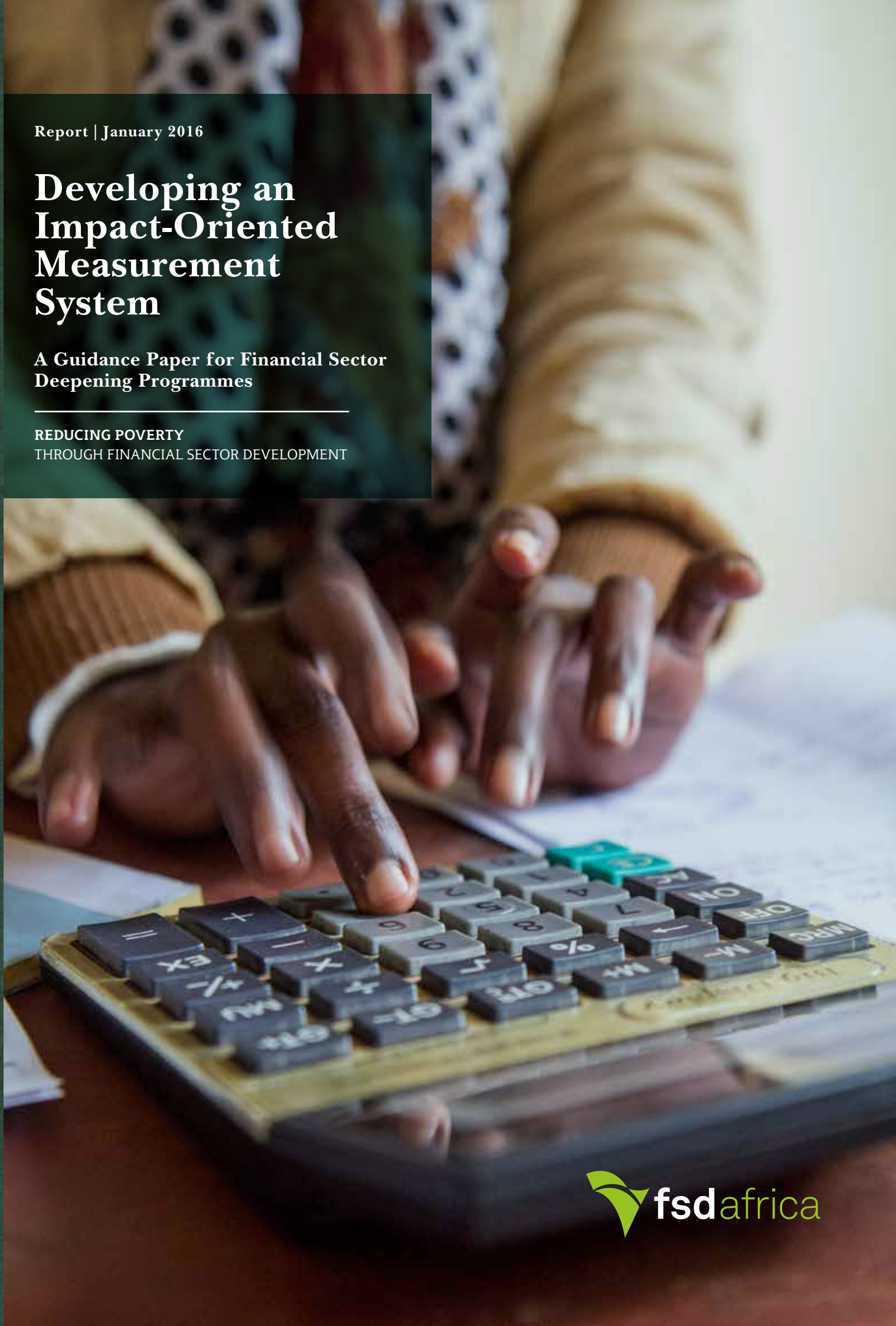


Report | January 2016

Developing an Impact-Oriented Measurement System

A Guidance Paper for Financial Sector Deepening Programmes

REDUCING POVERTY
THROUGH FINANCIAL SECTOR DEVELOPMENT

fsdafrica

Key points
Impact-Oriented Measurement

01

FSDs face increasing pressure to show results while implementing complex, multi-faceted, market-development programmes. Monitoring and evaluation practices have to keep pace with this changing context and respond robustly to more demanding expectations, particularly the need to measure medium-term market system outcomes and longer-term impacts on poverty reduction.

02

Impact-oriented measurement (IOM) aims to find the 'sweet spot' between monitoring and evaluation. This entails cycles of implementation and reflection to improve programme adaptation and build a robust evidence base to measure what is causing change in the financial sector.

03

IOM has two key objectives. First, to understand how FSD investments have contributed to observed changes in the financial sector, and how these changes have improved the livelihoods of the poor. Second, to track and improve the performance of FSD investments, by improving the evidence base on what works and what does not.

04

An FSD's Theory of Change (ToC), informed by regular market analysis that sets out the systemic constraints to financial sector development, is the bedrock of the measurement system. It guides the programmes and the funder to articulate both the expected market change process and which impacts and causal pathways to generate evidence on.

05

The term 'impact' is often used to refer to long-term changes at the end of a results chain, such as improved livelihoods or poverty reduction. We need to change the way we think about impact. A core focus of an individual FSD should be on how and why financial sector change occurs at every step of the ToC.

06

As market development programmes, FSDs' measurement lens needs to be recalibrated, moving beyond a focus on end numbers (e.g. level of financial access) to how the underlying structures, behaviours and incentives of the market have changed, or are changing, to support these outcomes.

07

Project-based measurement can become too focused on the specific interventions and fail to capture wider (and unexpected) changes in the financial sector. IOM augments this approach by offering a new framework that captures observed changes in the wider market system. We call these two approaches 'bottom-up' and 'topdown' measurement.

08

The change FSDs seek and the instruments they use to achieve such change have profound implications for results measurement. A range of approaches and methods are available for FSDs to set up an IOM, implement its core principles, and review its progress. These 'steps' are set out in this guidance.

Front cover image: Picture courtesy of FSD Africa

IOM Foreword

Since the creation of FinMark Trust in 2002, the FSD network has sought to be a leading force for poverty reduction through financial sector development in Africa.

Today, the FSD network comprises 9 like-minded organisations and over 120 in-country members of staff. It also spends around USD 55 million each year and represents a total invstment of USD 450 million by a range of international development donors, including the UK Government's Department for International Development (DFID) - its most significant supporter. Known as "FSDs", two are regional - FinMark Trust and FSD Africa, while seven are country-focused – Access to Finance Rwanda, Enhancing Financial Innovation & Access in Nigeria, FSD Kenya, FSD Moçambique, FSD Tanzania, FSD Uganda and FSD Zambia.

FSDs do not deliver financial services to the poor directly. Instead, they deploy financial resources, expertise and insights in collaboration with a range of public and private sector actors – from central banks and commercial banks to specialist training providers, telecommunication firms and microfinance networks - to create the market conditions that deliver financial inclusion, not only during the FSD intervention, but also beyond. This market facilitation (or M4P) approach works. Since 2002, support provided by the FSD network has helped extend the provision of financial services – from payment systems and loans to insurance and bank accounts – to millions of Africans and African businesses on a lasting basis.

Despite progress on the ground and a convincing weight of academic evidence, questions remain about the FSD approach. How can we be sure that it is the FSDs' inputs that cause changes in the financial market? And how do we know that these financial market changes lead to economic growth and poverty reduction? Are FSDs confident about which financial market changes are the most impactful? And are impact pathways similar between different African financial markets or does context take precedence?

As taxpayer-funded entities, it is incumbent on FSDs to monitor and measure results systematically. FSDs should celebrate their achievements, but also collect, analyse and communicate the evidence necessary to substantiate claims of success in a way that stands up to the scrutiny of a critical and informed audience. As a result, good monitoring and results measurement (MRM) helps to validate the FSD approach. It also enables a greater level of accountability.

For FSDs, MRM has another, equally important purpose. A clearer understanding what works, and what doesn't helps FSDs to become better market facilitators. First-rate MRM analyses emerging trends and provides actionable insights through regular feedback mechanisms that can guide ongoing programming choices. It can improve the way an FSD functions.

With these issues in mind, FSD Africa began an FSD network-wide consultative process in July 2014. Its objectives were two-fold: a) to strengthen MRM processes within individual FSDs at the project and programme level, and b) to develop a more consistent approach to MRM across the FSD network.

Through their participation and with the support of specialists from Oxford Policy Management (OPM) and the Consultative Group to Assist the Poor (CGAP), DFID and the Donor Committee for Enterprise Development (DCED), FSDs explored key MRM issues together. Topics included:

- Developing a common terminology for MRM work to avoid confusion within an FSD and with key partners (especially donors) and achieve consensus more quickly
- Consolidating a rich, sometime complex portfolio of FSD project work into a single MRM framework that is coherent and measurable
- Determining the core components (measurement tools, processes, indicators and management) of an MRM system to enable an FSD to quickly develop an approach that is well-understood, practical and which provides evidence in a timely, useful manner
- Better defining and measuring change in financial market systems (that is both expected and unexpected) to help prove and improve an FSD's market facilitation approach
- Determining an FSD network impact research agenda to create a better understanding of the causal relationships between certain kinds of financial sector interventions and the impact they are intended to generate

The result of this extensive consultation is the Impact Oriented Measurement framework, or IOM.

IOM is a comprehensive resource that helps FSDs, or FSD-like organisations, manage the challenge of measuring their contribution to changes in the market systems they seek to influence. IOM offers guidance, not a prescription. There is a high degree of consensus built-in to the model, which is informed by practical insights derived from FSD practice over a decade or

more, but also OPM, CGAP, DFID and DCED.

MRM will continue to evolve and adapt. However, it is hoped that FSDs, and similar organisations, will gravitate towards using IOM as their benchmark for MRM, even if full adoption of IOM will take time. Throughout, FSD Africa will look to support its implementation whether through the serialisation of its publication, MRM skills development, hosting MRM discussion groups and fora, and/or conducting FSD network-wide impact research.

We have little doubt that the complexity in financial market development will mean there will be a continuing need for hypothesis and research as we seek answers to tough questions concerning the links between financial sector development, economic growth and poverty reduction. We will undoubtedly get closer to some of these answers in the coming years. But with IOM, it is our ambition that at least the management of MRM will be a lot less mysterious for FSDs than it has been for many up until now. By sharing insights and practices, and by adopting a common measurement framework, we also aim for the FSD network as a whole, and not just its individual members, to claim credit for the collective impact it is achieving.

Mark Napier, Director, FSD Africa

Executive Summary

Financial sector deepening programmes (FSDs) promote the sustainable, pro-poor development of complex financial markets. To do this they work with market actors and policy-makers, and deploy multiple means of support, ranging from funding to research, while continually adapting to market changes. FSDs face increasing pressure to show results while implementing complex, multi-faceted market development programmes. Monitoring and evaluation practices have to keep pace with this changing context and respond robustly to more demanding expectations, particularly the need to measure medium-term market system outcomes and longer-term impacts on poverty reduction.

This impact-oriented measurement (IOM) guidance paper has two key objectives that are designed to assist FSDs in their measurement processes. First, to understand how FSD programme investments have contributed to observed changes in the financial sector, and how these changes have improved the livelihoods of the poor. Second, to track and improve the performance of FSD investments, by improving the evidence base regarding what works and what does not.

The change that FSDs seek – changes in the financial market system and the instruments they use to achieve such change – dynamic facilitation of a range of market actors – create profound challenges as regards measuring results. To address these measurement challenges, **the IOM is underpinned by five underlying principles:**

- Finding the 'sweet spot' between monitoring and (impact) evaluation. This entails cycles of implementation and reflection to improve programme adaptation and build a robust evidence base to measure what is causing change in the financial sector. This approach moves evaluation away from one-off actions.
- We need to change the way we think about impact. The term 'impact' is often used to refer to long-term changes at the end of a results chain, such as improved livelihoods or poverty reduction. A core focus of an impact evaluation for an individual FSD should be how and why financial sector change occurs at every stage of the theory of change (ToC).
- An FSD's ToC is the bedrock of the measurement system and is informed by regular market analysis that sets out the systemic constraints to financial sector development, and the framework for an FSD's multiple interventions.
- As market development programmes, the measurement lens for FSDs needs to be recalibrated, moving beyond a focus on end numbers (e.g. level of financial access) to how the underlying structures, behaviours and incentives of the market have changed, or are changing, to support these outcomes.
- Project/intervention-based measurement can become too focused on the specific interventions and fail to capture wider (and unexpected) changes in the financial sector. IOM augments this approach by offering a new framework that captures observed changes in the wider market system. We call these two approaches 'bottom-up' and 'top-down' measurement.

Different FSDs are at different stages of their strategy and funding cycles. This IOM guidance paper sets out the the three distinct stages involved in developing a measurement framework, tools and approaches for measurement for an FSD, and a way of bringing them together, from both an analytical and practical perspective. This includes the following steps:

Step 1: Ensuring that the FSD programme ToC is evaluable
Step 2: Developing impact measurement questions to guide the measurement focus over an FSD strategy period
Step 3: Developing impact-oriented indicators for both bottom-up and top-down measurement
Step 4: Data collection methods and sources
Step 5: Assessing causality and the contribution of FSD programme interventions
Step 6: The use of research to fill gaps in the testing of an FSD's ToC
Step 7: Articulating a credible narrative of the FSD's overall contribution to financial sector change, and poverty reduction

These steps are followed by guidance on embedding these measurement practices throughout FSD programme cycle.

Additional background papers and accompanying technical notes have also been prepared for the FSDs and for Financial Sector Deepening Africa (FSDA), which provide more detail on the conceptual underpinning of the guidance, and further technical detail on stand-alone topics. These papers can be found at the FSD Africa website – www.fsdafrica.org/knowledge-hub.

The IOM does not replace existing approaches, but it does provide the following benefits:

a. promotes greater coherence across FSDs as to how they approach impact evaluation;
b. provides guidance regarding areas in which FSDs have identified measurement challenges;
c. provides an in-depth and systematic approach to measuring outcomes and impact for FSDs;
d. encourages FSDs to focus not only on the measurement of outcomes (e.g. number of households with access to finance) but also on whether underlying structures, processes and incentives in the market for suppliers and users are affecting financial inclusion outcomes (i.e. are changes systemic?);
e. confirms that IOM need not necessarily involve complex/ expensive external studies; it is possible for FSDs to use their own staff and processes to collect important evidence of impact;
f. encourages FSDs to work with funders (and other stakeholders) to prioritise the learning questions and strengthen and speed up the feedback loops between programme design, implementation and review, and to facilitate learning to adjust investments and improve programme performance; and
g. helps FSDs to strengthen accountability to their funders, by generating a robust evidence base that can be used in FSDs' regular reporting and can be periodically tested by the independent evaluators.

Each member of the FSD network will determine the adoption and adaptation of the IOM guidance by considering its own context, funding and strategy timeline. FSDA is considering demand from the FSD network for support needed to implement the guidance and will accordingly finalise arrangements to provide support to those FSDs who take up implementation of IOM and require network support.

