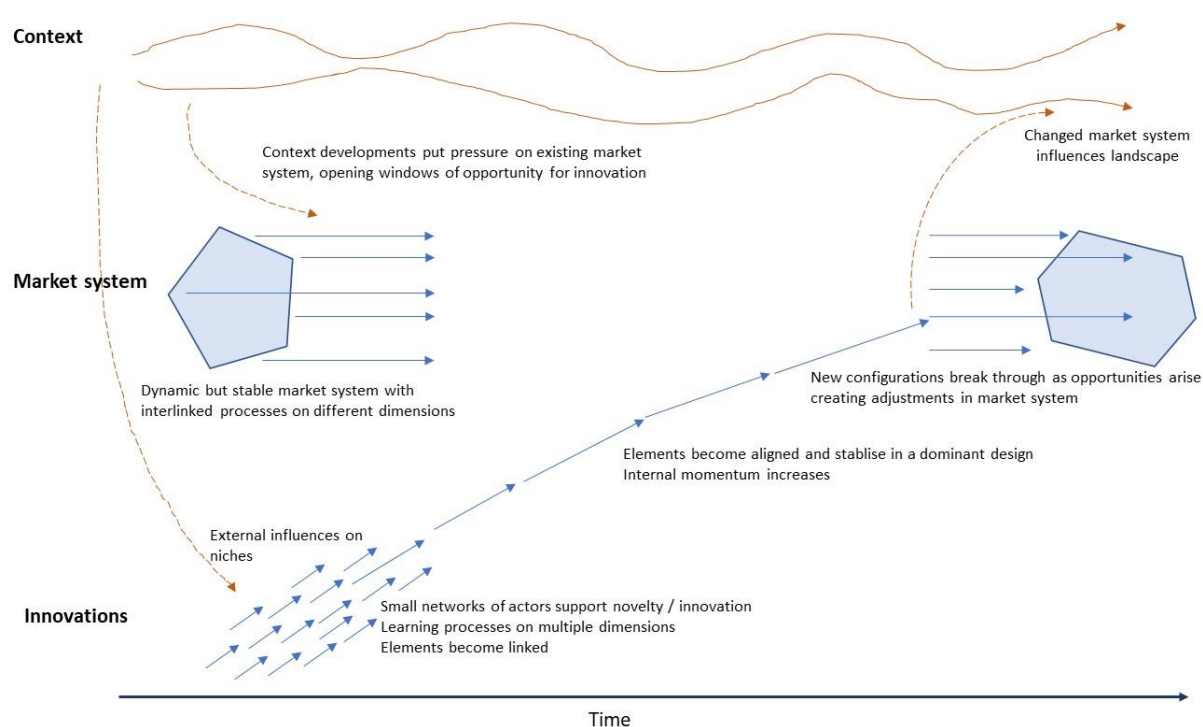
Given the size and complexity of most MSD programs, it is not feasible to document in-depth analysis of the effectiveness of every intervention. However, the current reporting practices do not provide visibility about what is happening at the intervention level, and perhaps contributes to challenges of assessing market systems. DFAT Post program managers find it difficult to understand what programs are doing and how efficient, effective and sustainable efforts are. Greater visibility of intervention effectiveness and market systems changes could contribute to greater program learning as well more cross-program and DFAT learning. However, increasing the length of progress reports to increase visibility does not seem feasible and therefore other avenues of communication need to be examined.

## Assessing sustainability of changes

- Few of the reports examined in this review provided evidence on the sustainability of intervention or market systems level changes.

Closely linked to the notion of systemic change is sustainable change, another key characteristic of the MSD approach. Ripley (2019, citing Taylor, 2016) emphasises that an uptake of a new product or service cannot be considered systemic change unless it alters the fundamental system structure or creates a 'new normal'. Figure E.5, adapted from Geels (2011), illustrates the pathway from an innovation to changes in market systems and contexts.

Figure E.5 Pathway from innovation to changes in market systems



Overall the sustainability of MSD outcomes (whether at an intervention level or market systems level) cannot be determined from the reports reviewed. The MTRs and evaluations on individual

<sup>62</sup> Strongim Bisnis MRM Plan (2019)

MSD programs have also found assessing sustainability difficult. For instance, the AMENCA 3 MTR (2019) noted:

*It was difficult for the MTR to determine the sustainability of the inclusive market opportunities as these, although promising, were still early-stage. The program level data shows significant achievements, yet most have been delivered in the last twelve months. Furthermore, the longer-term impacts are difficult to predict, given the multitude of market distorting factors and the security concerns in the Palestinian Territories, all of which are beyond AMENCA 3's control.*

Some MSD programs have included specific sustainability-related evaluation questions in their M&E plans that may make it easier in the future to understand sustainability. For instance, Strongim Bisnis plans to address questions such as:

- How likely is the continuation of market services to the poor in Solomon Islands without external support? How effective was the systemic change approach in achieving sustainability? Were there other factors affecting sustainability of results?
- Will market systems be able to adapt and adjust to environmental changes, so that benefits will continue for beneficiaries?

The PRISMA 2 design emphasises that the new phase will work to strengthen the sustainability of results achieved under AIP-Rural:

*strengthen sector adaptability and resilience so that sector actors have the capability to anticipate, mitigate or respond to opportunities and threats in the future ... PRISMA 2 will work with sector partners to ensure that the innovation is supported by other actors in the sector. This will mean strengthening other key functions or rules that support the innovation introduced by the private sector.*<sup>63</sup>

The CAVAC I evaluation noted that the primary outcome was the delivery of 20 rehabilitated irrigation schemes with varying levels of service and sustainability. However, the irrigation schemes were not implemented using an MSD approach.

The DCED Standard for Results Measurement notes interventions should be monitored for two years beyond the end of direct support. Programs may end before this date, and monitoring ceases. This suggests that it may be useful to undertake further analysis from existing information to draw insights on sustainability. For instance, disaggregated information on increased beneficiary income at intervention level could be used to further analyse the degree to which increases are sustainable, for whom and over what time. It would also be necessary to assess the financial sustainability of businesses who adopted innovations as this is also a driver for sustained income increases.<sup>64</sup> This beneficiary level analysis would also need to then be put in a broader analysis of changes in the market, the extent to which market constraints have changed and how, to consider the likelihood of sustainable changes.

Assessing sustainability requires a broader focus than on beneficiary income – it requires analysis at the market or sub-sector level, partner (e.g. business, government) and beneficiary level. The

---

<sup>63</sup> PRISMA II design

<sup>64</sup> The DCED Audit for Sri Lanka noted that while the program had developed indicators related to the financial viability of businesses these, in many cases, related only to sales volume and/or value. Additional key indicators relating to costs, profit and business viability or qualitative questions on the likelihood of sustainability are also needed.

type of analysis would also depend on the market constraints being addressed and the types of interventions. Aggregated quantitative indicators on increased beneficiary income do not provide insights into the sustainability of income for the more mature program interventions.

## Assessing value for money

### Key points:

DFAT's expectations regarding the purpose of and approach to VfM assessments is unclear and as such MSD programs use a variety of approaches to assess VfM, none of which convincingly demonstrates VfM or its absence. Few programs have plans to systematically collect and analysis information across all aspects of VfM to support an overall judgement of VfM.

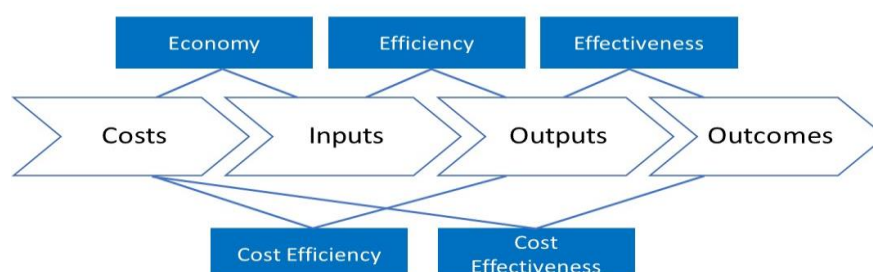
Only one MSD program makes a direct link to DFAT's VFM principles (PHAMA and PHAMA Plus). Programs emphasise and measure different elements of VfM, which normally relates to the easier to measure dimensions of economy or efficiency.

As drivers of costs and benefits differ across the various contexts in which programs operate, cross program comparisons are not very useful.

Most programs in the MSD portfolio do not present a clear understanding of value for money or articulation of what value for money looks like for their program, and on which to base a judgement about the extent to which the program is delivering value for money. Only one MSD program (PHAMA Plus) directly references DFAT's VFM principles. Most programs reference the more common VfM Framework (Box E.7). Most programs focus on:

- The relationship between cost and outcomes (cost-effectiveness). For example, MDF and PRISMA.
- The relationship between cost and efficiency (cost-efficiency). For example, AMENCA 3, CAVAC II and PRISMA mid-term reviews.

Box E.7: Common value for money framework



Value for money development should be:

**Economic:** inputs have been procured at the least cost for the relevant level of quality.

**Efficient:** efficiency is generally defined as considering the value of outputs in relation to the total cost of inputs (at the relevant level of quality).

**Effective:** achieving program outcomes (and impact) in relation to the total cost of inputs.

LIFT and Strongim Bisnis provides the clearest articulation of their understanding of VfM (see Box E.8).

### VfM assessment purposes and approaches

Three common purposes of value for money assessments are<sup>65</sup>:

- Putting the right management processes in place to enable managing for value for money
- Comparing value for money to drive improvement and inform decision-making
- Demonstrating value for money through evaluation.

DFAT's VfM principles (2015) emphasise managing for VFM. DFAT's *Making Performance Count* (2014) emphasise managing for VFM and comparing VFM for decision-making.

Individual MSD program's approaches to VfM assessment seem driven by their interpretation of the purpose of VfM and what DFAT wants. Most programs understand VfM as either comparing VfM to drive improvement and inform decision-making or demonstrating value for money through evaluation. Programs within DFAT's portfolio are conducting some assessments of value for money. For instance, ex-ante assessments are made during intervention design and selection processes. Programs assess budget and likely effort to deliver expected returns (e.g. expected number of beneficiaries and/or increased beneficiary income) as part of their decision-making processes.<sup>66</sup>

*Box E.8: Understanding of value for money in LIFT and Strongim Bisnis*

LIFT currently defines value for money<sup>67</sup> as:

- Economy: Contract for highest quality inputs at the best price
- Efficiency: How well inputs are converted to the outputs and transferred to the beneficiary
- Effectiveness: How well outputs are converted to beneficiary outcomes, and sustained
- Equity: Distribution of benefits across all beneficiaries: gender; region; wealth quintile; ethnicity.

LIFT also explains how it manages for value for money at various stages of the project management cycle – design, appraisal, inception and delivery. There are some slight variations in the approach used for different program components (agriculture production versus micro-finance).

Strongim Bisnis also uses the 4-E framework. It proposes to use a variety of indicators to assess each component. Examples are:

- Economy: Conversion of money into inputs (e.g. consultant time)
- Efficiency: Conversion of "inputs" into "outputs" (e.g. Strongim Bisnis partner launches online package for eco-stays in SI)
- Effectiveness: Conversion of "outputs" into "outcomes" (e.g. Eco-stay MSMEs receive more visitors)
- Cost Effectiveness: Conversion of "outcomes" into "impact" (Increase local income through tourism spend)
- Equity: The extent to which particularly poor or marginalised groups share in program impacts

<sup>65</sup> Based on Bond's (a broad network of UK-based international development organisations) articulation of VfM assessment purposes. See: Bond, 2012.

<sup>66</sup> Detailed guidance documentation was not reviewed across all programs. However, from documents reviewed and interviewed this is considered typical across projects. It is not possible to determine the degree of rigour applied to ex-ante assessments during the timeframe of this review.

<sup>67</sup> DFID is LIFT's largest funder and therefore has likely influenced the adoption of this framework. DFID uses economy, efficiency and effectiveness. However, a review by the Independent Commission for Aid Impact (ICAI, 2018) on value for money in DFID programs also added equity. This has influenced some DFID-funded programs to add this criterion to their assessments.

A few in-depth assessments of cost-effectiveness have been undertaken where program components lend themselves to more typical economic and/or financial analysis, e.g. the CAVAC II irrigation component impact study used quantitative and qualitative methods (although this component does not apply a market systems development approach).

Only one example was found where a program, LIFT, was seeking to improve how it assesses value for money. LIFT also has the broadest approach, providing information related to all three purposes – managing for VfM, comparing VfM to drive improvement and inform decision-making, and demonstrating VfM. PHAMA Plus and Strongim Bisnis focusing on managing for value for money by putting the right processes in place and comparing value for money to drive improvement and inform decision-making.

While none of the programs are systematically and thoroughly assessing value for money, this seems common to MSD programs. For instance, the 2019 BEAM Exchange Evidence (Conroy and Kessler, 2019) observed:

*... individual MSD initiatives sometimes generate very impressive impacts and hence the ratio of benefits to costs is likely to be positive. It is less clear whether the same situation would apply for entire whole programmes. Only 14% of new additions to the evidence map assessed VfM and the information in documents is very limited. The review found only half of these were already<sup>68</sup> demonstrating clear evidence of programme-wide benefits exceeding costs, but was cautious about drawing conclusions.*

Table E.5 summarises the range of assessment approaches and highlights examples across the portfolio of which programs are using specific approaches. Additional information is provided in Addendum 2 of this annex.

Table E.5: MSD programs approaches to assessing VfM<sup>69</sup>

Purpose	Possible approaches	Examples of VfM cited by programs or VfM questions programs intend to answer	
<b>Managing for value for money by putting the right processes in place</b>	Implementing transparent contracting procedures	PHAMA Plus Strongim Bisnis	e.g. A panel of pre-qualified international and locally-based STA experts was established, allowing the program to rapidly mobilise for specific assignments whilst reducing recruitment costs <sup>70</sup> e.g. Robust procurement processes and fiduciary risk management <sup>71</sup>
	Focusing on maximising economy in operational support	PHAMA Plus Strongim Bisnis	e.g. VfM was achieved by maintaining the full cohort of Phase 2 operations and finance staff. <sup>72</sup>
	Monitoring expenditure and avoiding corruption.	Strongim Bisnis	e.g. Funding guidelines refer to systems to prevent fraud and corruption
	Generating evidence-based theories of change for a program and ensuring that budgets allocated are	PHAMA Plus TOMAK	e.g. To what extent does PHAMA Plus consider if and how each intervention is likely to contribute to program outcomes when choosing and designing interventions?