Contents

List of figures, boxes and tables

List of abbreviations

AFI	Alliance for Financial Inclusion
AFR	Access Finance Rwanda
Cenfri	Centre for Financial Regulation and Inclusion
CGAP	The Consultative Group to Assist the Poor
DCED	The Donor Committee for Enterprise Development
DFI	Development finance institution
DFID	Department for International Development (UK)
EFInA	Enhancing Financial Innovation & Access, Nigeria
EIU	Economist Intelligence Unit
FGD	Focus group discussion
FMO	Dutch development bank
FMT	FinMark Trust South Africa
FSAP	Financial Sector Assessment Program
FSDs	Financial sector deepening programmes
FSDA	Financial Sector Deepening Africa
FSDK	Financial Sector Deepening Kenya
FSDMoç	Financial Sector Deepening Mozambique
FSDT	Financial Sector Deepening Trust, Tanzania
FSDU	Financial Sector Deepening Uganda
FSDZ	Financial Sector Deepening Zambia
FSP	Financial services providers
GPFA	Global Partnership for Financial Inclusion
HR	Human resources
IFC	International Finance Corporation
IMF	International Monetary Fund
IOM	Impact-oriented measurement
KPOSB	Kenya Post Office Savings Bank
KWFT	Kenya Women's Finance Trust
KYC	Know your customer
Logframe	Logical framework
LSM	Living standards measure
M&E	Monitoring and evaluation
M4P	Making markets work for the poor
MFI	Microfinance institution
MFS	Mobile financial services
MIS	Management information system
MNO	Mobile network operator
MSME	Micro, small and medium-sized enterprises
NBFI	Non-bank financial institution
NGO	Non-governmental organisation
OECD	Organisation for Economic Co-operation and Development
OPM	Oxford Policy Management
OPR	Output to purpose review
PAR	Project appraisal review
PPI	Progress Out of Poverty Index
RCT	Randomised controlled trial
SACCO	Savings and credit co-operative
SADC	Southern African Development Community
SILC	Savings and internal lending communities
SME	Small and medium-sized enterprise
TA	Technical assistance
TBE	Theory-based impact evaluation
ToC	Theory of change
VfM	Value for money

1 Introduction

1.1 Background

Financial Sector Deepening Africa (FSDA) and financial sector deepening programmes (FSDs) in eight countries[1] have taken part in a one-year consultation process facilitated by Oxford Policy Management (OPM) to develop an approach to improving how they measure the impact of their work. The objectives of this new approach are two-fold:

i. to understand how FSD investments, in terms of time, money and influence, have contributed to observed changes in the financial sector – their 'impact'; and
ii. to track and improve the performance in respect of FSDs' contributions, through improving the FSDs' evidence base regarding what works and what does not.

These objectives are being addressed to varying degrees by FSDs' existing monitoring and evaluation (M&E) systems. **This guidance paper presents an overall approach and practical steps to augment these existing practices.** This guidance will equip readers with the tools and frameworks needed to implement an impact-oriented measurement (IOM) system.

The context for this work is the increasing emphasis on transparency, accountability and value for money (VfM) for donor funded programmes. Specifically, there is a desire to establish the extent to which the modality of an FSD works (given the increasing use of market development instruments by funders), and where possible to establish how it works (i.e. to identify and test the various pathways to change, to enable further improvements and course corrections as needed). Finally, by bringing the FSD network together, under the facilitation of FSDA, efforts have been made to adopt more robust and coherent approaches to M&E, and to facilitate peer learning regarding what does and does not work.

1.2 Methodology

Following initial research by the OPM team and consultations with the FSD network during July–September 2014, a workshop took place in Nairobi on 28–29 October 2014, at which preliminary ideas were discussed and inputs were sought from FSDs and other stakeholders. The workshop provided an opportunity to discuss: a) key challenges in assessing the impact for market development programmes; b) FSD experience so far; and c) various options for impact assessment of the market development programmes. During the first workshop, four key themes were identified for detailed work: measurement of the role of FSDs in promoting systemic change; financial sector market measurement; assessing the quality of access; and links between financial sector development and economic growth/development.

After the October workshop, volunteers from the FSD network participated in follow-up discussions and commented on draft papers. Since the first theme mentioned above – measuring systemic impact – is central to FSD operations, this topic was integrated into the core guidance while separate papers were prepared for the other three themes. The core IOM guidance, together with the thematic papers, were discussed at a follow-up workshop in Nairobi on 12–13 March 2015. For both the workshops, background material was circulated in advance to give FSD directors and staff attending the workshop an opportunity to review the material and consult other FSD colleagues ahead of the workshop. The March workshop also benefited from contributions by representatives of the UK Department for International Development (DFID), the World Bank, the MasterCard Foundation, the Bill & Melinda Gates Foundation and selected members of the governing bodies of FSDK and FSDT.

Members of the OPM team also visited each of the FSD country programmes to discuss current measurement practices and to identify key challenges and opportunities. This was supplemented with a review of other guidance available and consultation with experts at Consultative Group to Assist the Poor (CGAP), the Donor Committee for Enterprise Development (DCED), the BEAM Exchange, ITAD and the Springfield Centre.

CGAP, FSDA's knowledge partner, worked with OPM and FSDA throughout the assignment (including the two workshops). In addition, CGAP was able to test the emerging guidance with the CGAP-led Measuring

1. Enhancing Financial Innovation & Access, Nigeria (EFInA); FSD Kenya (FSDK); FSD Tanzania (FSDT); Access to Finance Rwanda (AFR); FinMark Trust South Africa (FMT); FSD Zambia (FSDZ); FSD Uganda (FSDU) and FSD Mozambique (FSDMoç). These are referred to as FSDs or FSD programmes throughout this guidance paper.

Market Development Group during calls in November 2014 and March 2015, and meetings of group of experts in March 2015 and June 2015.

Robert Stone, Sukhwinder Arora, Richard Williams, Ian Robinson and Sarah Keen were the core members of the OPM team. The core team was guided by a panel of experts, including Thorsten Beck, Susan Johnson, Celina Lee and Alan Roe. The OPM team also greatly benefited from frequent consultations with and guidance from Mayada El-Zoghbi, Karina Nielsen and Krisana Pieper from CGAP, and Mark Napier and Joe Huxley from FSDA.

1.3 Guidance paper outline

The guidance paper includes the following chapters and sections:

- Chapter 2 outlines the purpose and scope of the guidance paper;
- Chapter 3 covers the foundational principles of the overall approach for FSD impact evaluation, including what impact(s) to focus on, and how to develop an overall impact narrative which is methodologically robust;
- Chapters 4–6 provide practical guidance, organised in a number of logical steps (one to seven), regarding implementing this guidance;
- Chapter 7 provides guidance as to how an FSD can implement IOM throughout its programme cycle; and
- a glossary that provides the definitions of terms used in this paper is provided in Annex A. Readers are recommended to familiarise themselves with these definitions before reading on.

This guidance note can be read from start to finish, since its chapters flow in a logical order, from the building blocks of the approach, to guidance for setting up an IOM system, to approaches for reporting results. However, readers from some FSDs may only want to focus on particular aspects of M&E that are a current focus area(s) for them. **Regardless of the reader's focus, before moving forward with any type of implementation of the concepts presented in this paper, Chapter 3 (foundational principles) should be clearly understood.**

Additional background and technical papers have also been prepared for the FSDs and FSDA,[2] which provide more detail on the conceptual underpinning of the guidance, and further guidance on stand-alone topics (Table 1).

Table 1 Additional papers in support of IOM

Technical guidance notes
1. Impact-oriented measurement for FSD macro interventions
2. Linking FinScope/ FinAccess and poverty surveys
3. Methods for undertaking causality analysis
4. Logframes in the context of FSD and market development programmes
5. Tracking financial sector development: A practical note
6. FSD logframe indicators[3]
7. Benchmarking financial sector development
8. Impact-oriented measurement: Frequently asked questions
Analytical background papers
1. Tracking financial sector development
2. The relationship between financial sector development, economic growth and poverty reduction
3. Assessing the quality of access

2. Hosted by FSDA at www.fsdafrica.org/knowledge-hub

3. At the time when this guidance was developed in 2014/15. FSDK indicators are taken from their strategic-level programme indicators, which is their primary results framework.

2 Purpose and scope of IOM guidance

2.1 Purpose

FSDs face increasing pressure to show results while implementing complex, multi-faceted, market development programmes. M&E practices have to keep pace with this changing context and respond robustly to more demanding expectations, particularly the need to measure medium-term market system outcomes and longer-term impacts on poverty reduction. This is a challenge because FSDs exhibit a number of characteristics that differ from those of conventional programmes and that influence how their impact is measured. These characteristics include the change that they seek – that of financial sector (or 'system') change – and how to achieve it (see Table 2). The purpose of this paper is to improve the ability of FSDs (and their funders) to overcome these challenges.

Table 2 FSD characteristics and their implications for programme evaluation

	FSD features	Implication for programme evaluation[4]
The FSD approach	Aim to stimulate change through partners (i.e. facilitative role)	– Need to assess **both how FSDs have influenced partners,** and the contribution of their partners to the financial sector. – FSD-supported projects on their own may not result in substantial changes in 'access' or 'improved livelihoods' (especially in the context of macro and meso level projects).[5] FSDs therefore need to **measure how projects affect market dynamics** to produce such outcomes.
	Adaptive nature of FSDs (and markets) and ability to respond to opportunities	– FSDs need to quickly leverage new opportunities and therefore **need timely approaches to assess evidence.** – There is only so much up-front planning FSDs can do in terms of setting out expected causal chains, given the unpredictable nature of markets. In some cases it will be more important for FSDs to start an intervention **and let the measurement framework evolve** as they gain greater understanding. – FSDs may change their approach or invest in new areas. **Baselines therefore need to be applied pragmatically** and should be adapted over time as they could become obsolete.
	Significant variety in the types of projects being implemented	– Different projects have very different results chain/ impact pathways (e.g. advising the central bank versus scaling up informal savings groups), requiring **flexibility in measurement (e.g. different approaches, timelines etc.)** – Projects may work together or only work when other contributory factors are present. There is a need therefore to **be careful when attempting to aggregate** individual project impacts.
FSD objectives	Focus on systemic change	– Measurement frameworks **should not only focus on the end-users of financial services** but also how the structures and dynamics of the sector have changed.[6] – Interactions between various agents lead to changes that **can only be observed through a broad sector lens,** and not just through individual project monitoring focused solely on the direct partners with which an FSD works. – Evaluation needs to **take into account the additional indirect and sustainable effects** generated by changes in how the market functions.
	Attempting to influence a complex set of activities in the financial sector	– There are **many other non-FSD influences** on the financial sector, which makes establishing a direct link to the FSD programme difficult and thus rules out experimental and quasi-experimental approaches to IE at overall programme level. – As **impact is unlikely to occur in a linear fashion,** unintended and unexpected influences and results need to be identified. – This complexity (with many interconnections and interdependencies) may mean that **small interventions can have significant consequences.** Measurement frameworks therefore need to be particularly careful about being driven solely by size of investment. – There is only so much of the sector that an FSD will understand. In its planning and measurement approaches an FSD therefore **needs to rely on the understanding and perceptions of market actors.**

4. See also, Ton et al. (2011)

5. As per the making markets work for the poor (M4P) approach where 'macro' refers to the rules and norms of the market, 'meso' to the market infrastructure (channels, support services, information), and 'micro' to the transactions between the producers and users of financial services.

6. Traditional M&E relates to changes in stocks and flows of goods (such as the number of trainees, beneficiaries' incomes, number of packages of seed sold per month) whereas structural changes include a focus on relationships, behaviours, power etc. (see Ton, op. cit.).

All FSDs have, to some extent, existing measurement processes to meet these challenges. However, these vary greatly throughout the FSD network. These processes include FSDs tracking their own activities and the performance achieved by their investments against a logframe that summarises annual programme results for accountability and communication purposes. FSDs also produce annual reports that provide an overall narrative of the types of activities that they have undertaken and some indication of their contribution to changes in the financial sector. Five out of nine FSDs (including FSDA) are less than four years old, and to date only one FSD has conducted an evaluation of its programme impact. Existing practices tend to focus on the direct results of their projects, potentially missing the broad sectoral range of impacts an FSD may have.[7] **The IOM is therefore aimed at augmenting existing measurement processes, ensuring that they are oriented towards providing robust evidence for FSDs to carry out impact evaluation.**

2.2 Defining impact evaluation for FSDs

We use an adapted version of the Organisation for Economic Co-operation and Development (OECD) definition for impact: **'positive and negative, primary and secondary *medium* to long-term effects produced by a development intervention[8], directly or indirectly, intended or unintended'**.

The words '*medium* to' have been added to indicate that this framework is not only concerned with long-term impacts (normally associated with changes in the welfare of the poor), but is also interested in what may be relatively shorter term impacts on changing financial markets. The guidance paper is therefore concerned with 'impact' on a range of variables, not just the end objective of improved livelihoods for the poor.[9] These impacts can occur at project and at the overall FSD programme level.

For FSDs, impact evaluation should seek to answer two fundamental evaluation questions:

1. **What are the medium to long-term results of the programme?** This means moving beyond a focus on activities and direct outputs to a focus more on substantial and sustainable changes in the financial market and in the situation of end-users (households and enterprises).
2. **Are these results because of interventions made under the programme?** This means analysing the contribution of the programme in conjunction with other external factors.

This guidance paper hones in on specific steps that can be taken to ensure that the impact of the FSDs can be properly evaluated.

2.3 Benefits for FSDs and other stakeholders

This guidance recognises that FSDs are at different stages in their thinking and application in regards to results measurement and impact evaluation. **This guidance does not replace existing approaches, but provides the following benefits:**

a. promotes greater coherence across FSDs as to how they approach impact evaluation – leveraging existing approaches to gather insights and facilitate peer learning across the FSD network;
b. builds on existing processes (see Table 3 below) as well as providing guidance regarding areas in which FSDs have identified measurement challenges;
c. provides an in-depth and systematic approach for measuring *outcomes* and *impact* for FSDs (existing measurement manuals of FSDs largely focus on monitoring);
d. encourages FSDs to focus not only on the measurement of direct outcomes (e.g. number of households with access to finance directly caused by interventions) but also on whether underlying processes, and formal and informal rules (incentives) reduce the risks and costs of the financial services providers and users to provide/use these on a large scale (i.e. are changes systemic?) This helps to assess not only the current achievement of outcomes but the likely prospects of achieving scale and the sustainability of the outcomes in the future. This also confirms whether or not the modality of market development/making markets work for the poor (M4P) is working;
e. confirms that IOM need not necessarily involve complex/ expensive external studies. FSDs can initially use the implementing teams (and FSD partners) to assess themselves why the key changes of interest are occurring (or not occurring) and what factors (FSD-related and others) are driving this change in outcomes (and underlying factors);
f. encourages FSDs to work with funders (and other stakeholders) to prioritise the learning questions and strengthen and speed up the feedback loops between programme design, implementation and review, and to facilitate learning in order to adjust investments and improve performance; and
g. helps FSDs to strengthen accountability to their funders, by generating a robust evidence base that can be used in FSDs' regular reporting and that can be periodically tested by the independent evaluators.

This framework is primarily focused on FSD management and staff, but its audience also includes those that commission impact evaluations, including FSD governing bodies/investment boards, comprising bilateral and multilateral donors (referred to as 'funders' throughout this document), as well as – in some cases – representatives of governments. Other stakeholders (for example, policy-makers and market actors) also benefit from clear and timely evidence of what change is occurring and why, and how better data can be collected and used for improved decision-making by them.