<sup>68</sup> Program implementation was still underway

<sup>69</sup> Purposes and approaches based on Bond (2012)

<sup>70</sup> PHAMA Plus Inception Report (2019)

<sup>71</sup> Strongim Bisnis MRM Plan and Manual (2019)

<sup>72</sup> PHAMA Plus Inception Report (2019)

Purpose	Possible approaches	Examples of VfM cited by programs or VfM questions programs intend to answer	
	commensurate with predicted outputs and outcomes.		To what extent is TOMAK making appropriate use of resources to achieve its outcomes? <sup>73</sup>
	Ensuring appropriate monitoring, evaluation and learning systems for programs	PHAMA Plus LIFT, Strongim Bisnis	e.g. To what extent is PHAMA Plus using evidence from its MRM systems to inform decision-making and improvement?; Pathways to scale mapped for each activity. <sup>74</sup> LIFT tracks the percentage of projects on track with expected implementation rate
	Linking resources to performance data e.g. based on need, alignment with strategic priorities	LIFT	e.g. Funding windows introduce competition into project selection and provide an opportunity to modify the LIFT portfolio in light of emerging evidence and shifting priorities. <sup>75</sup> Making Vegetable Markets Work Evaluation Question: To what extent has value for money been used to inform management decisions? <sup>76</sup>
<b>Comparing value for money to drive improvement and inform decision-making</b>	Comparing cost per beneficiary for a service /outcome	LIFT, MDF II, PHAMA Plus, PRISMA, Strongim Bisnis <sup>77</sup>	Narrow focus is on cost-efficiency and cost-effectiveness e.g. The average cost per farm household with a net-income increase due to the intervention; ratio of intervention costs to value of increased income
	Comparing stakeholder perceptions of the effectiveness of different activities relative to costs		No examples identified
	Benchmarking staff and consultant salaries across the sector.		e.g. Average consultant rates and expenses (total consultant spend ÷ total consultant days) <sup>78</sup>
	Comparing the expenditure and efficiency of different methods of achieving similar outputs		No examples identified
<b>Demonstrating value for money through evaluation</b>	Cost benefit analysis	CAVAC I and II	e.g. Thorough analysis of costs of benefits CAVAC II's irrigation impact study assessed the financial internal rate of return <sup>79</sup> for three irrigation schemes, using a 30-year time horizon. It also undertook qualitative analysis to understand benefits that could not be easily monetised. The CAVAC I MTR also used a cost-benefit analysis comparing costs of the intervention to

<sup>73</sup> PHAMA Plus MRM Manual (2019)

<sup>74</sup> Strongim Bisnis MRM Plan and Manual (2019)

<sup>75</sup> DFID Annual Review for LIFT (2017)

<sup>76</sup> LIFT Making Vegetable Markets Work Evaluation (2018)

<sup>77</sup> Strongim Bisnis also measure women's and youth economic empowerment improvement rate (income) (Proportion of women and youth reporting increased access and agency)

<sup>78</sup> Strongim Bisnis MRM Plan and Manual (2019)

<sup>79</sup> Financial analysis concerns the analysis of costs and benefits for an enterprise. It differs from economic analysis that has a broader perspective and considers costs and benefits for broader society. Economic analysis considers different stakeholders views. Costs and benefits may not only be monetary and often include intangible benefits.

Purpose	Possible approaches	Examples of VfM cited by programs or VfM questions programs intend to answer
		increased farmer income (using a 10-year horizon and 6% discount rate)
	Social return on investment	No examples identified <sup>80</sup>
	Other evaluative approaches <sup>81</sup>	Selected LIFT projects <sup>82</sup> Evaluation of projected and actual project costs and benefits such as increased farmer incomes, increased use of technology, using mid-line and end-line data with examination of factors affecting change.

Only some programs make explicit judgements about a program's or interventions' value for money during implementation or ex-post. For example, various VfM-related indicators are reported but no interpretation or conclusion is provided, so that judgement seems to be left to the reader. It is expected that programs are making judgements in the day-to-day process of managing interventions or during portfolio reviews. This may be implicit. AIP-Rural's portfolio review process uses quantitative and qualitative indicators (outreach, benefit, rationale, value for money, quality of deal, quality of collaboration, systemic change potential, poverty, gender and environment) to rank interventions to supporting decision making regarding dropping, changing or improving interventions.<sup>83</sup>

Mid-term review and evaluation teams are more likely to make judgements about value for money, although sometimes cautiously. Most reports provide little information on the methodology used or assumptions underpinning the analysis. Most judgements of VfM assess some elements of value for money using quantitative data that is most easily collected and/or a few anecdotal examples of management practices to illustrate managing value for money. Likely influencing factors include the difficulties in assessing value for money, particularly less easily quantified values, as well as the limited time available to MTR and evaluation teams.

A range of issues appear to affect current practices across the portfolio. These include:

- DFAT has provided a lack of clear guidance on its expectations, the specific value for money questions DFAT would most like programs to answer, the quality of assessments expected, and the amount of time and resources that programs should be devoting to them.<sup>84</sup> This is considered the most significant issue.
- Few DFAT investment managers are making demands on programs to do more than they are currently doing, therefore providing little or no incentive to do more.

<sup>80</sup> Some programs refer social return on investment when calculating a ratio of cost to increased farmers income. However, they do not use the Social Return on Investment methodology that aims to monetise all benefits.

<sup>81</sup> For instance, 1) cost effectiveness analysis (see <https://www.betterevaluation.org/en/evaluation-options/CostEffectivenessAnalysis>); 2) OPM's guide to assessing value for money takes an evaluative approach whereby the criteria and standards of economy, efficiency and effectiveness are developed with specific reference to contexts, perspectives, values and trade-offs relative to the program concerned (See OPM and King, 2018)

<sup>82</sup> For example, LIFT Making Vegetable Markets Work, a \$4m / 3.75 year project (Evaluation report, 2018)

<sup>83</sup> See [https://aip-prisma.or.id/data/public/uploaded\\_file/2017-11-14\\_10-11-58am\\_PRISMA\\_and\\_SAFIRA\\_PRIP\\_S1Y17\\_Public\\_Version.pdf](https://aip-prisma.or.id/data/public/uploaded_file/2017-11-14_10-11-58am_PRISMA_and_SAFIRA_PRIP_S1Y17_Public_Version.pdf)

<sup>84</sup> A generic question such as 'what is the value for money' does not provide sufficient guidance to programs.



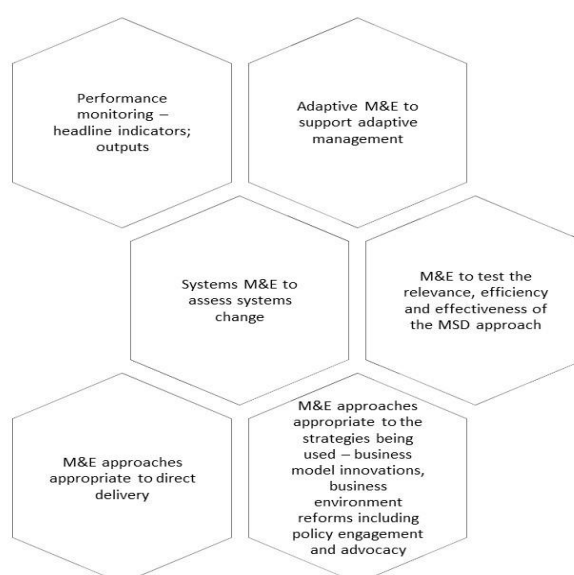
- DFAT already regulates most costs through the regulation of procurement (goods and services, including consultants).
- Implementing organisations consider managing for VfM is part of management and they do not see the value of additional frameworks designed to help ensure they manage for VfM and are able to demonstrate how they do so.
- Implementing organisations consider VfM indicators as a compliance requirement and/or do not see more thorough mid-implementation or ex-post VfM assessments as beneficial. Increased VfM requirements would be seen as an additional burden and cost that does not add value<sup>85</sup>.
- DFAT and programs are uncertain how to assess ‘value for money’ through a process that brings the different VFM elements together to make an overall assessment. In the absence of such a process, VfM assessment defaults to most easy to measure economic dimensions.
- Mid-term reviews and evaluations are undertaken in limited timeframes. VfM is not the core of these evaluations and there is limited time to assess VfM thoroughly.

A common VfM assessment framework based on DFAT’s principles could provide the basis for understanding expectations between DFAT and implementing organisations and thereby facilitate more efficient and constructive assessments and discussions on value for money. However, such a framework is only of value if DFAT is clear on its purpose and how it would use the information. Given there are different perceptions about how highly DFAT regards value for money assessments, a risk is that it is seen as another compliance requirement by implementing organisations.

## Final observations - M&E for the MSD portfolio and programs

As highlighted previously, DFAT’s and programs’ information needs are wide-ranging and varying in their breadth and depth. This, together with the complexity of MSD approaches, demands more than one approach to M&E (including VfM assessments). Figure E.6 highlights some of these. There will also be trade-offs between the different demands. While M&E is often portrayed to aid accountability and learning, there are tensions and contradictions in practice with demands for accountability coming at the expense of learning (Reinersten et al, 2017).

Figure E.6: M&E approaches for MSD programs



<sup>85</sup> For financial and human resources.



Performance monitoring remains a central purpose of most development program M&E systems. Performance monitoring systems are designed to collect data for indicators measuring intended results achieved by change pathways, delivered according to program strategies. This approach is compatible with interventions operating in both simple and complicated contexts where certainty and agreement on what needs to be done to achieve change are relatively high. Monitoring headline indicators is normally linked to performance monitoring, even though MSD programs operate in contexts with higher levels of uncertainty.

MSD programs that are adaptive require more than performance monitoring. M&E to support adaptive management is strongly embedded in the DCED Standard as it promotes<sup>86</sup> the adaptation of results chains and indicators, creating regular feedback loops of information for program managers. M&E requirements for adaptive and complex programs need to put learning at the centre to be flexible and responsive to changing contexts and needs (Pasanen, 2017; Simister, 2018). Pasanen (Pasanen, 2017) provides five reality checks for adaptive management monitoring and evaluation. These are:

1. bigger M&E budgets
2. adaptive M&E budgets
3. more people involved in monitoring and analysis
4. ensure that (managers and technical) staff have the right competencies; skill set of M&E staff, in particular encouraging more strategic thinking.
5. select evaluation approaches carefully.

For programs seeking to change market systems through an adaptive approach, the needs are greater. Given the centrality of systemic change to the MSD approach, further work on the assessment of market systems change is warranted. DFAT's program portfolio provides an excellent opportunity to explore and test cost-effective approaches relevant to different contexts (e.g. shallow markets), building on data and information already collected by programs. Sharing experiences across programs is expected to provide valuable learning opportunities for programs, DFAT and the broader MSD community.

Policy engagement, business environment reform, business model innovation and direct delivery will likely require different approaches to M&E for specific interventions. There is a broad array of experience, approaches and methods outside of market systems development to draw on.

To test the MSD model would require additional approaches that include tracking programs over time. Based on the programs covered by this review, the evidence of the effectiveness of the MSD principles is anecdotal and has not been the result of systematic analysis. This could be another area for DFAT to explore. For the future, specific evaluation questions might include:

- to what extent and how, and in what circumstances, did adaptive management contribute to better/more outcomes? and
- to what extent did analysis of market systems contribute to the understanding of constraints and identification of opportunities for the program to intervene?

Lastly, better communicating progress, results and lessons from MSD programs for a variety of audiences is important. There is also a need for analysis to be communicated in ways relevant for different audiences. This does not negate the need for technical analysis and reports but there is

---

<sup>86</sup> This is done by including 'control points' relating to program staff using information regularly to guide their decisions about interventions.

also a need for summary findings, lessons and conclusions to be communicated in laypersons language.

The complexity of MSD programs means that their M&E is not straightforward. Some programs may question the value of doing more, but it might also mean doing things differently, less of some things and more of others. Further consideration of what is feasible and what is of most value to DFAT and programs is important. The cost of *not* improving areas like assessing systemic change also needs to be factored into discussions.

## Addendum 1: Examples of different indicator definitions

Table E. 5: Different definitions used to count the value of private sector investment leveraged

<p><b>DFAT's Target Definitions:</b> 'means expenditure (financial or in-kind) by a commercial enterprise – including sole traders, small and medium enterprises, large locally-owned firms and multinational corporations or representative organisations – that contributes to a development outcome'.</p> <p><b>DFAT Guidelines state:</b> 'This indicator should capture all expenditure by the private sector that has been directly influenced by DFAT ODA funding. The indicator is purposefully broad and program areas should attempt to quantify the value of all private sector contributions to development outcomes. This includes the financial value of in-kind contributions of staff and facilities to development projects as well as the value of private investments that have been incentivised or assisted through technical advisory or project preparation facilities'.</p>
<p><b>PRISMA:</b> Included: New personnel for the partnership; Existing personnel that have been fully reallocated to the partnership; Assets purchased or rented for the partnership; Consultants for partnership activities; Operation costs towards the partnership (all costs related to the new personnel, travel, etc.); Raw materials (including samples provided to farmers) for partnership activities; Loans from the partner to farmers or ISPs for partnership activities; Direct activity costs;</p> <p>Excluded: Management fees; Staff and management commitment from existing personnel (except if the staff or manager has been fully reallocated to the partnership)</p>
<p><b>MDF II:</b> Measures the amount of money the partner invests in the development and implementation of the innovations or regulatory reform. The investment can be made directly in partnership activities or in further improvements to products or services resulting from a partnership. The investment can be made directly by partners or additional investment leveraged by partners from private funding sources. The figure reported therefore captures the amount of expenditure made by the partners within and outside of the investment commitment made as per the Partnership Agreement with MDF.</p>
<p><b>AMENCA 3:</b> 'Financial' contributions involving the transfer of funds, either directly from one commercial entity to another partner who is providing services or produce, or to purchase services, agricultural supplies or equipment from a third party. In-kind' investments are usually where a partner agrees to provide inputs, supplies or the services of its staff free of charge. This might be the provision of free seeds or fertiliser, agricultural equipment or commercial extension services. From the farmer's perspective, this includes the provision of land, for instance for demonstration sites. It is important to estimate the cost to that partner of making this provision at the current commercial price paid within the sector for the goods and services in question, as the provider is not only forgoing the cost of these items but also its regular profit. In the case of in-kind contributions of staff or facilities, the calculations of funds leveraged should take account of the cost to the business of providing these services and facilities.</p>

## Addendum 2: MSD program's approaches to assessing VfM

This addendum, firstly, provides additional information relating to how DFAT's MSD programs are assessing VfM; and secondly provides a limited number of examples of how other organisations or programs are assessing VfM with links to additional resources.

### VfM assessment approaches

#### *Managing for value for money by putting the right processes in place*

DFAT's VfM principles relate to managing for value for money - by putting the right processes in place. Only PHAMA Plus has articulated data collection, analysis and reporting processes related to these principles into their monitoring and evaluation plan<sup>87</sup>. PHAMA Plus has not yet implemented these processes as the inception period only concluded in March 2019.

While not referencing DFAT's VfM principles, other programs are likely to provide a partial, but perhaps not systematic assessment, of some aspects of managing for VfM. For example:

- LIFT's Annual Report (2018) provides some relevant information about how it manages for value for money. LIFT is a large multi-donor funded initiative that has common processes and procedures that fund members agree to, to mitigate overburdening the program with multiple reporting requirements. DFAT's contribution to the fund is small when compared to the total pool of funds.
- Strongim Bisnis highlights specific policies and processes that contribute to the 4Es and cost-effectiveness.
- TOMAK has also included at least two questions in its monitoring, evaluation, reporting and learning framework that relate to managing for value for money.
- AIP-Rural's portfolio review processes include consideration of the possibility of success and scale in decisions to drop, improve or change.

The lack of a common framework for assessing 'managing for value for money' covering DFAT's VfM principles<sup>88</sup> means most MSD programs do not systematically collect, analyse and report information related to the principles. Where programs do (or intend to do – PHAMA Plus), this is based on their own interpretation of the information DFAT is looking for.

#### *Comparing value for money to drive improvement and inform decision-making*

Comparing value for money can take two perspectives: within program and across programs. Within-program analysis could assist program managers, whereas across-program analysis could help DFAT with its programming decisions.

### **Within programs**

Most programs' approaches to value for money seem implicitly tied to VfM to drive improvement and inform decision making, for example by comparing the cost per beneficiary of delivering a service or achieving an outcome.