Table 3 Adding value to existing FSD processes

	Independent impact evaluation (programme level)	Annual reviews (by funders) / annual reports (by FSDs)	Logframe	VfM analysis	Impact evaluations (project level)	Project monitoring
Current focus	Annual overview of how programme is performing (as represented by progress against logframe indicators) Some focus on VfM and risks faced by the programme Communication to a wide-ranging audience (e.g. illustrative project stories)	Sets out a simple overview of how FSDs may contribute to impact (i.e. a programme intervention logic) Provides annual review of progress of the programme, largely for accountability purposes (i.e. to funders)	In-depth analysis of economy, efflciency, effectiveness and equity Cost–benefit analysis of monetary benefit created by (some/all) projects, against cost of the programme	In-depth analysis of economy, efflciency, effectiveness and equity Cost–benefit analysis of monetary benefit created by (some/all) projects, against cost of the programme	Provides an in-depth assessment of a specific FSD intervention and its direct impact	Captures qualitative and quantitative results to confirm if an intervention is working or not, and what changes are taking place
IOM's added value	Provides a coherent and robust approach to developing the evidence base (including objectives) for impact evaluation, while providing real-time evidence for programme adaptation	Provide a robust evidence base for interrogating numbers within the logframe, and the percentage of results attributable to the FSD Focus measurement beyond logframe to capture evidence of all FSD programme contributions, including to systemic market change Highlight broader changes in sector ('sector tracking') to strengthen FSD narrative	Can provide evidence of whether/ how the programme ToC is working, and if logframe needs updating Over time, IOM can provide logframe indicators that measure changes in the underlying structure of the market (i.e. not just direct FSD impact)	Sets out in-depth ToC (at programme level) and results chains (project levels), to identify potential areas of monetary benefit Provides robust evidence base for calculating percentage of monetary benefit attributable to the FSD	Somewhat limited – provides assistance in regard to choosing which projects may require standalone evaluations, and what types of questions and methods might be used, but IOM is largely focused on programme impact	Builds on current monitoring processes and focuses on evidence that will be useful for assessing impact (including causal pathways, systemic change indicators, sector tracking, and programme aggregation)

7. We do note, however, that some FSD programmes are relatively recent and would not have been expected to undertake a programme evaluation at this point.
8. Intervention is the action or process of intervening with the intent of modifying an outcome.
9. This framework uses 'improved livelihoods' as a broad category for FSDs' final impact, acknowledging that there are variations across FSDs in regard to how this is defined.

2.4 Scope

The scope of this guidance paper has three important parameters:

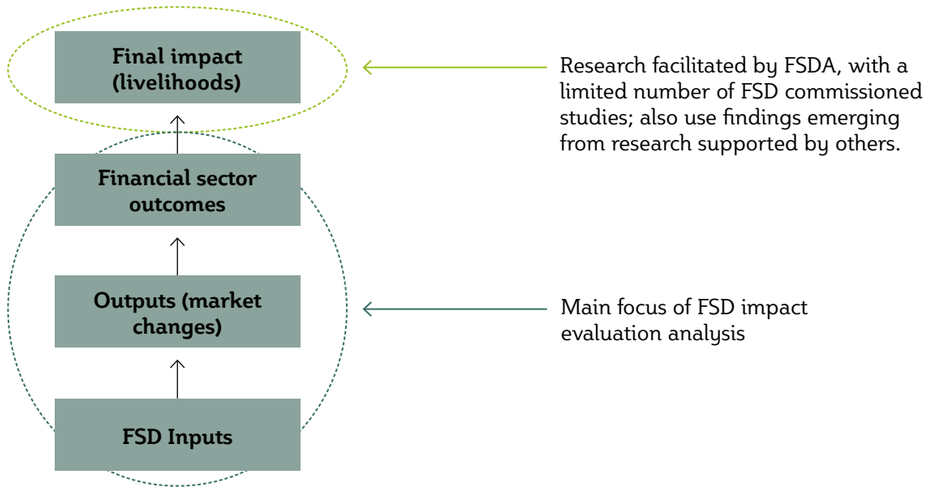
1. it primarily **focuses on inputs to financial sector outcomes:** for most FSDs, impact on livelihoods and economic growth is largely beyond the scope of what they measure;
2. it focuses on **overall FSD programme impact evaluation,** and therefore does not provide in-depth guidance on how to undertake an impact evaluation for the variety of individual FSD projects; and
3. it does not provide a prescriptive step-by-step manual for FSDs, but provides overall guidance to FSDs **to integrate, and augment their existing M&E systems.**

Focus of impact for an individual FSD: As highlighted in Figure 1, the approach adopted in this framework suggests that the main focus of FSD measurement should be on examining the effect of its inputs on developing a financial sector that works for the poor. This is what the IOM concentrates on. FSDA–FSD-facilitated studies, and other non-FSD/ global research, can complement this measurement by focusing on assessing links between financial sector outcomes and livelihoods (both directly and indirectly through promoting economic growth).

Discussion point: During the consultation it was highlighted that in some instances, analysis on financial usage at a household level can be very close to discussing welfare effects in terms of issues such as consumption smoothing and resilience to shocks. Indeed the diagram in Figure 1 does provide some artificial separation and where possible the IOM recommends an FSD trying to understand results of its contribution on poverty reduction, and how the poor are using financing services. However, IOM recommends that the majority of an FSD's measurement resources focus on its contribution to changing the financial system.

Tip: Individual FSDs may still undertake impact evaluation studies (particularly approaches that prioritise learning) that focus on poverty reduction but the IOM provides guidance as to where the balance of resources should be invested. However, FSDs will still need to continue to invest in non-impact evaluation research on livelihoods (for example, financial dairies) to improve their understanding of how the poor use financial resources and services in their particular country.

Figure 1 FSD impact measurement focus



Programme impact: The IOM is focused on providing guidance on the overall FSD programme impact, the sum of all FSD interventions, including indirect results (also called spill-over effects). It provides some guidance on measuring specific types of interventions, particularly in terms of identifying useful project indicators, and types of causality methods to the extent that this helps inform an overall assessment of programme impact. Undertaking impact evaluation for each individual intervention or project in isolation is methodologically a relatively simple proposition and is already well covered in existing global impact evaluation guidance in the financial sector.

Guidance (not instructions): The guidance contained in this paper takes for granted many of the building blocks that make up good measurement for all programmes.[10] Its aim is to provide core approaches and principles that will allow FSDs to extend their existing systems to improve their measurement of outcomes and impact (see Table 3). It does not provide a comprehensive step-by-step manual that must be followed in its entirety; intervention in complex market systems requires flexibility and creativity, as does measurement of such intervention (see Box 1).[11] However in areas that are relatively new or challenging (for example, measuring systemic change) it does provide a number of practical steps that FSDs can follow, as well as checklists at the end of each step of the guidance.

Box 1 IOM system – an evolving and pragmatic approach

Impact evaluation for FSDs (and market development programmes more generally) is a relatively new and dynamic arena of analysis. Indeed, there are areas of measurement where the IOM does not provide the complete answer. Rather, the FSD community will need to undertake further analysis (for example, in measuring the quality of access from a user's perspective). Moreover, FSD programmes themselves are dynamic: they change the way they intervene in response to their own learning, and to market changes. **Therefore this framework should not be viewed as the final word in impact evaluation for FSDs. It provides a platform for FSDs, on the basis of which they can experiment and learn.** Future learning events, organised by FSDA, will help to update sections, add new sections, and delete parts of the guidance that are no longer relevant.

On a related note, **FSD measurement frameworks should evolve as the FSD's internal capacities increase.** Some FSDs are new and are still setting up their operations. The IOM will help establish some of these building blocks for measurement (e.g. thinking around project indicators and logframes), including assisting those following the DCED standard. FSDs will need to be pragmatic in regard to how far they implement the more complex aspects of the guidance, alongside more straightforward monitoring of their initial interventions. FSDs also have to ensure that any improvements in the measurement system continue to meet the core reporting requirement of funders.

2.5 The outline of the IOM

The IOM can be embedded in FSD operations at a number of FSD programme cycle points. These are noted throughout this guidance paper and summarised in Chapter 7. These points include:

a. the start of an FSD programme (in terms of developing a strategy and corresponding results framework);
b. project investment decisions; and
c. periodic monthly, quarterly, annual, and end-of-strategy review processes.

Embedding IOM at these points in turn provides opportunities for assessing project-level and programme-level impact to varying levels of depth and robustness, as well as for reporting to different stakeholders.

10. For example, setting realistic results, using indicators that are aligned with interventions and are verifiable, with data collected and analysed in robust ways.

11. This is consistent with the M4P operational guide (Springfield Centre, 2014).

The following table sets out the chapters in this document that cover the main steps in the process of embedding and implementing the IOM:

Table 4 Summary of IOM guidance

Chapter/stage	Steps	Timeline
Foundational principles (Chapter 3)	Embedded throughout measurement practices	
Stage 1: Clarity of purpose (Chapter 4)	1. Ensuring the FSD ToC is evaluable 2. Developing impact measurement questions	Start of the strategy period
Stage 2a: Measuring change: What happened? (Chapter 5)	3. Developing indicators 4. Data collection methods and sources	Throughout implementation
Stage 2b: Measuring change: Why it happened (Chapter 5)	5. Assessing causality and contribution 6. The research agenda	
Stage 3: Bringing it all together (Chapter 6)	7. Developing a credible narrative	End of strategy (with periodic checks over the programme strategy cycle)
Implementing the IOM (Chapter 7)	Embedded throughout measurement practices	

3 Foundational principles of IOM

This section outlines five foundational principles FSDs should follow in order to implement IOM:

1. **aligning monitoring with measuring impact (the 'sweet spot');**
2. **using the FSD ToC as a strategic framework for planning and impact evaluation;**
3. **focusing on the FSD programme's primary interest (as regards measurement), which is assessing changes in inclusive financial markets;**
4. **identifying and measuring systemic change; and**
5. **measuring impact from the perspective of both the FSD programme ('bottom-up') and the sector/ market system ('top-down').**

Principle 1: Aligning monitoring with measuring impact (the 'sweet spot')

Monitoring has traditionally been seen as an in-house and somewhat narrow technical process of tracking performance against pre-defined indicators, usually quantitative. It tends to focus on the short term: mainly activities and outputs or, at best, immediate or very short-term outcomes. In this traditional approach monitoring was designed to reassure, or, in the event of divergence from expectations, to trigger restorative action – to get the intervention back on track. This is important but monitoring needs to go further; **in particular, monitoring needs to extend further into outcome territory, as well as to involve periodic reflection on progress and results to obtain a perspective regarding how and why change might have happened.**

When implemented in isolation, the traditional concepts of M&E have been shown to fail (see Box 2), particularly in the case of dynamic and complex programmes such as FSDs.[12] A central proposition of the IOM system, as shown in Figure 2, is therefore to bridge the gap between traditional monitoring and impact evaluation – finding the 'sweet spot' that **aligns an FSD's monitoring system with the objectives of impact evaluation.** In other words, the idea is that FSDs improve their existing monitoring system so that they produce the reliable, credible data that will be needed later to perform an assessment of their impact.[13]

Box 2 Why do evaluations fail?

There is a gulf between programme directors/managers and evaluators. Managers sometimes lack access to (or even appreciation of) key evaluative skills, whereas evaluators complain about poor monitoring data. Evaluations also come too late (in terms of programme implementation) and too infrequently to give useful management information. Furthermore, evaluators struggle in short visits to understand complex and dynamic programmes.

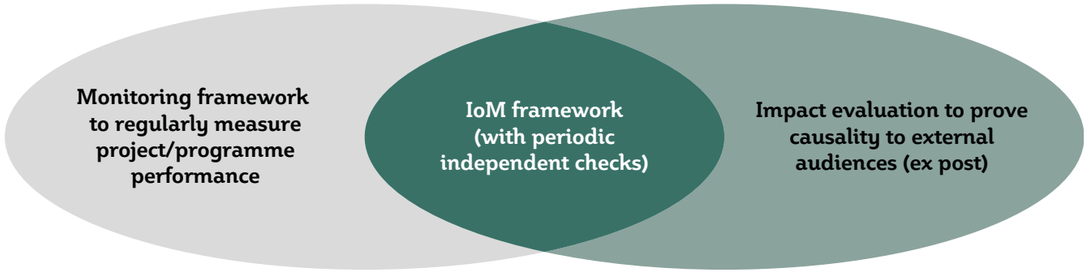
This complexity has made proper monitoring more urgent. Programmes increasingly aim to use relatively small amounts of money to catalyse pro-poor change in large markets. Interventions are necessarily experimental, meaning that managers really need to know in real time whether they are working or not. Adjustments in approach will very often be needed, and such management decisions need to be informed by data on outcomes – which needs evaluative expertise.

Evaluation also needs all stakeholders to be clear about the logic of the programme. Evaluations of programmes with well-articulated logic – and high-quality monitoring data to test that logic – will always be quicker, easier and cheaper. Evaluators will need to validate the monitoring data through triangulation, random checks and other means – but at least there will be something to work with. *Source: Adam Kessler and Jim Tanburn, DCED (see DCED 2014a)*

12. As the recent DFID (2014) evaluation policy states, 'Programme monitoring and review activities can be used to inform evaluations. The complementarity of evaluation to these other evaluative activities lies in the potential for evaluation to provide a deeper and broader understanding of an intervention or look across a set of interventions to reach robust conclusions and form useful recommendations about what needs to change to reach development goals.'

13. At a practical level, with multiple interventions/ projects, it is also prohibitively expensive to measure all of them through stand-alone impact evaluation measures. Therefore pragmatic approaches are needed to strategically establish and use routine monitoring systems.

Figure 2 The sweet spot (between monitoring and impact evaluation)



To hit the 'sweet spot' means that:

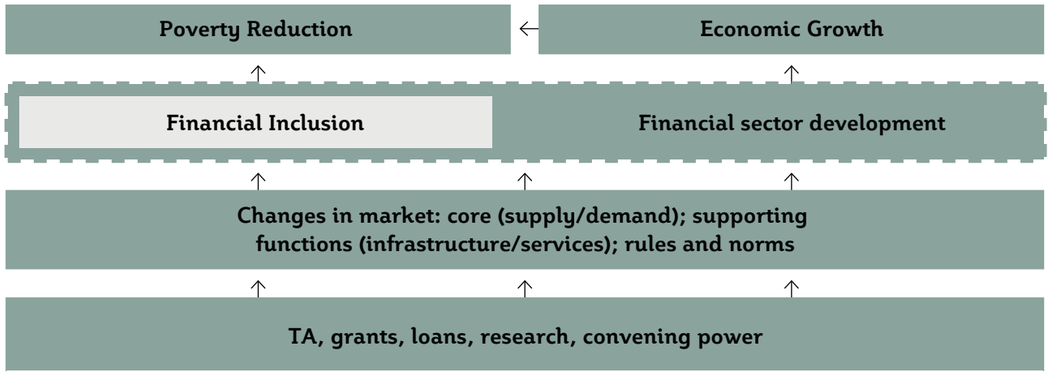
1. **indicators that are monitored regularly are designed to be useful for impact measurement,** i.e. they measure outcomes beyond direct results of interventions and look for evidence of an FSD's contribution;
2. **periodic assessments of critical causal links within the ToC will be embedded in the regular M&E activities** during programme implementation, not just at the end when the entire programme is evaluated. FSD programmes aim to adapt to changes in the market environment and timely information about their impact is therefore important;
3. the ToC will define the changes that need to be monitored and the causal links that need to be evaluated periodically. **As evidence is collected, the ToC itself will need to be reviewed and revised;**
4. **measurement frameworks are evolving and focus on learning, given the unpredictability of ex ante identification of the precise nature of change.** An FSD will need to balance its theory of what it expects to happen with implementation experience regarding the types of impacts that are most important to the programme; and
5. **the evidence collected will also feed into the ex post impact evaluation** of the programme.

Principle 2: ToC – a strategic framework for planning and impact evaluation

Made up of various interventions, FSDs are coherent programmes that normally focus on implementing a multi-year strategy to promote financial sector development.

The strategy should be underpinned by an explicit ToC, informed by market analysis that sets out the systemic constraints to achieving pro-poor outcomes in the financial sector.[14] **It is the evaluation of this strategy, and its underlying ToC, that will define the focus of an impact evaluation.**[15] That is, how FSDs' inputs lead to changes in the market system (at the macro, meso, and micro levels) – the outputs[16] – and how these in turn promote financial sector development, and subsequent impacts on livelihoods (see Figure 3).