Ex-ante assessments of likely intervention costs and benefits guide decision making related to resources and effort. Program examples<sup>89</sup> include:

---

<sup>87</sup> PHAMA Plus MRM Manual (2019)

<sup>88</sup> The PPA's focus on the economy element of value for money and lack of articulation about what behaviours illustrate the principles in action mean it is not adequate as an assessment tool.

<sup>89</sup> These examples are drawn from procedure guidelines. Detailed examples of these assessments in practice have not been reviewed.

- PHAMA Plus' portfolio review processes includes an assessment of the degree to which the program is meeting the VfM criteria as outlined in its monitoring and evaluation plan.
- Strongim Bisnis' notes underlying funding guiding principles include value for money and value for money is noted as an investment criteria, although it is not explicit about how the ex-ante assessment is made.

Most MSD programs are required to report private sector investment leveraged for DFAT and most also record the program's investment costs, reporting this as a sign of value for money (efficiency). Some programs, such as PRISMA II, aggregate data and present the relationship as a ratio. Program and private sector investments are usually recorded at an intervention level (e.g. GREAT, MDF, AIP-Rural and PRISMA), but not necessarily reported or analysed. Program reports examined in this review do not include comparisons and analysis of ex-ante and ex-post results or ratios. PRISMA I analysed trends in changes at the aggregate level but does not provide insights into reasons for changes or variations within the portfolio.<sup>90</sup> Notably, some programs reported they do not use this information beyond reporting to DFAT.

The advantage of these basic cost-benefit indicators, e.g. intervention cost to net increased income, is that data is already being collected and recorded. Some indicators are relatively straight-forward and easy to understand. Assuming a consistent application of the definition of costs and benefits is used across interventions, programs could compare across interventions. LIFT, for example, compares the cost per beneficiary between agriculture production and microfinance projects. However, to interpret the numbers further information on factors affecting and impacting costs and income would be necessary. At times, it may be easier to sensibly make comparisons of the VfM of interventions within a common geographic area, sector or by type of intervention. For instance, comparing intervention costs and benefits of MDF interventions in Pakistan and Papua New Guinea may say little since there are different contextual cost drivers. Similarly, costs associated with more remote locations (e.g. East Java versus Papua) may make it is hard to directly compare ratios of costs and benefits. MDF also measures the net additional value of market transactions as an indicator of value for money.<sup>91</sup> However, without reference to the significance of the additional value to specific markets or sectors, and number of businesses operating, the number alone is not meaningful. Therefore, as standalone VfM indicators they have limited utility.

The general focus on a limited number of quantitative indicators fails to appreciate benefits that are not easily quantified or monetised. In situations where there are a variety of intervention-types e.g. business environment or advocacy-related activities, the program's full value or the value of specific types of interventions may be under appreciated. This may also inadvertently incentivise programs to prioritise interventions where there is a more immediately direct relationship to beneficiary incomes.

Even within programs, approaches to assessing value for money may differ over time. For example, CAVAC I and II MTRs and evaluations used slightly different approaches, limiting their comparability over time.

### **Across programs**

Benchmarking value for money across programs is broadly seen as challenging, and even problematic given different program structures (such as MDF's multi-country structure),

---

<sup>90</sup> PRISMA Activity Completion Report, 2018

<sup>91</sup> PRISMA have a similar indicator but do not consider it a value for money indicator.

implementing contexts and associated cost drivers (e.g. costs of labour, access to implementation sites).<sup>92</sup>

While some programs report against similar indicators e.g. intervention costs to beneficiary income, the definitions can differ across programs and therefore a reliable cross-program comparison would require a thorough understanding of what each program is counting. Some programs also measure increased income at beneficiary level while others measure it at the household level. An example of this type of challenge is provided below (Box E.9) for the indicator private sector investment leveraged. DFAT's guidelines provide some flexibility as to what programs count, taking their context into account. This provides some challenges for reliable and meaningful comparisons.

*Box E.9: Private sector investment leveraged – What is counted?*

#### **Private sector investment leveraged – what is counted?**

- Amongst other things PRISMA counts loans from businesses to farmers or implementing service providers for partnership activities.<sup>93</sup>
- MDF captures investments made directly by partners or additional investments leveraged by partners from private funding sources.
- AMENCA 3 includes opportunity costs associated with the provision of land for demonstration sites as farmers' in-kind investments.

DFAT usually conducts some cross-program comparisons, most obviously through the Annual Program Performance Reports that are based on AQC scores. Cross-program comparisons of criteria such as relevance, efficiency, effectiveness are made within a country annually.

#### *Demonstrating value for money through evaluation.*

Few programs are undertaking thorough value for money analysis. CAVAC II and LIFT provide the best examples with specific projects or parts of their program conducting more detailed analysis.

#### *Other organisations' approaches to assessing VfM*

The BEAM Exchange notes that VfM systems are often among the least understood or developed components of the monitoring frameworks for market systems programmes.<sup>94</sup> The challenges experienced by DFAT's MSD programs are not new, or experienced only by MSD programs.<sup>95</sup> Across development programs, the VFM concept has often been applied 'vaguely, so meaning different things to different people, or narrowly, defined as relating simply to cost'<sup>96</sup>. The current state of VfM assessments are influenced by:

- programs that are under pressure to demonstrate VfM before the 'value' of the program or intervention is realisable;
- a tendency to focus on what is measurable (or comparable through benchmarks) rather than what is important.

---

<sup>92</sup> Conway, K and Kessler, A (2019)

<sup>93</sup> It is not clear from documents reviewed how PRISMA accounts for the nature of loans in that they are repayable capital and the same \$1000, for example, may be used multiple times. PRISMA also reports total co-investments from private and public sector as well as from farmer households.

<sup>94</sup> BEAM Exchange (2019)

<sup>95</sup> ICAI (2018)

<sup>96</sup> Barr and Christie (2015)

- disconnects between different sets of people, from those in finance, procurement and administration that have a more granular approach to money, to technical advisers, team leaders and specialist consultants who concentrate on results and value.<sup>97</sup>

These factors may lead to an over-emphasis on unit costs and cost savings; prioritisation of cost reduction over maximising impact; a focus on maximising beneficiary numbers rather than reaching the poorest; a distorted picture and judgement. A common assumption that seems to underpin VfM assessments is that over the life of the program, VfM will improve. Given that individual indicators are often measured, this translates to an expectation of a downward trend for measures of economy, such as fee rates and overheads, while an upward trend on ratios such as benefit to cost.

Outlined below are four brief examples from other organisations and MSD programs and their approach to assessing value for money. It is not possible with the time limits of this review to provide a substantive review of VfM assessments on MSD programs. Each example includes lessons drawn from undertaking the assessments, highlighting that some of the challenges noted above have not yet been overcome and that greater consideration of the purpose of VfM assessments, and how the information will be used by donors and programs is warranted

### Example 1: MADE Ghana

The Department for International Development's (DFID) MADE program in Ghana (Box F.5) completed studies in 2016 and 2018 to benchmark key indicators with other MSD programs. The purpose for the study or how the information will be used by the program or DFID are unknown. Box F.5 provides an overview of the key questions behind the 2018 study and the types of comparisons made. The study highlighted a range of issues that can affect the usefulness of the findings.

*Box F.5: Approach to benchmarking MADE Ghana's value for money*

In 2018 MADE Ghana, a UK DFID funded program, compared selected costs and benefits with seven M4P programs that were considered most relevant (e.g. similar budget, start date).<sup>98</sup> Key questions were:

- How well are inputs converted into outputs?
- How well are the outputs from an intervention<sup>99</sup> achieving the desired outcome on poverty reduction?
- How fairly distributed are program benefits?<sup>100</sup>

Based on information available, comparisons were made across: 1) expenditure to outreach (including female outreach); 2) expenditure to market/actor change; 3) rate of expenditure to budget as a proxy for the strength of financial management; and 4) delivery against timeframe, implementers performance, asset management and monitoring and results measurement, as proxies for the strength of management.