Figure 3 A generic FSD ToC

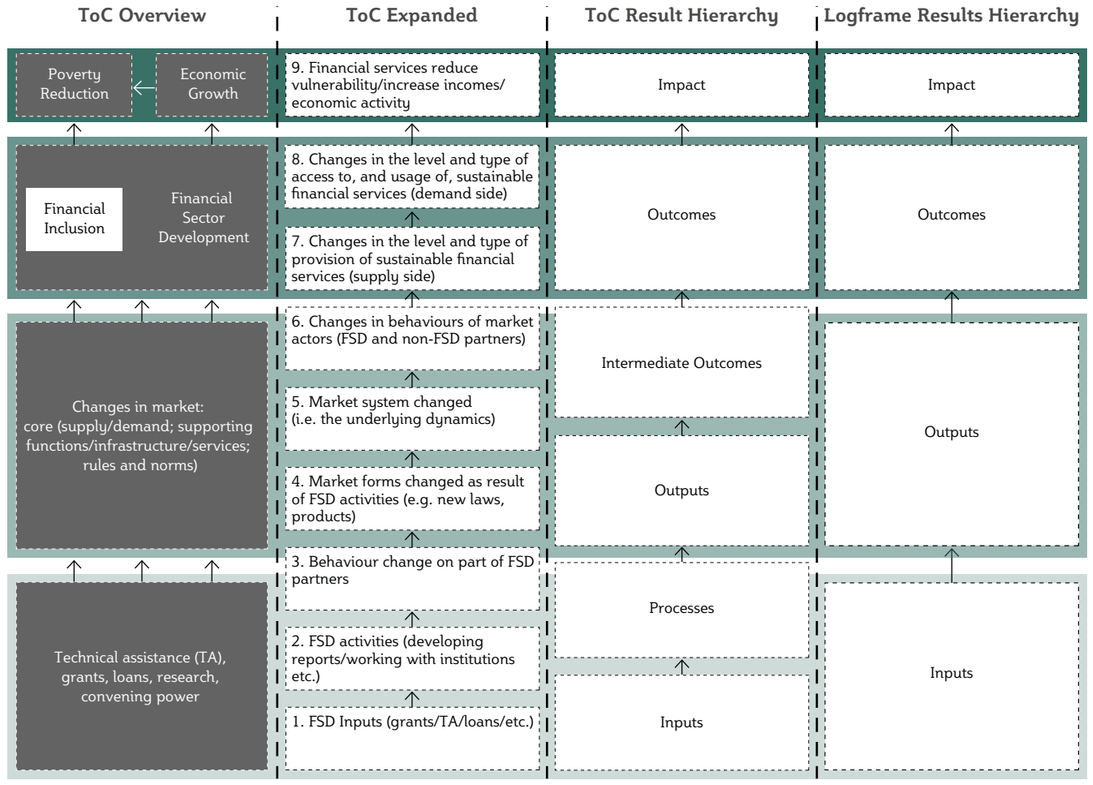


14. This may include a broad focus on the sector, or specific areas, such as mobile money, or rural finance.
15. As discussed in Chapter 7 this does not imply that assessment of impact will only occur at the end of the strategy period; FSDs will need to build in periodic checks to internally assess their progress towards intended impacts.
16. As set out in the DFID logframe (see Figure 4).

Tip: It is important to note that many of the indicators used to measure 'programme outputs', as referred in the DFID logframe (which all FSDs use), will sometimes be articulated as outcome indicators that are used to assess changes at the market level. Figure 4 shows how this 'output' level of the logframe can be unpacked, from assessing the direct change in a market form (e.g. a law, or a financial institution) to how these changes may affect, or are affected by, the underlying dynamics in the market. It is important that the FSD and its funders have a clear and shared understanding regarding indicators for which an FSD should be held accountable, reflecting how much control an FSD can have over a potential change in that indicator. A logframe may include these as a long term output measure (i.e. a target indicator to achieve after a number of years), or as an outcome measure, to which an FSD contributes. Indeed with more flexibility being introduced into accountability frameworks, such as logframes, it may be possible to include these separate stages, outputs (market forms), intermediate outcomes (underlying market changes), and outcomes (changes in provision and use of financial services).

We will refer to this diagram throughout this guidance document. Figure 4 illustrates how it corresponds to different terminologies for results reporting at a FSD programme level.[17] It is important to note that some of the indicators applied to measure 'programme outputs', as referred in the DFID logframe (which all FSDs use), will actually be articulated as outcome indicators that are used to assess changes at the market level. Figure 4 therefore notes that there are direct outputs that are in FSDs' control, and FSDs are thus accountable for them. The IOM refers to these as changes in market forms, for example, a change in rules, or a new innovative business model an FSD has helped bring about. In an FSD ToC, underlying systemic change resulting from these market form changes (and other influences) can be viewed as an intermediate outcome, leading to final outcomes related to increases in sustainable financial inclusion.

Figure 4 ToC and results hierarchies



17. Note: EFInA has a slightly different results framework, with final impact also focused on financial inclusion rather than livelihoods.

A theory-based impact evaluation (TBE) approach is an appropriate framework – or evaluation design – for measuring FSDs' impacts (see Annex A definitions). The basis of TBE is that *causal inference* can stem from the identification/ confirmation of causal processes or 'chains' and the supporting factors at work (i.e. the financial sector context), rather than from a specific counterfactual such as another financial sector against which comparisons can be made.[18] **To that end, claims regarding FSDs causing impacts are best described as contributory –** rather than saying that impact can be attributed solely to interventions (see Box 3).

Box 3 Assessing contribution or attribution at programme and project level

Attribution is a measurement of how much of the observed impact has been achieved solely due to (and can be attributed to) the FSD intervention. **Contribution** makes a causal claim about whether and how an intervention has contributed (along with other factors) to an observed impact (Stern, *et al.* 2012). The difference between these two broad approaches is sometimes mistakenly equated with that of quantitative vs. qualitative methods. Whilst theory-based approaches (which measure contribution) do tend to include the use of qualitative methods, they should also use quantitative methods to examine the evidence, in order to confirm whether the results chain is working in practice. In other words, a mixed methods approach is best when seeking to identify contribution.

A TBE essentially has two parts:[19]

- **a conceptual framework** that sets out the programme theory, such as the ToC and strategy outlined above; and
- **an empirical approach** that investigates how an FSD programme has caused the intended or observed outcomes (i.e. those outcomes that may or may not have been intended). Within this empirical enquiry a range of methods can be used to assess causality.

Tip: It is not possible to conduct a study that attributes all of the observed changes to the FSD. Contribution analysis is better able to tell the causal story of programme impact. However, for certain specific projects the observed impact may be more clearly attributable to the FSD, and thus methods which rely on statistical attribution may be appropriate. See Step 5 in Section 5.3 on causality methods for more guidance on this.

Principle 3: The primary measurement interest for FSD programmes is in assessing changes in inclusive financial markets

It is necessary to recalibrate our use of the term 'impact'. In the field of impact evaluation the term 'impact' is generally used to refer to the long-term changes at the end of the results chain – in this case improved livelihoods linked to economic growth and poverty reduction. As illustrated in Figure 4, however, FSDs (like all market development initiatives) have a long and complex intervention logic. FSD inputs entail a lot of what might be described as intermediate impacts on the pathway to the final impact (these intermediate impacts are identified as outputs or outcomes in the logframe hierarchy). Such impacts occur, as shown in Figure 4, in such areas as the behaviour of FSD partners (box 3 in the figure); market forms (box 4); market systems (box 5) etc. Some of the impacts, especially at the higher outcome level, will go beyond the effects on the direct FSD partners to changes in the behaviour of non-FSD partners or the functioning of the system as a whole (through replication, demonstration or crowding in, as explained in Section 5.1.3 below).

The IOM approach does not focus solely on the long-term final impact – it uses the tools of evaluation and impact evaluation to test the operation of the ToC on a continuous basis, producing insights that can be used to improve the effectiveness of the FSD programme. The market development approach on which FSDs are based requires such acts of investigation, reflection and learning at various points of the ToC throughout a strategy period. The primary purpose of these acts of investigation, reflection and learning is to provide real-time information to improve the targeting and management of the programme – but the process will also significantly improve the evidence base for an overall impact evaluation in the traditional sense, using a TBE design.

It is important for an FSD and its governing body to decide: (a) how far along their ToC chain towards the final impact on poverty reduction they think it is feasible and cost-effective to go when seeking to measure impact; and (b) at what points along that chain they want to concentrate their resources to collect evidence. This in turn will shape the focus of the IOM, i.e. the types and sources of data collected, the 'granularity' of data for different parts of the ToC, and the rigour and complexity of the analysis for particular causal pathways. It is important to note that these decisions will relate to the allocation of resources for *collecting evidence to measure impact* at various points along the results chain, not to the resources invested in the inputs for a particular intervention: there is by no means a direct and automatic correlation between the portion of FSD resources invested in implementing a specific input (and therefore pathway) and the proportion of M&E resources devoted to investigating that specific pathway.

From this perspective examples of impacts that an FSD may define as the impact(s) of interest for their IOM system are:

- **Overall financial sector development:** Are the various functions of the financial sector (e.g. mobilisation and allocation of resources) being performed effectively?
- **Financial inclusion:** Whether the financial sector is providing and successfully delivering appropriate (quality) and affordable financial products and services, and if these are being used by poor people and micro, small and medium enterprises (MSMEs).
- **Market development:** How financial sector outcomes were achieved through the M4P approach and whether these offer prospects of sustainable access at scale (see Section 3.4 below).
- **Improved livelihoods:** (a) what are the *direct* effects financial services have on the livelihoods of the poor (e.g. reduced vulnerability to shocks, increased incomes and increased employment)? (b) what are the *indirect* effect of financial sector development and subsequent economic growth on improved livelihoods?[20] Evidence as to the causal pathways for this impact is likely to be largely generated outside of a specific FSD's IOM system.

Principle 4: Measuring systemic change

A core objective of FSDs is to transform the sector or system such that it become sustainable and resilient for large numbers of poor people and small businesses. Thus, there is a need to assess whether **'systemic change'** has occurred. This contrasts with traditional investments to promote financial sector development and financial inclusion that tend to focus on the direct impact of a specific project(s), and not the underlying dynamics of the system.

We have defined systemic change 'as a transformation in the structure or dynamics of a system. *The systemic change in which FSDs are interested is change that leads to impacts on large numbers of poor people and/ or small businesses, either in their material conditions or in their behaviour.*'[21]

Discussion point: All FSDs noted that the main challenge with measuring impact comes from the many other factors that contribute to financial sector change, thus making it hard to disentangle effects of FSDs' contributions. The IOM accepts this, and provides guidance on how FSDs can use their ToC to bring some focus to this question, putting parameters around what it is trying to measure. But it also encourages an FSD to not shy away from trying to understand and measure these broader, and complex financial system changes

This means that the measurement lens must focus not just on end numbers (e.g. in respect of financial access) but also on how the underlying structures of the market have changed to support these outcomes. For example, how financial sector skills/capability and infrastructure, formal and informal rules (incentives), and inclusive business models collectively reduce the risks and costs of the providers and users of the financial services. This helps to assess not only the current achievement of outcomes but the likely prospects of achieving scale and the sustainability of these outcomes in the future.

Tip: To aid measurement, an FSD should have a shared understanding (i.e. amongst staff members) regarding their definition of what systemic change is, and what it looks like if it occurs.

Currently, FSDs are arguably better at measuring the number of people on which they have had an impact (through numbers of accounts, for example) rather than how the underlying dynamics of the system have changed to facilitate this, and to what extent financial sector outcomes are sustainable and resilient. Systemic change is not in itself the final objective for FSDs; rather, it is a process for achieving sustainable and appropriate financial services for the poor. **Therefore the IOM argues that FSDs need to clearly measure both the systemic changes that have occurred and how they have occurred.** The contribution of these changes to growth and improved livelihoods is also important. Often only limited human and financial resources are available for measurement. In such a context, FSDs need to prioritise analysing their contribution to changes in financial inclusion. They can draw on wider research or use proxy indicators (see Step 3 in Section 5.1 below) to confirm the links between financial inclusion and economic growth.

18. See, Stern et al. (2012). Experimental (e.g. randomised control trials (RCTs)) and quasi-experimental approaches are considered inappropriate for assessing programme impact.

19. Coryn et al. (2011).

20. A separate paper provides an in-depth discussion on the links between the financial sector development, economic growth and poverty reduction. See "The relationship between financial sector development, economic growth and poverty reduction.

21. Adapted from Ruffer and Wach (2013). This definition allows for systemic change to also capture negative impacts.

Principle 5: Measure impact from the perspective of both the FSD programme and the sector/market system

Perhaps the easiest way to measure the impact of an FSD is to assess and aggregate its various interventions. This would analyse the extent to which the FSD intervention has led to market-level change, and then the extent to which this change has led to an improvement in programme outcomes and impacts. Such an approach would largely rely on data provided by an FSD M&E system. We refer to this as a **'bottom-up' approach,** in that it attempts to track changes from the bottom of the ToC (Figure 3) to the top.

However, as noted in Table 5, there are potentially significant problems with such an approach, namely the difficultly of connecting changes in the broad financial sector to specific FSD interventions, given the multitude of other players and influences. In other words, an FSD risks being intervention-centric (e.g. as a result of 'self-importance bias') and might assume that changes are largely brought about by FSD interventions. **This is a particular risk for market development programmes where changes are unpredictable and are influenced by many players.**

Table 5 Advantages and disadvantages of bottom-up and top-down approaches

	Advantages	Disadvantages
Bottom-up	– Clear link with projects, and aligned with logframe and annual reports – Real-time information for understanding performance and contribution – Partly reliant on monitoring data, avoiding costly large-scale (programme-level) surveys	– Extent to which you can aggregate indicators (being project focused) may be limited, missing synergies across projects – Risk of self-importance bias linking changes in the overall system to relatively small projects[22] – Risk of failing to capture the results of systemic changes, such as replication and adaptations in the market – Limits as regards understanding all the factors influencing sector change and livelihoods when having to attempt to attribute them to FSD projects
Top-down	– Focuses directly on how things have changed at the level of the impact of interest – Helps FSDs identify unexpected impacts (as not primarily focused on expected project impacts) – Looks at all contributory factors, including non-FSD influences	– Attribution back to the programme is limited (thus reducing accountability), and potentially very expensive if tried – Approaches that try to measure the whole programme from baseline to end-point may become obsolete given the dynamic nature of FSDs (e.g. target groups change) – Linking FSD intervention to broader change is difflcult

Source: Adapted from Ruffer and Wach (2013)

We therefore argue that this bottom-up approach should be augmented with (not replaced by) other methods (as shown in Table 5, these also have weaknesses if used by themselves). We recommend two additional types of measurement:

1. A top-down monitoring approach – also referred to as 'sector tracking': this would track indicators of interest across the broad financial sector (for example, FinScope is already one method that FSDs use to do this).
2. A top-down evaluation approach: assessing what has caused the changes observed in top-down monitoring.

Tip: Top-down analysis is likely to be done less often, meaning that changes in the market can occur between updates.

The differences between the two approaches are further summarised in Table 6.

22. The emphasis on proving a causal link between the intervention and outcomes along predicted, linear causal pathways means that monitoring can be blind to a broader range of outcomes associated with the intervention or system (intended, unintended, positive or negative), alternative causes from other actors and factors, and the full range of non-linear pathways of contribution. See 'Evaluating systems and systemic change for inclusive market development: Literature review and synthesis, ACDI/VOCA, LEO Project', USAID June 2014.

Table 6 Characteristics of top-down and bottom-up approaches

Direction of measurement		Description	Objective	Example
Top-down	**Monitoring approach ('sector tracking')**	Tracking changes in livelihoods, economic growth, financial sector	– Identifying changes in the financial sector beyond those captured (or thought about) in FSD interventions/ frameworks – Helps improve programme design, and provides an evidence base to confirm if expected changes are occurring	Is efficiency of the formal financial sector improving?
	Evaluation approach	Assessing what has caused changes in livelihoods,[23] economic growth,[24] financial sector	– Assesses the causes of changes in the financial sector from a non-FSD programme perspective – Provides in-depth focus on specific complex pathways found towards the top of the ToC[25]	What has caused the changes in the efficiency of the financial sector?
Bottom-up	**Monitoring Approach**	Tracking the performance of FSD projects	– Provides information to help improve project performance – Provides record of key achievements	Have FSD partners improved their efficiency?
	Evaluation approach	Assessing if a project's intervention led to observed outcomes	– Provides evidence regarding if FSD-specific interventions worked or not, and why	To what extent have FSD interventions (e.g. use of better processes, technology or expanding agency banking) contributed to changes in an FSD partner's efficiency?

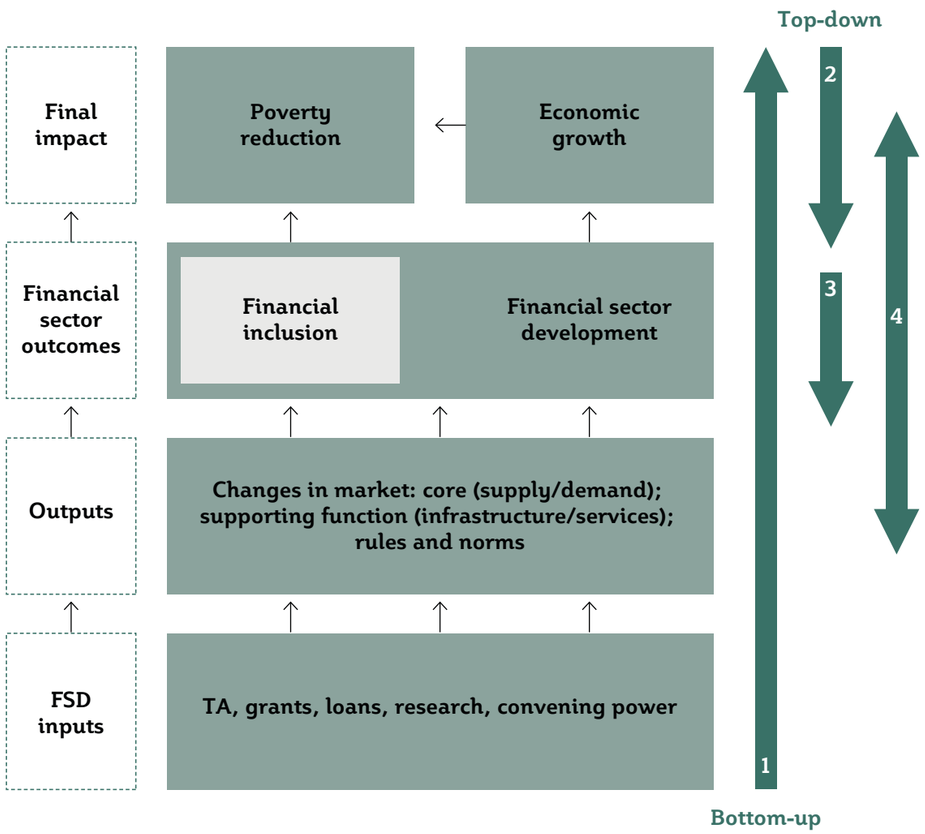
The robustness of evidence independently collected through a bottom-up or a top-down approach becomes weaker as the investigation extends further from the starting point (as other factors influence the causal chain) – as illustrated by the lighter parts of the shaded arrows in Figure 5. For example, it is hard to assess how small changes in one or two banks led to an observed change in overall financial access (top-down). **A combination of both approaches is therefore recommended, in order to maximise their usefulness: one should seek a syntheses of both approaches at the point where they meet, in order to develop an overall FSD contribution narrative.**[26] This approach would help assess the extent to which results achieved by an FSD's interventions (measured 'bottom-up') are consistent with the market-level factors that have driven changes in outcome/ impact level indicators (measured 'top-down'). Step 7 in section 6.1 discusses how to build an overall credible narrative of the FSD impact.

Tip: Shown in relation to the overall programme ToC in Figure 5, a top-down and bottom-up measurement approach can also be applied to FSD themes or flagship projects (see Step 1 and Step 2 below).

Combining both approaches may seem like additional work for an FSD but it is important to ensure that we are identifying all changes that matter, drawing the right conclusions regarding the magnitude of that change, and identifying all factors (including the FSD contributions) that are driving these changes.

23. As noted above, the examination of the links between the financial sector and livelihoods may be assisted by research coordinated by FSDA, rather than separate research by each FSD.
24. As noted above, the examination of the links between the financial sector and economic growth may be assisted by research coordinated by FSDA, rather than separate research by each FSD.
25. This recognises that linking FSD interventions directly to these broad changes can be difficult.
26. There are other examples of such a framework in practice, both in market development programmes (see GEMS, 2012), and in other complex evaluations (see the OECD's budget support methodology).

Figure 5 Bottom-up and top-down measurement framework



1. FSD programmes measuring direct results of investments/ interventions (from inputs, to outputs, to outcomes, to impacts)
2. Based on changes in households and/or the economy, FSD programmes can assess the range of key factors that contribute to this change (FSDA-supported knowledge management/ research can help)
3. Based on changes in levels of financial sector development and financial inclusion, FSD programmes can assess the range of key factors that contributed to this change
4. Provide an overall contribution narrative of FSD impact by triangulating the two perspectives of evidence – assessing to what extent changes in impact that are of interest were contributed to by FSD programme's interventions

Box 4 Foundational principles checklist

- Is the monitoring system set up to assist evaluation? Are there processes to:
  - articulate the programme's anticipated impacts through a theory of change (see Step 1)
  - ensure that there is a clear definition of the systemic change that the FSD is trying to bring about and indicators (see Step 3) to track this change are available? Is this understanding shared throughout the organisation?
  - ensure that the IOM framework is embedded within normal FSD programme processes and that the responsibilities of the implementation team (e.g. project managers) and the M&E team are clearly
    articulated (i.e. not just seen as the responsibility of the M&E team)
  - identify the sector/market changes beyond monitoring of specific FSD projects? (also see Steps 3 and 4)
  - alert you to emerging outcomes that indicate that the interventions are not operating as anticipated, and that can help you to adjust your strategy and/or your ToC?
- Do you have an agreed plan with funders regarding the arrangement for evaluation?
- Do you also have a process in place to learn about what works and what does not work in measuring your impact?
- Do the principles and guidance contained in the IOM help FSD directors to better manage their programmes? If not, what should change in future editions of the IOM guidance to ensure that this happens?
- What additional guidance and support (TA or training) might you need, and can FSDA assist in this regard?

IOM – Chapter 4: Clarity of Purpose (Stage 1)

Chapter 4, Clarity of purpose, covers Stage 1 of the process of implementing the IOM guidance.

This stage is focused on setting up the IOM for the FSD, ensuring that there is a shared understanding of the programme's main objectives, and what the FSD is seeking to measure. This stage is more relevant at the start of a FSD's strategy period.

Stage 1 is broken into two steps:

Step 1 – Ensuring that the FSD ToC is evaluable: This step sets out the FSD's programme theory of change, and how this will influence the measurement process.

Step 2 – Developing impact measurement questions: This step provides a coherent set of measurement objectives to orientate the implementation of the IOM.

4.1 Ensuring that the ToC is evaluable (STEP 1)

4.1.1 Overview

- This step focuses on setting up the IOM, which is most likely to take place at the start of a new strategy period
- Setting out what is to be evaluated – the Programme ToC – is a pre-requisite, to help determine what impacts the FSD is seeking to achieve, and how they are expected to be achieved
- This step provides guidance on how to assess whether the ToC is evaluable, and how to improve its ability to guide subsequent measurement
- To provide detailed analysis of causal pathways it will often be necessary to link a programme ToC with thematic ToCs and specific individual project results chains

4.1.2 The ToC and its role in the IOM system

A ToC provides the basis for IOM. It is important that FSDs explicitly articulate a ToC that reflects the programme (and through it the projects, which nest within this ToC) that will be evaluated. Annex B lists some issues to consider in developing a programme ToC, with the focus here on assessing if it will provide a framework for implementing the IOM guidance.[27]

As elaborated in Box 5, throughout this guidance we use the term ToC to represent causal pathways and the context (and assumptions) in which they operate at programme level or in a thematic/pillar area (i.e. a group of interventions closely connected to a broad objective of an FSD, such as digital financial services). For measurement purposes, the ToC should clarify the main impacts and causal relationships it wants to test as part of the impact evaluation, and the context in which the FSD is operating.[28]

This will form the basis on which the entire measurement system will be developed, including the impact measurement questions (see Step 2), definition of indicators (Step 3), and overall programme evaluation approaches (Step 7).

Tip: It is important to remember that ToCs are best-informed guesses of how market development change will proceed, and will need to be periodically reviewed and updated in the light of changing circumstances or emerging evidence. Unexpected changes in the financial sector or the environment may require adjustments to the causal pathways in the ToC, or evidence arising from the implementation of a programme may reveal that one or more of the assumptions on which the ToC was built was mistaken and needs adjustment.

Box 5 ToC, results chain, logframe – how do they relate?

Many different terms are used to describe how a programme intends to achieve its impact. This can be confusing, not just because of the number of terms used, but also because the same term can mean different things to different people, and sometimes the terms are used interchangeably (see Figure 4 above). We provide common definitions of these terms in Annex A, although we are conscious that these definitions are not the only ones in use.

For multi-project programmes such as FSDs, these tools fulfil different functions. The **programme logical framework (logframe)** is a standard requirement for development funding, with FSDs having to develop and report on a logframe using a set matrix template for many funders. The primary use of a logframe is therefore an accountability tool.

A **ToC** is typically seen as a precursor to and expansion of the logframe, articulating why the series of results is expected to occur and focusing not just on the 'boxes' but also the 'arrows connecting the boxes' of a traditional logframe. ToCs look at the assumptions (implicit or explicit), risks and external factors that are important in relation to results being achieved.

The term **results chain** is used to describe a more detailed causal pathway, focused on an individual intervention or project. The project results chain should be consistent with the overall ToC (see 4.1.6), and from a measurement perspective it will provide much of the evidence base ('the results') that will be used to test the overall programme.

Tip: Though ToC and results chains are often presented as visual figures, the narrative explaining the programme logic is key and should supplement any visual, in order to help the reader and to avoid making the figures very complex.

4.1.3 Articulating systemic change in a ToC

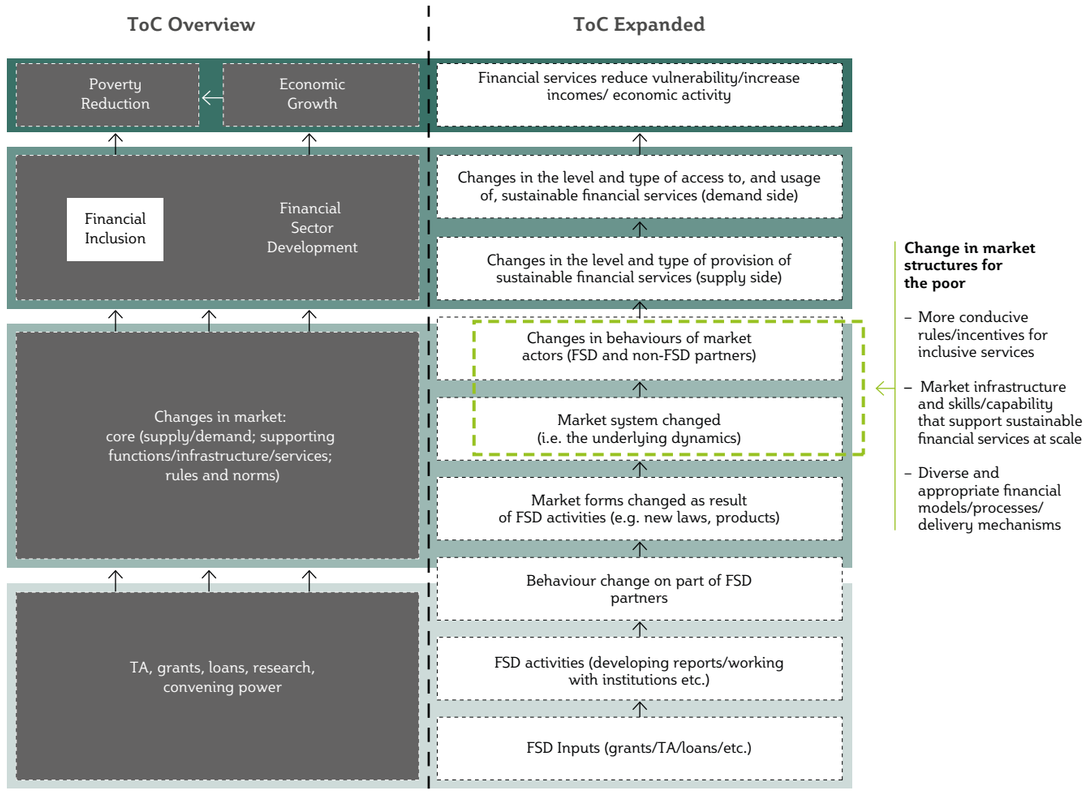
Systemic change refers to the broader transformation that takes place in the sector as a result of the programme. It is important for FSDs, at programme and theme level, to articulate how they anticipate their

27. Every development intervention is necessarily based on some kind of ToC, but the ToC is not always made explicit. If an FSD has a ToC that is only implicit, then it is useful to set it out explicitly, along with the assumptions regarding how parts of the causal process are expected to work, and how contextual factors may affect the programme – see Vogel (2012).

28. In particular, this will require FSDs to revisit their core assumptions regarding the binding constraints on the market working effectively for the poor, and to what extent they have removed/eased these constraints, and to what extent, by doing so, the market is producing better financial sector outcomes for an FSD's target group.

interventions will lead to changes through direct acts of facilitation, **but also how these interventions will fundamentally change the market in a way that enables even greater and more sustainable impact on users beyond the changes directly supported by the FSD and its partners.** Figure 6 provides a brief overview of how systemic change can be elaborated in the context of the generic ToC.

Figure 6 ToC and systemic change

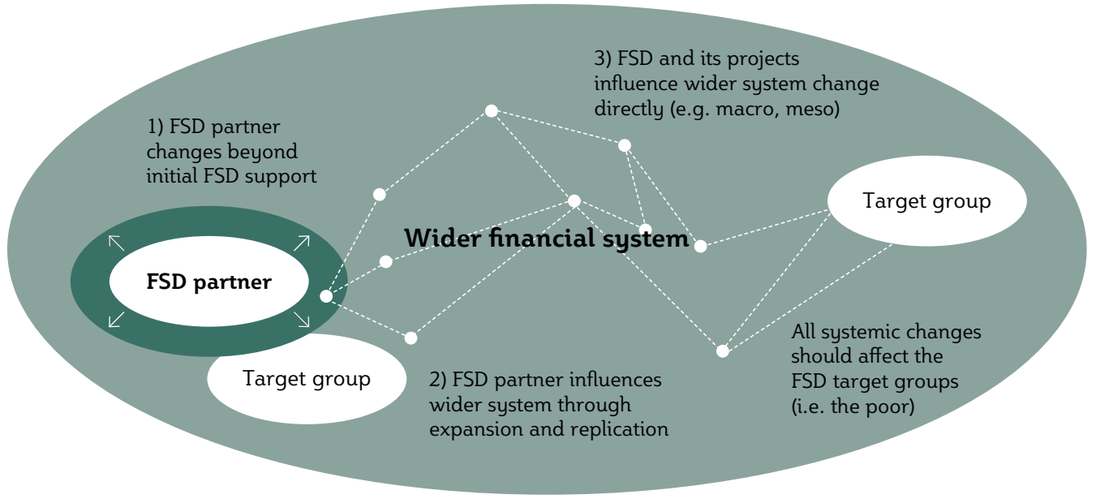


More specifically, Figure 7 **shows three main pathways by which FSDs can promote systemic change.** FSDs can consider including these pathways in their ToCs (and project results chains). These three pathways are:

1. the effects of the project on the partner and their immediate sphere of influence (for example through supporting a bank to move downmarket);
2. the effects that their partners have on the market system by demonstrating expansion and attracting replication (or championing reform with policy-makers) through their networks and others participating in the wider system; and
3. the effects that changing the structures of the system can have on the incentives for system actors. This may be triggered by FSD partner successes as well as directly by FSD-facilitated activities/projects, not just with the financial services providers but also policy-makers and other stakeholders (dialogue, research, advisory and advocacy activities).