Due to information gaps, comparisons could not be made across: proportion of management costs to overall program spend; proportion of management overhead to total spend; management costs to intervention costs; management overhead to outcome (beneficiary using/accessing inputs and

<sup>97</sup> Barr and Christie (2015)

<sup>98</sup> Nathan Associates (2018). MADE 2018 Benchmarking Study (Prepared by Holly Lard Krueger)

<sup>99</sup> Interventions refer to programs

<sup>100</sup> These reflect the 4E framework used by DFID – economy, efficiency, effectiveness and equity.

services); total private investment to program co-investment; and total additional income to female beneficiary experiencing improved income.

The study observed that factors to consider when benchmarking VfM include:

- Interpreting VfM analysis results requires careful consideration of context to provide relevant findings and insights.
- Determining what is 'good' value for money is not straightforward. A balance between economy, efficiency, effectiveness and equity is required as VfM concerns both being economical while maximising actual outcomes.
- Benchmarking exercises can 'miss the forest for the trees', they provide an assessment of relative performance not absolute performance. Ex-ante assessments remain important. VfM analysis and benchmarking can be important for learning and should consider a diverse range of data.
- Too much focus on VfM may disincentivise adaptive management, innovation and understanding or pursuit of larger development impact.

#### Example 2: Making Vegetable Markets Work

The LIFT-funded Making Vegetable Markets Work (MVMW) tracked a range of quantitative indicators to assess VfM against a target of five to one (5:1) return on investment. This means that for every LIFT dollar spent, the MVMW project interventions would generate five dollars in e.g. additional net attributable income change by farmers from program activities, increase in partner business revenues, farm-level investment, cost-share or match and crowding-in by private sector partners. The implementing organisation proposed the ROI effectively captures value for money, and therefore impact, because it is consistent with how our private sector partners and farmers will measure the impact of their own investments in improving their own businesses, and by extension, the market system. The ToR for the external end of project evaluation linked VfM to efficiency and included the following questions:

- To what extent has the MVMW project delivered value for money against its results framework, where material/tangible benefits are measurable?
  - To what extent have value for money considerations been taken into account in the selection and design of projects?
  - What are the costs of the interventions compared with the monetary benefits from the interventions?
  - To what extent has value for money been used to inform management decisions?

The evaluation examined the value for money for individual interventions in the project, including business advisory, contract farming, capacity building. They found the MVMW's ROI approach was useful for some interventions, but not others e.g. those related to capacity building of research organisations, where outcomes were more difficult to measure (and quantify).

#### Example 3: Financial Sector Deepening Program, Kenya

The Financial Sector Deepening Program in Kenya (Arora, 2012) aimed to demonstrate that VfM analysis could be practically applied to a large, complex and multi-dimensional donor project designed to enhance financial sector access in Kenya and to reduce the rate of financial exclusion. The program applies a MSD approach. It focused on measuring outcomes that could be quantified even though it also identified a range of benefits that could not be quantified. Quantifiable benefits included: Increased number and penetration of bank accounts; money transfers lower costs and greater volumes; increased availability of bank and non-bank financial outlets. Examples of non-quantified benefits that were identified but not included in the analysis were: reduced costs of international transfers; lower costs in Government cash transfer programmes; benefits from new



regulations; and benefits from improved management systems in community-based financial institutions.

Perhaps anticipating some criticism about the methodology that only focused on the more easily quantifiable benefits, the authors (2012: 6) noted:

*These non-quantifiable elements are arguably of central importance to the overall impacts on financial markets, wealth creation and thus poverty reduction that DFID and its donor partners sought to achieve when they launched the FSD initiative in 2005. So, how can these elements be omitted from a quantification of the VFM of the project overall? The answer in this case, and we suggest in many other similar donor projects, is that the absence of any credible means to put quantifiable values on these elements effectively forces their omission from any fully quantified VFM analysis. However, for many strong donor projects, including FSD, it may still be possible to demonstrate the VFM in the FSD programme as a whole by simply concentrating on a few (in this case four) quantifiable elements as identified above. In marginal cases, the non-quantifiable elements can be brought back into the analysis, while explicitly recognising the limitations of any quantification that is then attempted in those elements. This may be a messy approach compared to one that purports to link up all the elements of a project in one neat package of numbers. But it is, we suggest, a superior and more honest approach than one that regards all elements as being equally amenable to quantification.*

#### Example 4: Samarth-Nepal Market Development Program<sup>101</sup>

The final evaluation of the Samarth-Nepal Market Development Program assessed the program's VfM framework, which was based on a typical set of quantitative VfM indicators for DFID-funded projects (e.g. average fee rates, overheads as a percentage of total spend, cost of outreach, proportion of women reached, and beneficiary impact to cost ratio). Benefits that were not monetised, and included in the assessment, included government and local staff capacity building on private sector development, and knowledge management. The review examined VfM by sector across the program and some VfM indicators (outreach and cost of outreach) were benchmarked with MSD programs in other countries. There was very limited analysis as to what the variations meant and their significance. However, the review noted some recommendations that may also be useful for other MSD programs considering undertaking VfM assessments. These were:

- Programs should document lessons including failures, as well as their costs.
- Programs should undertake internal comparisons of sectors/ portfolios in order to identify winners from a VfM standpoint.
- A program VfM framework should be designed from the outset, and ideally be aligned to the logframes.
- Programs should incorporate more substantial aspects of economy, relating to key cost drivers such as share of fees in overall spend.
- MSD programs should use a more consistent approach and calculation method regarding calculation of leveraging of private investment.
- Future programs should consider using non-monetised measures of effectiveness, as well as the monetised metrics, as not everything about a program's value can be captured in periodic metrics.

---

<sup>101</sup> E-pact (2019)

## Alternative approaches to assessing VfM

In recognition that there is room for improvement in value for money assessments, some organisations have developed new approaches or tools to assist programs. Two are outlined below:

Barr and Christie (2015) propose that a **VfM indicator matrix** that distinguishes between types of indicators types of measurement to strengthen the formation of VfM judgements.

Indicator types:

- Monetary indicators report the monetary value of a point on a program's results chain (for example, an output or an outcome) in relation to the associated cost.
- Quantitative indicators report how much (in numbers) a program has achieved in relation to the associated cost.
- Qualitative indicators report the kind of change a program has achieved (in descriptive terms – for example, an improvement in the quality of a process or product) in relation to the associated cost.

Measurement types:

- Benchmarked measurement compares program achievements with similar achievements outside the program (within country or outside country). They are thus external, relative indicators, and can provide strong evidence of best value or best cost or both.
- Comparative measurement shows progress over time (for example, years) or space (for example, districts), demonstrating cumulative effect or showing comparative improvement between 'cases'. They are internal, relative indicators.
- Stand-alone measurement shows what has been achieved within a reporting period. These are 'one-off' realisations of value, and not likely to be repeated. They can be compared against the planned target for that period, in which case the value in VfM terms depends on the credibility of the original plan as both realistic and stretching. They may be important as denoting a results step-change.

This approach has been applied in a number different types of programs including economic development, governance and health. Program characteristics have led to adaptations to the approach.

OPM and King 's (2018)VfM framework takes a more evaluative approach to VfM assessment. It uses criteria (dimensions of VfM) and standards (levels of performance) to provide a transparent basis for making sound judgements about performance and VfM. It aims to use quantitative and qualitative information to provide a more nuanced indication of VfM than quantitative indicators can provide. The approach also aims to accommodate economic evaluation (where feasible and appropriate) without limiting the analysis to economic methods and metrics.

The approach has been used on MUVA women's economic empowerment project in Mozambique

## Annex F: References and documents reviewed

Arora, S., Roe, A and Stone, R., 2012, Assessing Value for Money: The case of donor support to FSD Kenya. Oxford Policy Management, Oxford.

AusAID, 2012, From Seed to Scale-Up: Lessons learned from Australia's rural development assistance, <https://dfat.gov.au/aid/how-we-measure-performance/ode/Documents/from-seed-to-scale-up.pdf>

Australian Government, 2017, 2017 Foreign Policy White Paper, <https://www.fpwhitepaper.gov.au/foreign-policy-white-paper>

Barr, J. and Christie, A., 2015, Improving the Practice of Value for Money Assessment. Centre for Development Impact Practice Paper, No. 12. Institute of Development Studies, University of Sussex.

BEAM Exchange, 2019, Module 7: Assessing value for money. Downloaded from: <https://beamexchange.org/guidance/monitoring-overview/assessing-value-money/overview/>

Bear M, and Bekkers H, 2018, In Search of the Sweet Spot in Implementing MSD Programmes, MDF Case Study No 5, [https://beamexchange.org/uploads/filer\\_public/dd/82/dd826973-b8c8-45e3-b1c2-68733cc43a82/in\\_search\\_of\\_the\\_sweet\\_spot\\_in\\_implementing\\_msd\\_programmes\\_-\\_part\\_3\\_compressed.pdf](https://beamexchange.org/uploads/filer_public/dd/82/dd826973-b8c8-45e3-b1c2-68733cc43a82/in_search_of_the_sweet_spot_in_implementing_msd_programmes_-_part_3_compressed.pdf)

Bond, 2012, Value for Money – what it means for NGOs. Bond, London. [https://www.bond.org.uk/sites/default/files/value\\_for\\_money\\_-\\_what\\_it\\_means\\_for\\_ngos\\_jan\\_2012.pdf](https://www.bond.org.uk/sites/default/files/value_for_money_-_what_it_means_for_ngos_jan_2012.pdf)

Conroy K. and Kessler A., 2019, The results achieved by programmes that use the market systems development approach, BEAM Evidence Review, [https://beamexchange.org/uploads/filer\\_public/a9/b5/a9b5c464-b21a-4537-85b2-aba11ee381a3/beam\\_evidence\\_review\\_2019\\_final\\_compressed.pdf](https://beamexchange.org/uploads/filer_public/a9/b5/a9b5c464-b21a-4537-85b2-aba11ee381a3/beam_evidence_review_2019_final_compressed.pdf)

Cunningham S. and Jenal M., 2016, Rethinking systemic change: economic evolution and institutions, BEAM exchange technical paper, [https://beamexchange.org/uploads/filer\\_public/17/14/17149bf0-9944-4699-8c2e-1057097d914e/systemic\\_change\\_technical02\\_compressed.pdf](https://beamexchange.org/uploads/filer_public/17/14/17149bf0-9944-4699-8c2e-1057097d914e/systemic_change_technical02_compressed.pdf)

Dart J., 2016, View of the DCED Standard prepared for DFAT Indonesia Post, unpublished Donor Committee for Enterprise Development (DCED), 2017, Standard for Measuring Results in Private Sector Development, Version VIII, [https://cdn.enterprise-development.org/wp-content/uploads/DCED\\_Standard\\_VersionVIII\\_Apr17.pdf](https://cdn.enterprise-development.org/wp-content/uploads/DCED_Standard_VersionVIII_Apr17.pdf)

Davis, G., 2016, Getting to scale: Lessons in reaching scale in private sector development programmes, [https://beamexchange.org/uploads/filer\\_public/ab/fc/abfcb2d3-7847-448e-9ed6-c99fd99f3b94/getting\\_to\\_scale.pdf](https://beamexchange.org/uploads/filer_public/ab/fc/abfcb2d3-7847-448e-9ed6-c99fd99f3b94/getting_to_scale.pdf)

DCED, 2018, The 2018 Reader on Results Measurement: An introduction to the DCED Standard, [https://cdn.enterprise-development.org/wp-content/uploads/DCED\\_Reader\\_RM.pdf](https://cdn.enterprise-development.org/wp-content/uploads/DCED_Reader_RM.pdf)

DFAT 2010, Market Development Facility, Program Design Document, <https://dfat.gov.au/about-us/publications/Documents/market-design-document.pdf>

DFAT, 2014a, Australian aid: promoting prosperity, reducing poverty, enhancing stability, <https://dfat.gov.au/about-us/publications/Documents/australian-aid-development-policy.pdf>

DFAT 2014b, Making Performance Count: enhancing the accountability and effectiveness of Australian aid, <https://dfat.gov.au/about-us/publications/Documents/framework-making-performance-count.pdf>

DFAT 2015a, Strategy for Australia's aid investments in private sector development, <https://dfat.gov.au/about-us/publications/Documents/strategy-for-australias-investments-in-private-sector-development.pdf>

DFAT 2015b, Creating shared value through partnership: Ministerial statement on engaging the private sector in aid and development, <http://dfat.gov.au/about-us/publications/aid/Pages/creating-shared-value-throughpartnership.aspx>

DFAT 2015c, Strategy for Australia's aid investments in agriculture, fisheries and water, <http://dfat.gov.au/aboutus//publications/Pages/strategy-for-australias-aid-investment-in-agriculture-fisheries-and-water.aspx>

DFAT, 2015d, Development for All, Strategy for strengthening disability-inclusive development in Australia's aid program, <https://dfat.gov.au/about-us/publications/Documents/development-for-all-2015-2020.pdf>

DFAT, 2015e, Value for Money principles, <https://dfat.gov.au/aid/who-we-work-with/value-for-money-principles/Pages/value-for-money-principles.aspx>

DFAT, 2016, Gender equality and women's empowerment strategy, <https://dfat.gov.au/about-us/publications/Documents/gender-equality-and-womens-empowerment-strategy.pdf>

DFAT, 2017a, Market Systems Development: Operational Guidance Note, <https://dfat.gov.au/about-us/publications/Documents/operational-guidance-note-market-systems-development.pdf>

DFAT, 2017b, Cambodia Agricultural Value Chain (CAVAC) Phase One: Completion Evaluation, <https://dfat.gov.au/aid/how-we-measure-performance/ode/strategic-evaluations/Documents/cambodia-agricultural-value-chain-cavac-phase-one-evaluation.pdf>

DFAT 2017c, DFAT Monitoring and Evaluation Standards, April 2017, <https://dfat.gov.au/about-us/publications/Documents/monitoring-evaluation-standards.pdf>

DFAT, 2018b, Independent Facilities Review, <https://dfat.gov.au/about-us/publications/Documents/independent-facilities-review.pdf>

DFAT, 2018a, MDF and Systemic Change – A Summary for Program Managers, <https://dfat.gov.au/aid/topics/investment-priorities/agriculture-fisheries-water/agriculture-food-security/Documents/mdf-and-systemic-change-summary-for-program-managers.docx>

DFAT, 2018b, Review of DFAT Investment Level Monitoring Systems. Office of Development Effectiveness, Canberra. <https://dfat.gov.au/aid/how-we-measure-performance/ode/strategic-evaluations/Documents/evaluation-of-investment-level-monitoring-systems.pdf>

DFAT, 2018c, Aid Programming Guide, <https://dfat.gov.au/about-us/publications/Pages/aid-programming-guide.aspx>

DFAT, 2018d, Monitoring and Evaluation, Practitioner Level, Education and Development Module, <https://dfat.gov.au/aid/topics/investment-priorities/education-health/education/Documents/practitioner-monitoring-and-evaluation.pdf>

DFAT, 2019, Operational Framework for Private Sector Engagement in Australia's Aid Program, <https://dfat.gov.au/about-us/publications/Documents/private-sector-engagement-in-australias-aid-program-operational-framework.pdf>

e-Pact, 2019, Samarth-NMDP Programme, Final Evaluation Report, 21 July 2019, unpublished.