Tip: An FSD can consider which pathways are relevant ex ante (i.e. during the investment decision phase), as well as monitor if expected pathways are working and/or others appear during implementation.

Figure 7 Pathways to systemic change



Source: Adapted from Osorio-Cortes and Jenal (2013

4.1.4 How to check that your ToC is evaluable

For a ToC to be evaluable it must be clear, relevant, plausible, testable, and contextualised, and it must take account of complexity.[29] Whilst these criteria can also be applied to individual results chains, they are likely to be too resource-intensive to be applied to all interventions, especially as many will not be subject to individual evaluation.

Clear: For a ToC to be clear, two elements must be considered:

- **Are the final impact, the financial sector outcomes and the outputs clearly identified?** FSDs should check that their impact, outcome and output statements are clearly defined. This means specifying, at each level, the change(s) that you want to see, who should benefit (or what beneficial change should happen) and by when.
- **Are the proposed steps towards achieving these clearly identified?** FSDs that achieve change rely on complex interactions and feedback loops. Key causal strands need to be unpicked to show how change is expected to occur. FSDs should, in particular ask:
  - Does your ToC explain how your programme outcome is expected to lead to poverty reduction and/or economic growth? The links discussed in the accompanying paper *The relationship between financial sector development, economic growth and poverty reduction* could be applied here.
  - Does your ToC explicitly incorporate systemic change (see below)? For example, does it capture both the direct and indirect effects of your interventions? Does it show how change in rules and norms or supporting functions could lead to changes in the interaction between suppliers and consumers of financial services?

Relevant: Are the programme objective(s) clearly relevant to the needs of the target group, as identified by any form of situation analysis, baseline study or other evidence (undertaken by FSD or others)? Is the intended target group clearly identified? FSDs should particularly consider:

- Could you use your market diagnostic to check the relevance of your programme objectives and if the intervention logic still holds?
- Do you have a focus on a particular group—for example, small and medium-sized enterprises (SMEs) or women? This will help you to define your target group.
- Have you explained why an M4P approach is more relevant to the target group than alternatives such as direct interventions/ service delivery?[30]

29. These criteria and general questions are adapted from Davies (2013). We have not included Davies' criteria of 'valid and reliable', because these are more about indicators, which we discuss in Stage 2, or 'consistent' and 'agreed', because these are more about organisational arrangements.

30. In their review of M4P evaluations, Ruffer and Wach (2013) find that no evaluation explicitly looked at this, but the assumption that the M4P approach is optimal should be tested.

Plausible: For a ToC to be plausible, two elements must be considered:

- **Is there a continuous causal chain connecting the FSD with impacts at the outcome or final impact levels?** Markets are complex and it can be hard to capture this complexity. There is a danger that a programme ToC will be too simplistic, or be based on past projects or an existing logical framework; it can also be too linear, with every step in the chain expected to follow automatically on from the next. To avoid this pitfall, FSDs should, in particular, consider:
  - Does the target group at the final impact level logically follow from the target group at the financial sector outcome level? For example, it is less plausible that poverty reduction could be achieved for all poor people in a country if the outcome is improved financial inclusion for only a small sub-set of people.
  - Can project or thematic results chains be 'nested' within the programme ToC? It will be difficult for FSDs to capture the entire causal chain, particularly all the potential links from inputs to outputs, in one diagram. However, assumptions about how the mechanisms might work between inputs and outputs are often unstated or simplistic, with limited consideration given to contextual factors or unintended consequences. This is why we are recommending that FSDs nest their project or thematic results chains within the programme ToC (see Box 7).
- **Is it likely that the programme objective will be achieved, given the planned interventions, within the programme lifespan? Is there evidence from elsewhere that it can be achieved?** FSDs should particularly consider:
  - To what extent existing evidence can be used to substantiate the causal links in the ToC. For example, a separate paper "The relationship between financial sector development, economic growth and poverty reduction" could be used to explain the links between financial sector outcomes and 'final impact.' This is a difficult area, one that is still being tested by global research, so the ToC should be clear regarding the extent to which evidence from other studies substantiates the FSD ToC in a particular context.
  - Are longer-term effects adequately captured? Apart from initial 'quick wins', FSDs are more likely to be aiming for changes that require time to take root, and in some cases for changes that are not solely related to their own interventions, but that are dependent on a variety of factors coming together.

Testable: Is it possible to identify which linkages in the causal chain will be most critical to the success of the programme, and which should thus be the focus of evaluation questions? Assumption testing is essential to the robustness of TBE. FSDs should particularly consider:

- What did your market diagnostic identify as the primary constraint on the programme objective being achieved? Testing to what extent this constraint was overcome is important. If it was overcome, then the impact on financial sector development of the removal or easing of the constraint should then be captured in the evaluation questions (see Step 2 for guidance on this).
- Are assumptions about the systemic nature of your interventions critical to the success of your programme? Have you defined exactly what you mean by systemic change, and how to measure it?

Contextualised: Have assumptions about the roles of other actors outside the programme been made explicit? A risk of taking a theory-based approach is the overstatement of causal contribution. For FSDs, which seek to facilitate change, it is important that the interaction of the programme with the context (i.e. the financial sector and the economy as a whole) is elaborated, to help ensure that impact is not overestimated (or underestimated). Understanding what others are doing and ensuring that FSD interventions are coordinated with such actions is part of a good market diagnosis. FSDs should explain their incremental strategic role (i.e. relative to others) in the narrative part of their programme ToC.

Complexity: Are there expected to be multiple interactions between different components, thus complicating the attribution of causes and identification of effects? How clearly are the expected interactions defined? FSDs are complex programmes that are working in complex contexts. At a minimum they should ask:

- Have you identified potential unintended effects? Assessing impact is about understanding the unintended as well as the intended consequences of action — particularly the negative unintended consequences. One way of doing this is by developing a 'negative' programme theory; for example, some households or enterprises may be negatively affected because others are benefitting from programme activities (displacement). When is this likely to occur? What exactly are the implications for households and enterprises, and how will this affect the achievement of the programme objective?
- Have you identified sequential dependencies in your ToC? For example, your market diagnosis may have identified that achieving a policy change will be important in regard to the effectiveness of other interventions (a classic example being the need for a suitable agent banking regulatory framework when promoting remote access banking technology).[31]

Tips for assessing the evaluability of the programme ToC include the following:

- A facilitated workshop may be an effective way of checking the evaluability of your programme/theme ToC. This can occur as part of the strategy design process. Working through the evaluability criteria above may reveal some weaknesses in the current version, or perhaps even the strategy that it underpins, and these weaknesses will need to be addressed. A facilitated workshop should help to build consensus on these points, and also confirm which ToC should be used as the framework for the evaluation.
- External perspectives, either obtained at the workshop or gathered separately, may be helpful for ensuring, in particular, that contextual factors are included, that large attribution gaps are not present, and that secondary evidence is marshalled. This also helps confirm if someone not closely involved with developing the ToC (e.g. a member of the FSD Programme Investment Committee (PIC)) can understand the core logic and underlying assumptions and deal with any 'self-importance bias'.
- Keep a record of the different versions of your programme ToC and of the reasons for the changes. This will be useful for evaluation questions such as:
  - To what extent has the programme been implemented as envisaged by the programme level ToC?
  - If an envisaged result was not achieved, was this due to a failure in the original theory or a failure in implementing the project or programme?

4.1.5 Project results chains

At any given time an FSD programme is made up of a cluster of projects. Each of these projects has its own project logic, and key evaluation questions. The logic for different projects and the evidence collected to confirm performance will make up the core evidence base for testing the overall programme ToC. All FSDs attempt to set out the intervention logic for a specific intervention – referred to as 'results chains' throughout the guidance. Results chains help to:

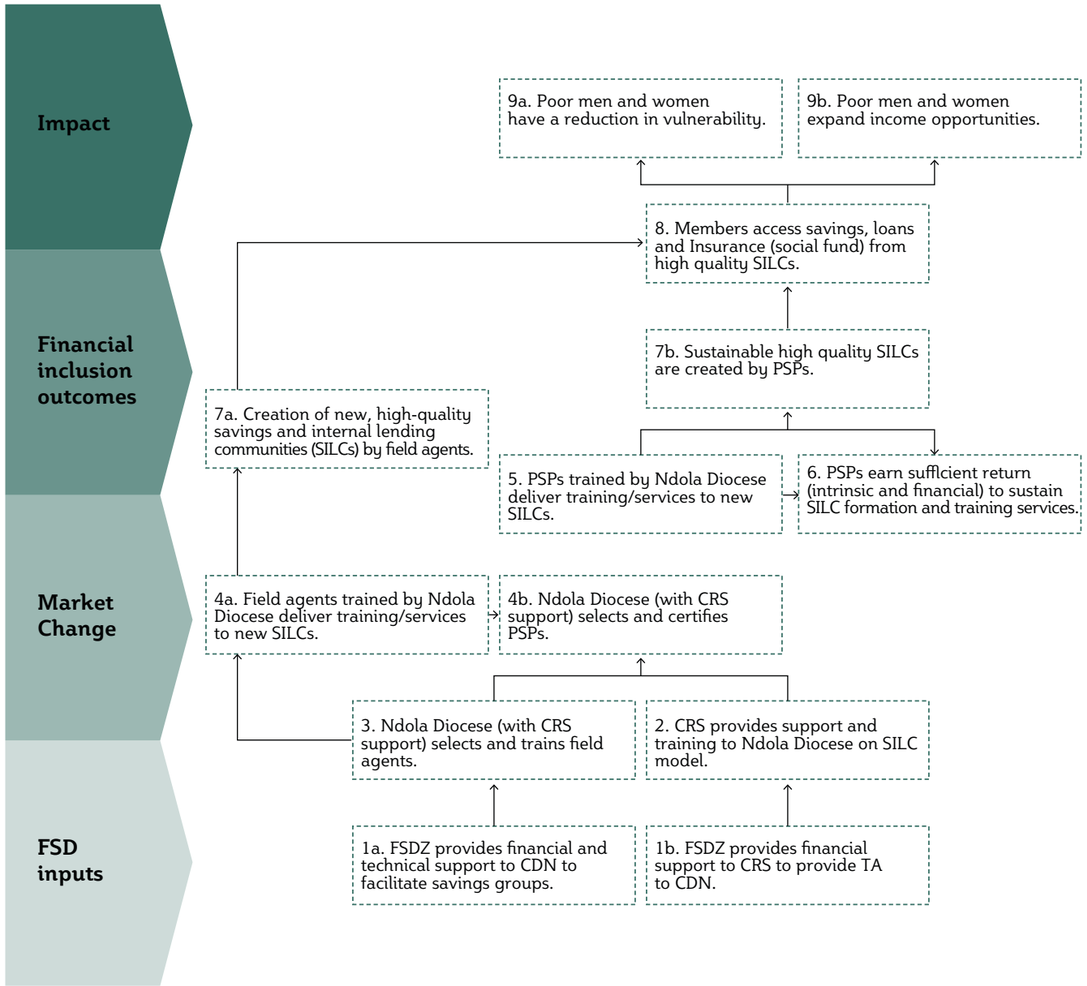
- set out the activities that the FSD will undertake, the outputs of these activities and the outcomes and eventual impacts expected;
- provide the basis for assessing if and to what extent changes are taking place; and
- provide the basis for assessing to what extent changes are due to programme activities.

Results chains, as illustrated by the example in Figures 8 and 9, provide an articulation of specific FSD interventions to help ensure coherence with the programme ToC, and ensure they are generating the appropriate evidence given the IOM objectives.

Discussion point: FSDs have noted the importance of in-depth analysis in measuring impacts along a results chains, not just the end impact. This breaks down the complex results chain into something easier to measure. Moreover, it allows them confidently to adjust their interventions if the results are not seen to be as intended.

31. FSDs that would like their ToC to be informed by recent thinking on applying complexity science to development interventions and FSDs should see the work undertaken by FSDK (Boulton and Johnson, 2013).

Figure 8 Example of a project's results chain



**PSP = private service provider, CRS = Catholic Relief Services, CDN = Catholic of Ndola Diocese*
Source: Example from a savings group intervention by FSDZ;

Whereas in a logframe the relation between the different levels of results will not always be stated, the purpose of a results chain is to make this causality of change explicit, clear and measurable. Indicators are then chosen to closely correspond to the different results specified, ideally with milestones (forecasts) of how the indicators are expected to change, helping the FSD team to think through and communicate this logic (see Step 3). The section below will show how fairly straightforward results chains can (and should) be adapted to help identify and measure systemic change effects. However, as Table 7 indicates, results chains are not a panacea, and there are a number of potential pitfalls to avoid when setting out results chains for measurement purposes. Further, the level of resources devoted to articulating and measuring the results chain, should be comparable with its importance to the FSD.

From both a measurement and a design standpoint. being clear as to the additionality that an FSD intervention is bringing to the market is also important, in terms both of articulating the change that the FSD wants to see and of being able to measure an FSD's contribution to broader market change (i.e. what would have happened to the market without the intervention) – see Box 6.