DFID, 2015, DFID policy framework for the provision of grants or concessional finance to for-profit firms. UK Department for International Development, London.

DFID, 2016, Technical Competency Framework Private Sector Development, 2016, [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/553188/Private-sector-dept-Sept16\\_.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/553188/Private-sector-dept-Sept16_.pdf)

e-Pact, 2019, Samarth-NMDP Programme, Final Evaluation Report, July 2019.

ESID, 2015, Making political analysis useful: Adjusting and scaling (ESID Briefing Paper No. 12). Effective States and Inclusive Development Research Centre (ESID) School of Environment and Development, The University of Manchester, Manchester.

Geels F.W., 2011, The multi-level perspective on sustainability transitions; responses to seven criticisms. Environmental Innovation and Societal Transitions, 1; 24-40 cited in Getting a nudge: moving from niche innovations to systems change, <http://www.dovetailnz.com/news/16-September-2019>).

Heinrich M., 2014, Demonstrating Additionality in Private Sector Development Initiatives, DCED <https://www.enterprise-development.org/wp-content/uploads/DCED-Demonstrating-Additionality-final.pdf>

HEKS/EPER, 2015, Market Systems Development: Guideline to plan and facilitate system changes, <https://en.heks.ch/sites/default/files/documents/2017-03/HEKS-MarketSystemsDevelopmentGuideline-2016-E.pdf>

ICAI, 2018, DFID's approach to value for money in programme and portfolio management - A performance review. ICAI, London.

Itad, 2018, Evaluation of the market systems development approach: Lessons for expanded use and adaptive management at Sida, <https://www.sida.se/contentassets/bfe15e8902fa4dbb864bd478c2f14df1/2018-2a-evaluation-market-systems-dev-approach-vol-1.pdf>

Jenal, M. 2019, Market systems resilience – a concept to reframe systemic change and sustainability? <https://beamexchange.org/to-pdf/?url=/community/blogs/2019/7/11/market-systems-resilience-systemic-change-sustainability/>

Jordan, S., Davies, G., Bell, E (2017). Synthesis Document: How to Integrate Gender and Women's Economic Empowerment into Private Sector Development Programmes. DCED, Cambridge. <https://www.enterprise-development.org/wp-content/uploads/DCED-WEEWG-How-to-integrate-gender-into-PSD-programmes.pdf>

Kershaw I., and Warner B., 2019, Advisory Group Management Review.

Koh, H et al, 2016, Shaping Inclusive Markets: How Funders and Intermediaries can Help Markets Move toward Greater Economic Inclusion. FSG, Boston. Downloaded from [www.fsg.org/publications/shaping-inclusive-markets](http://www.fsg.org/publications/shaping-inclusive-markets)

- Lomax, J., 2019, Why can't we measure systemic change'. BEAM Exchange blog.  
[https://beamexchange.org/to\\_pdf/?url=/community/blogs/2019/5/30/why-cant-we-measure-systemic-change/](https://beamexchange.org/to_pdf/?url=/community/blogs/2019/5/30/why-cant-we-measure-systemic-change/)
- McCulloch, N., Barnett, A., Duncan, A., Kingsmill, W., Kydd, J., Owen, O., and Williams, G., 2017, The Evolving Role of Political Economy Analysis in Development Practice. Policy and Practice Brief 11.  
[www.policyandpractice.co.uk](http://www.policyandpractice.co.uk)
- MDF, 2015, Achieving Changes in Markets: The MDF Framework for Defining and Populating Pathways for Systemic Change, Strategic Guidance Note .3, [https://www.enterprise-development.org/wp-content/uploads/Case\\_8SystemicChangeMDF.pdf](https://www.enterprise-development.org/wp-content/uploads/Case_8SystemicChangeMDF.pdf)
- MDF, 2017, Market Development Facility Phase I Activity Completion Report,  
<https://dfat.gov.au/about-us/publications/Documents/timor-leste-market-development-facility-activity-completion-report.pdf>
- Miehlabradt A., Swete Kelly D. and Warner B., 2018, Promoting Systemic Change in Shallow Markets” Lessons from Phase One of the Market Development Facility,  
[https://beamexchange.org/uploads/filer\\_public/b5/a4/b5a45d87-c0b4-4925-9e00-30c86c443b68/promoting-systemic-change-in-shallow-markets.pdf](https://beamexchange.org/uploads/filer_public/b5/a4/b5a45d87-c0b4-4925-9e00-30c86c443b68/promoting-systemic-change-in-shallow-markets.pdf)
- Nippard D., Hitchins R. and Elliott D., 2014, Adopt-Adapt-Expand-Respond” a framework for managing and measuring systemic change processes, The Springfield Centre,  
<http://www.springfieldcentre.com/wp-content/uploads/2014/06/2014-03-Adopt-Adapt-Expand-Respond-Briefing-Paper1.pdf>
- OECD, 2018, Making Blended Finance Work for the Sustainable Development Goals, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264288768-en>
- OPM and King, J., 2018, OPM’s approach to assessing Value for Money. Oxford Policy Management, Oxford. <https://www.opml.co.uk/publications/assessing-value-for-money>
- Pasanen, T., 2017, Monitoring and Evaluation; five reality checks for adaptive management,  
<https://www.odi.org/blogs/10588-monitoring-and-evaluation-five-reality-checks-adaptive-management>
- Reinersten, H., Bjorkdahl, K, and McNeill, D., 2017, Confronting the contradiction – an exploration into the dual purpose of accountability and learning in aid evaluation. Expert Group on Aid Studies, University of Oslo. Downloaded from: <https://eba.se/wp-content/uploads/2017/05/Webbversion-FINAL-blanksida.pdf>
- Ripley, M., 2019, Systemic Change: walking the talk? ILO the Lab – Market systems development for decent work.
- Ruedy L., 2018, Six Models for Understanding Impact, blog,  
<https://www.democracyfund.org/blog/entry/six-models-for-understanding-impact>
- Simister, N (2019). Praxis Series Paper No.6. Complex M&E Systems – Raising standards, lowering the bar. Intrac, Oxford, [Simister, N \(2019\). Praxis Series Paper No.6. Complex M&E Systems – Raising standards, lowering the bar. Intrac, Oxford.](https://www.intrac.org/publications/complex-m-e-systems-raising-standards-lowering-the-bar)

Springfield Centre, 2008, Perspectives on the Making Markets Work for the Poor M4P Approach, [https://beamexchange.org/uploads/filer\\_public/8d/45/8d458e51-cdba-4ba2-94db-be956f5ec2cb/perspectives2008\\_compressed.pdf](https://beamexchange.org/uploads/filer_public/8d/45/8d458e51-cdba-4ba2-94db-be956f5ec2cb/perspectives2008_compressed.pdf)

Tanburn J., Kershaw, I and Pyburn R., 2016, Market Development Facility (MDF), Mid-term review, unpublished.

Taylor, B., 2016, Systems and Systemic Change – Clarity in Concept. Springfield Centre for Business in Development.

Tewes-Gradi C., von Blomberg I., and Scholl J., 2018, Minimising the Risk of Negative Market Distortions in Private Sector Engagement: A practical framework, [https://www.enterprise-development.org/wp-content/uploads/DCED\\_Minimising-the-Risk-of-Market-Distortions-in-PSE\\_Practical-Framework.pdf](https://www.enterprise-development.org/wp-content/uploads/DCED_Minimising-the-Risk-of-Market-Distortions-in-PSE_Practical-Framework.pdf)

In addition to these referenced documents, the review examined the following material, wherever they were available:

- Various Aid Quality Checks between 2014 and 2019
- Program progress reports (quarterly, six monthly, annual)
- Mid term and program completion reviews
- Activity completion reports
- Inception reports
- Program strategies
- Monitoring and Evaluation Manuals
- Advisory group reports