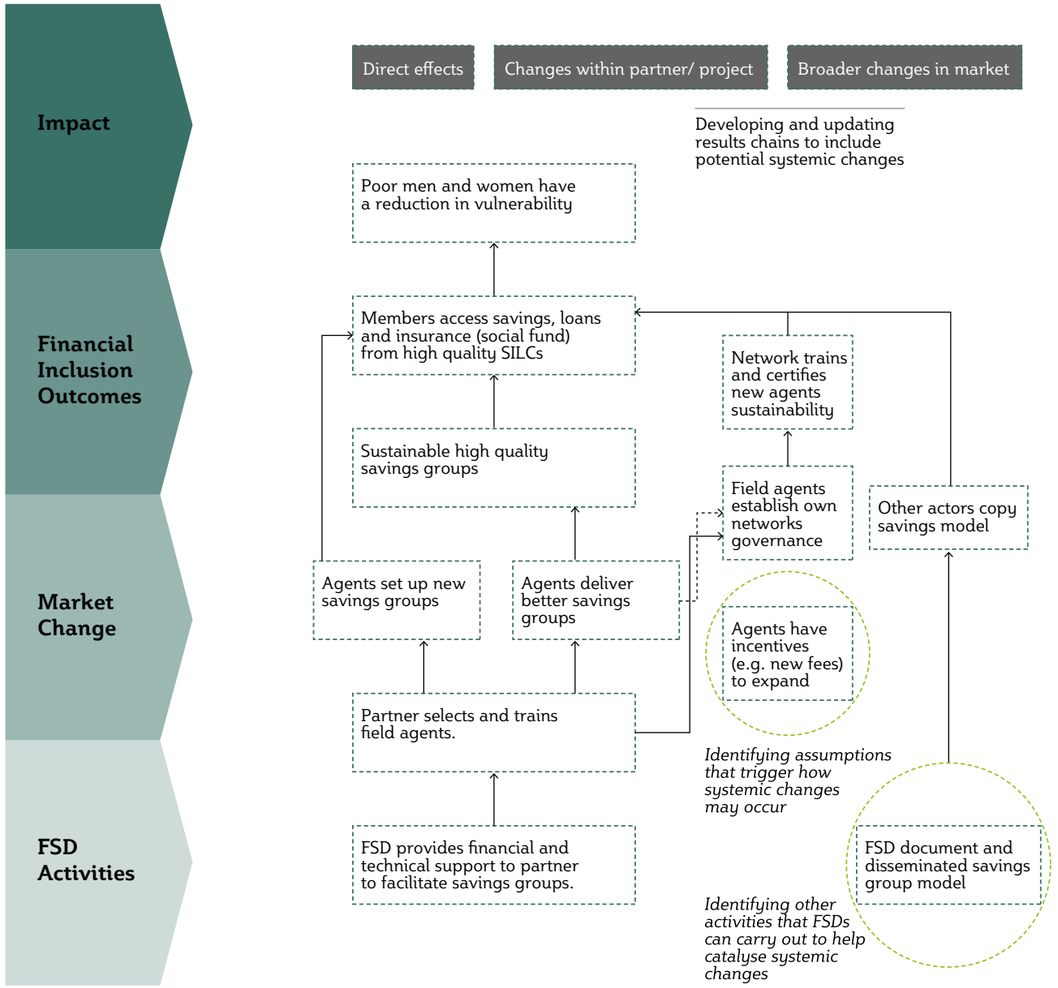
Table 7 Tips for setting out result chains

Challenge	Mitigation strategy (tips)
Project logic is not clearly set out	– Invest resources (e.g. a facilitated workshop) and seek the perspectives of others (e.g. government/private sector) when articulating/ confirming this logic – Ensure intervention logic at project level is consistent with the ToC at programme level – Avoid trying to capture too much information in one diagram (hence the use of nested approaches to results chains)
Numerous results chains become unwieldy	– Ensure that M&E staff and project managers work together to contribute to the design and updating of results chains – Prioritise flagship results chains: i.e. results chains that are more complex, account for a large part of the FSD impact and/or need substantial FSD resources
The project logic does not evolve	– Build in critical reflection points (e.g. quarterly and annual reviews) to test and update the ToC within the FSD team and with key FSD partners
The impact pathways are not detailed enough to really explain how outcomes/ impact will be achieved	– Indicators need to be mapped and tested against the results chain at each stage to assess if they remain relevant – Ensure that gaps between indicators at different levels, for example from outputs to intermediate outcomes, are not too large – Use available evidence globally to ascertain where specific gaps might exist with regard to how impact pathways might be working (and what data to gather) – Use additional studies to fill in gaps where required, especially in relation to understanding the causality between different links in the impact pathways
Measurement becomes too intervention-centric	– Use additional methods such as focus group discussions (FGDs) to obtain a reality check as a result of a range of perspectives. Compare with sector tracking results. (See Step 3 for further discussion.)

Box 6 Improving the contribution story: thinking about additionality in results chains

Additionality is typically considered as part of an FSD's investment decision and can help inform the project result's chain, by identifying what changes can result from the FSD intervention, and why these would not have occurred without the intervention (i.e. the FSD intervention was additional). The additionality analysis helps develop a robust contribution narrative of impact. Additionality can mean two things: first that the changes would not have occurred without the FSD intervention, or, secondly, that they would not have occurred as fast without the FSD intervention. Development finance institutions (DFIs), like the International Finance Corporation (IFC) and the Dutch Development Bank (FMO), use these criteria to identify evidence of additionality in areas such as: if the financing they are offering has a longer term tenor or is provided in a more appropriate currency than what is currently being offered in the market; if their finance is mobilising other resources; and if their finance is not crowding out other sources of capital.

Figure 9 Results chain with systemic change



4.1.6 'Nesting' ToCs and project results chains in the overall programme ToC

For simple projects or programmes it is possible for a single overall ToC to achieve a sufficient level of detail to be a useful basis for an impact evaluation. FSDs, however, have a large number, and different types, of projects and interventions that cannot be usefully represented in a single programme ToC. We therefore recommend a 'nesting' approach, which several FSDs have already adopted. This involves:

- An **overall programme ToC,** which summarises the overall logic of the FSD programme —the final impact that it aims to achieve and how broadly this is expected to take place. The programme ToC is used for developing high-level impact measurement questions and the overall narrative about the programme. The programme ToC can also provide a relatively simple snapshot, for FSDs to communicate their programmes to external stakeholders.
- Nested within this overall ToC, **thematic ToCs,** which show more detailed impact pathways for particular themes, such as digital financial services or SMEs or 'enabling environment'. It is these thematic ToCs that will be useful for developing more nuanced impact measurement questions, providing a sufficient level of detail to serve as a conceptual and empirical framework for measurement.

This ToC also provides an indication of whether a new project is consistent with the objectives (and understanding) of the programme.

Tip: Whilst thematic ToCs provide an extra depth to FSD analysis for measurement, care should be taken to ensure that synergies between themes across the FSD programme are articulated where possible.

- Nested within these ToCs, **project results chains,** which provide even more detail about the impact pathways for individual projects. The measurement of project-level results is not the focus of this guidance (though see the guidance in Section 4.1.5), but results at this level will contribute to the evidence base about programme impact, and are particularly important for generating evidence of how interventions promote changes in the market (see Step 3).

Advice about how to nest thematic ToCs and project results chains within an overall programme ToC is given in Box 7.

Box 7 How to nest project results and thematic results in the programme ToC

For FSDs, the programme ToC will be less detailed than projects results chains and will focus more on the overall programme intervention logic from outputs to final impact, showing how systemic constraints are reduced to achieve financial inclusion, financial sector development and ultimately poverty reduction and/or economic growth. Some FSDs will also have thematic results chains (e.g. agricultural services or market information), which fall between the programme ToC and project results chains.

It is important that thematic ToCs and individual project results chains are consistent with the programme ToC, so that project-level evidence can be used as part of the programme-level impact evaluation.

Each project will work at a different level/ distance from the programme impact. In practice, output/outcome/impact levels will have to be harmonised, i.e. even if a specific project's outcome is X, in the programme ToC that outcome may be called an output.

For macro/meso level interventions, the advice is to align outcomes in the project results chain with the output level of the programme ToC (i.e. that of market change). Project impact will then be at the financial sector outcome level of the programme ToC (i.e. changes in the behaviour of market actors as a result of the macro or meso level change in the market). For most micro level interventions, which may have a more direct path to financial sector outcomes, the advice is to align outputs in the project results chain with the output level of the programme ToC (although additional performance indicators will be needed to assess if activities have been carried out as planned, for example how many people were trained etc.). Project impact will then be at the final impact level of the programme ToC.

Box 8 Step 1 checklist

- Set out programme, thematic and project ToCs/ results chains. Ensure that these are nested and aligned.
- Have you developed results chains for at least your key flagship projects?
- Have you explicitly set out the assumptions that link your ToCs/results chains?
- Can the relevant staff (and not just the results advisor) explain the ToC and results chain(s)?
- Identify the additionality of the overall programme and of specific themes (both before and after the programme implementation).
- Test your ToCs with informed outsiders, but only those who have not been closely involved in the development of the intervention logic but understand the overall context.
- For the purposes of measuring the programme impact, the programme ToC and thematic ToCs should be tested to confirm that they are clear, relevant, plausible, testable, context specific, and take proper account of the complexities of the financial environment in which your FSD is operating.
- Has systemic change been considered in developing the intervention logic, at programme and project level?
- Do you have processes/ systems to periodically check if these ToCs still accurately reflect the FSD interventions and are still evaluable? Keep good records of these processes/ systems, especially in cases where you need to adjust any ToCs or results chains in the light of new evidence.

4.2 Developing impact measurement questions (Step 2)

4.2.1 Overview

- Once the Programme ToC is agreed, the second task is to develop impact measurement questions. These questions should ensure consistency between an FSD (including its partners where appropriate) and its funders as to mutually agreed measurement priorities, and should orientate the IOM system to collect appropriate evidence.
- There are many issues to consider when developing impact measurement questions, including which impacts to focus on, and how answering the questions will inform decision-making.
- Questions can be formulated to aid top-down and bottom-up measurement processes.
- Once you have developed your questions, there will then be an iterative process of reviewing and revising them with stakeholders, but also checking that they can be answered practically, with available resources and at a reasonable cost.

4.2.2 Impact measurement questions

Step 2 focuses on the development of questions for assessing the impact of the programme, not individual projects. Much of the advice, however, can also be applied to the development of project evaluation questions.

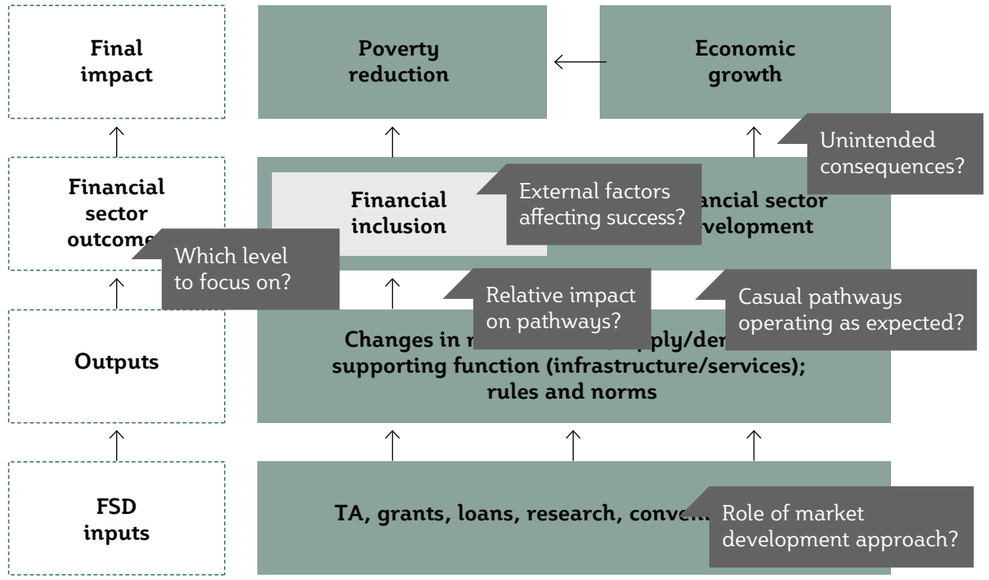
Identifying and prioritising the measurement questions is often a neglected part of the evaluation design process. These questions will be used to direct the focus of the evaluation of the programme's impact. **Measurement questions provide a crucial link between the FSD and its funders, helping to ensure a shared vision of what the FSD is trying to achieve,** as well as highlighting the practical challenges and opportunities from both an implementation and a measurement perspective. The questions will help to define the type of evidence the FSD wants to generate from the IOM system. These questions will focus the programme evaluation on specific aspects of the ToC, and will therefore determine the indicators that will be collected and additional analysis/studies that need to be conducted. The set of impact measurement questions should be addressed systematically during the periodic programme reviews and evaluations – see Step 7.

Tip: Taking time to think carefully about impact measurement questions will pay dividends and will help focus both measurement and research activities.

4.2.3 How to develop impact measurement questions

For this step we have mapped onto the stylised ToC some types of question that FSDs may want to consider asking about their impact (Figure 10). These questions are explained in further detail below.

Figure 10 FSD ToC and types of impact measurement questions



Questions should be developed in consultation with stakeholders and by considering the programme (and/ or theme) ToC, with a clear understanding of what an FSD and its funders hope to achieve by answering these questions. Using the simplified ToC above as the framework, the next sections focus on the questions posed in the small black boxes in Figure 10. FSDs should therefore consider:

- **At what level do you want to assess impact?** E.g. at the financial sector outcome level, are you interested in your impact on financial inclusion, or financial sector development more broadly? Are you interested in measuring impact at the final impact level for the whole programme, measuring it for a sub-set of (micro) projects, or will you rely only on secondary evidence at this level?
- **Are causal pathways operating as expected?** Answering this question would involve examining if the overall logic of the programme is operating as expected, and would form the core of most impact evaluations.
- **Which impact pathways are the most critical to success?** Which ones are the most risky and/or which are less likely to work? In their review of M4P evaluations, Ruffer and Wach (2013) found that the majority of evaluation questions were based on a ToC in which most questions did not adequately test the linkages contained in the theory. Effective TBE means you will need to ask (sometimes tough) questions about these linkages.
- **Are you interested in the relative impact of different types of interventions?** For example, are you interested in understanding whether your work with policy-makers or your TA to financial service providers was a more significant driver in unblocking a market constraint?
- Other key points have been discussed in Section 4.1 above, on Step 1. These include whether an FSD's programme is being implemented as envisaged in the ToC, capturing cross-cutting themes, identifying unintended consequences and the risks of a market development approach relative to one based on alternative options (i.e. using a counterfactual).

Stakeholder objectives for the impact evaluation should also be taken into account when developing questions. This could be accomplished by involving stakeholders in the discussion of the programme theory (perhaps through a facilitated workshop) or by asking them to comment on a set of draft questions. It will be particularly important to take funders' interests into account.

Tips for developing impact measurement questions

- **Formulate your questions as questions.** This may seem obvious, but so-called evaluation questions are sometimes just a set of issues to explore. Having an actual question helps define the focus of the evaluation and prompts you to think about how to collect and analyse data to answer it.
- **Be aware of language.** '*Did* the programme make a difference?' indicates that a yes or no answer is required and implies that quantitative methods will predominate. In contrast, '*How and why* has the programme made a difference?' indicates that the answer will be an explanation and implies that qualitative methods will predominate.

  At the programme level, FSDs are not going to be able to answer an impact evaluation question phrased as follows: 'What was the net impact of the programme?' – because this requires a statistical comparison group, which, as discussed in earlier Chapter 3, will not be feasible for assessing programme-level impact.

  If you are interested in the differential impact of the programme, you may also want to consider asking 'for whom' and 'under what circumstances'.
- **Define the cause and effect that is of interest in your question.** In general, an impact evaluation question will try to relate the programme, or an element of the programme, to an effect. But as discussed below, the IOM recommends the use of both bottom-up and top-down focused questions.
- **Check that the questions are clear and concise.** It can help to develop sub-questions under your primary evaluation questions and, if the question is becoming too lengthy, to define key terms separately.
- **Prioritise and potentially eliminate questions** by assessing: a) stakeholder interest in the questions (especially FSD funders); b) their potential for providing new information that will influence the decision-making of stakeholders; and c) time, budget and skills implications.

4.2.4 Example of impact measurement questions

Figure 11 gives some indicative examples of impact measurement questions for FSDs, separated into top-down and bottom-up approaches: answers to both are required to provide a comprehensive narrative as to whether the ToC has held up in the light of the evidence. FSDs will, of course, need to adapt such questions to their own ToC and stakeholder priorities. Table **8** below highlights potential misconceptions regarding types of bottom-up/top-down questions, based on consultations with FSDs.

These figures and tables focus on measuring the cause and effect of relationships in the financial sector and FSDs' contributions to them. However, there may be additional questions of the kind that are normally considered in traditional programme evaluations and that are not focused on impact (e.g. questions that focus on issues such as programme relevance, efficiency and broader learning). Nonetheless, the IOM guidance will provide FSDs with significant evidence to help answer such questions. For example:

- Which market segments have performed relatively better in terms of the FSD's objectives?
- To what extent has the programme been implemented as envisaged by the programme-level ToC?
- How are market constraints being identified and how this is reflected in FSD actions?
- Are there any significant linkages between projects, and what (if any) impact do they have?
- Is the programme delivering VfM?

Figure 11 Examples of impact measurement questions

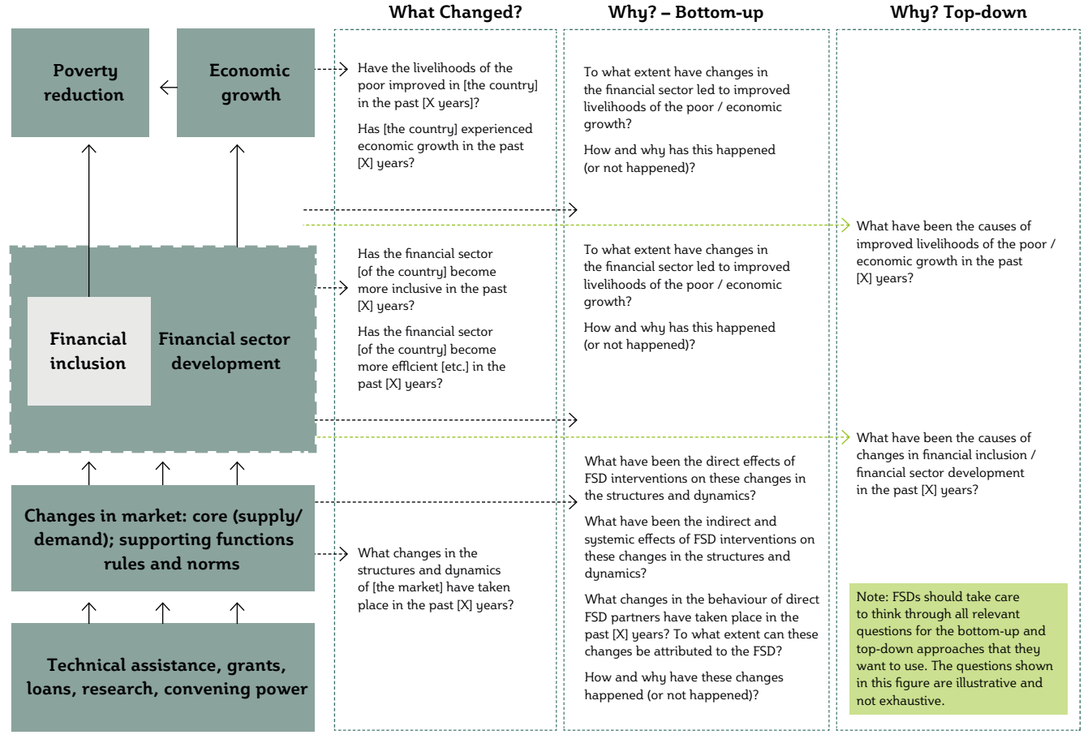


Table 8 Programme-level top-down and bottom-up impact measurement questions – good and bad examples

	Good examples	Bad examples *(with explanation)*	Ambiguous *(with explanation)*
Bottom-up	– What has been the contribution of the FSD programme to the observed changes in market competitiveness and efficiency? – How have FSDs contributed to enabling policy changes that promote the growth of mobile money services for the poor?	– To what extent will gaining access to financial services improve livelihoods? *(More of a top-down question, as it is focused on how livelihoods have been influenced by (all the) changes in the financial sector, with no focus on the FSD programme)*	– What has been the contribution of changes in the structures and dynamics of the market to changes in financial sector development *(Could be both bottom-up and top-down; market changes are often measured in terms of how FSD programmes have caused them, although this more open question would provide the evaluator with the opportunity to explore many non-FSD contributions)*
Top-down	– To what extent has growth been driven by finance? – What have been the main causes of changes in financial inclusion in the past [10] years?	– Is FSD programme a key driver of financial sector development in the past [x] years? *(This is more of a bottom-up question focusing on how an FSD programme has driven change in the sector)*	– What changes can be tracked to an increase of mobile money? Is there any contribution from FSD interventions? *(The first part of this question could be top-down but the second part is bottom-up, focusing on how an FSD programme has driven change. It would be better to split this into separate questions)*

Source: Adapted from March 2015 FSDA workshop exercise

An example of the application of this approach is provided in Figure 12 and Table 9 below, which set out impact measurement questions for an indicative thematic ToC of an FSD programme that is working on the enabling environment.

Figure 12 Indicative example top-down and bottom-up impact measurement questions

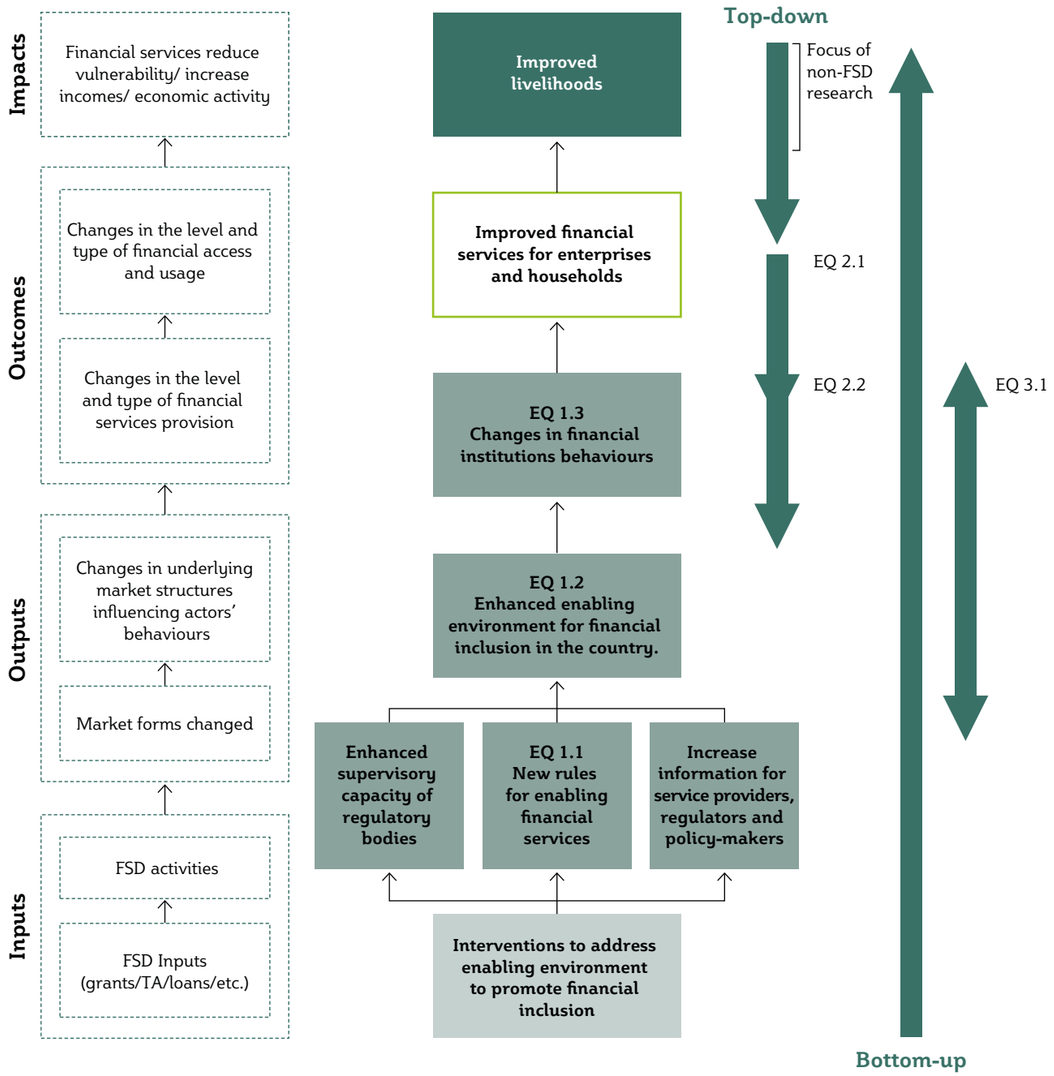


Table 9 Bottom-up and top-down impact measurement questions and methods (indicative example)

Type of evaluative approach	Impact measurement questions	Indicative measurement methods (See Steps 3–5)
Bottom-up: From intervention to market system change (and outcomes)[32]	EQ 1.1 What have been the direct outputs from FSD interventions? (or signs of progress, such as behaviour changes, towards these outputs)	– Intervention monitoring data – Key informant interviews and FGDs
	EQ1.2 What changes have occurred in the enabling environment for financial inclusion?	– Intervention monitoring data – Enabling environment surveys
	EQ1.2 To what extent can these changes be attributed to FSD?	– Key informant interviews and FGDs – Case studies
	EQ1.2 What other causal factors have there been?	– Political economy analysis
	EQ1.3 To what extent have these enabling market changes led to improvements in the inclusiveness of the system? (i.e. through changing institutions' behaviours)	– Key informant interviews and FGDs – Institutions' published information – Supply-side studies
Top-down: Changes in financial outcomes (behaviour of financial institutions and level and appropriateness of financial services)	EQ 2.1. How have the level and composition of financial services changed (in the relevant sector, given the enabling environment changes)?	– Demand-side studies – Supply-side studies
	EQ 2.2 What market system changes led to the changes in financial inclusion? EQ 2.2 What have been the causes of any changes in the enabling environment?	– Key informant interviews and FGDs – Supply-side studies – Case study approaches – Outcome harvesting – Most significant change
Credible narrative: Linking market changes (and outcomes of those changes) to the FSD's interventions	EQ 3.1 To what extent can the market-level changes observed in top-down answers be attributed to FSD interventions?	– Findings from Steps 1 and 2 – Synthesis and triangulation – Contribution analysis

There is no set rule regarding when the top-down and bottom-up measurement questions should meet (see footnote 33), and where synthesis measurement questions are most useful. **The choice of questions will depend on what is sensible for that ToC (or individual result chain), as well as on what the limits are in terms of gathering evidence for bottom-up and top-down assessment.** Even for micro interventions, where one can obtain good bottom-up data all the way to the financial inclusion level, FSDs still need to assess what market system drivers are supporting that inclusion. This provides a way of checking that it is not just the FSD project that is having a direct impact, which may be unsustainable in the long term, but that underlying market drivers have also changed.

Discussion point: During the second IOM consultation workshop (March 2015) funders encouraged FSDs to think about the IOM at the start of the programme/strategy period. They stressed the difficulty in retrospectively evaluating existing programmes, where there is no budget, no baseline, no ex-ante evaluation objectives, and limited data, with it being particularly hard to demand this from private sector actors after a project has finished.

32. It is more likely that direct outcomes from projects, in terms of provision and usage of financial services, will be captured for micro-level projects rather than for those operating at macro level, given the results chains are longer and more diffuse (suggesting that bottom-up measurement may go further than this illustrative example for some micro projects). However, even for micro projects, whilst measuring those directly impacted (for example through investing in savings groups) it is still important to assess how the underlying dynamics of informal finance markets are changing (hence top-down measurement will still be relevant, in order to develop a comprehensive picture).

Once you have identified and prioritised your impact measurement questions, using an evaluable programme ToC as your framework, you are ready to start thinking about how to measure your impact. **The evaluation matrix template shown in Table 10 below will be a useful tool for summarising your evaluation questions and how you are planning to answer them.** An FSD programme should be able to complete the first two columns even before the strategy implementation starts. The remaining columns will be completed using the guidance in Chapter 5 of this paper (which deals with Stage 2).

Table 10 An evaluation matrix template (with indicative example)

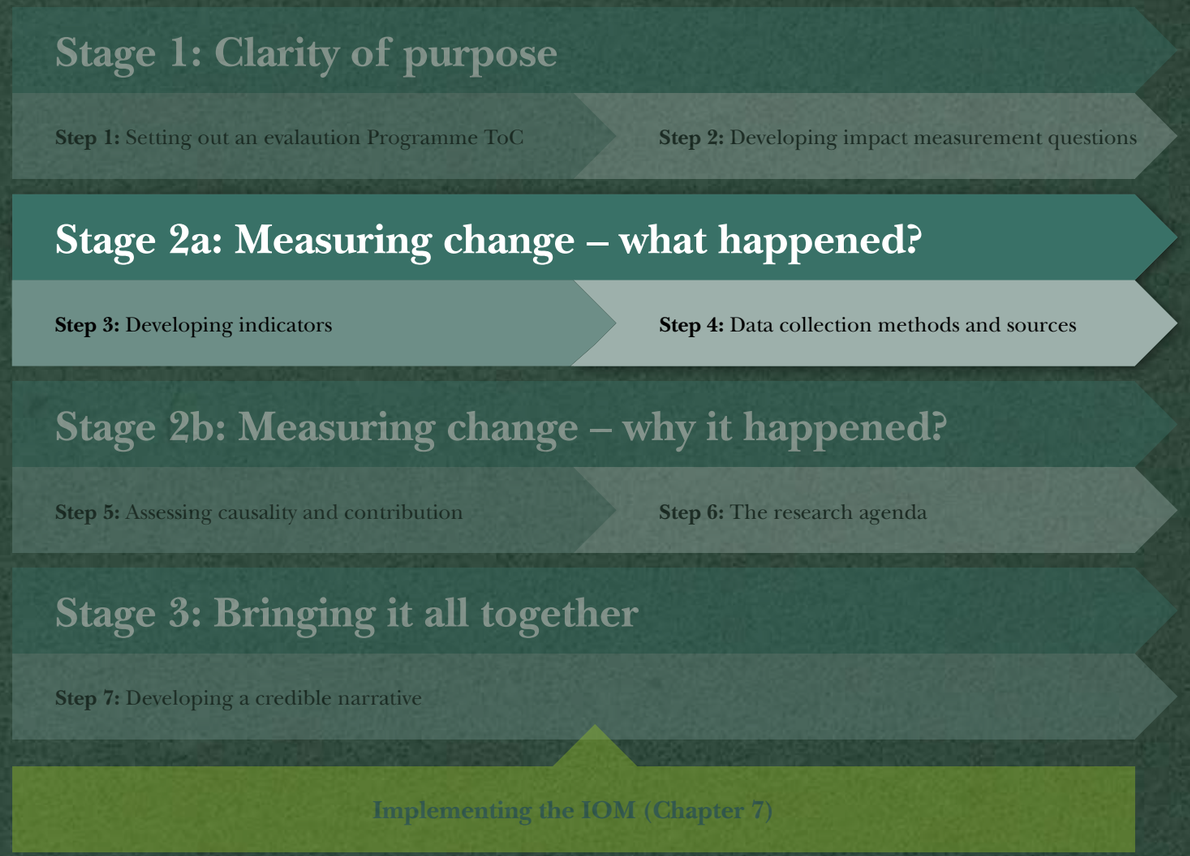
Evaluation question	Sub-question	Indicative indicators	Data collection methods	Data analysis methods	Evaluability issues
Bottom-up: To what extent has the FSD contributed to an improved enabling environment?	How have FSD projects affected policy-making?	– Number of organisations demonstrating improved effectiveness in advocating for financial sector reform – Number of policy-making bodies engaged by FSD with improved capacity to formulate and implement effective financial sector policies and regulations – Attitudes to financial inclusion changing amongst policy-makers	– Intervention-based monitoring data – Capacity index[33] – Key informant interviews – Case studies – Supply-side studies	– Compilation of interview and questionnaire responses – Analysis of case studies – Cross-checking of supply-side findings with primary data collection	– Key informant interviews conducted with only five organisations, limiting the extent to which information can be triangulated for all
	[Based on the example above, additional sub-questions and their respective indicators and data methods would be included here by the FSD team. Additional examples of indicators are provided in Section 5.1 and Annex D]				

Box 9 Step 2 checklist

- Have you identified which impact pathways are likely to be the most critical to success, and discussed these with the funders?
- Decide on the level at which you want to evaluate impact (i.e. financial inclusion, financial sector development, economic growth, or livelihoods), and the extent of resources you are willing to devote.
- Have you considered other important criteria (e.g. changes in the wider market, understanding the market development modality you are using or cross-cutting themes such as a gender, geography or youth) that the FSD programme and its funders are interested in? However, be pragmatic regarding how many core questions your IOM can focus on.
- Formulate top-down and bottom-up questions for programme and thematic ToCs (and subsequently for flagship projects). Will these act as useful and feasible measurement questions?
- Do these key questions need to be agreed with funders?
- Decide on how far down top-down questions should go, and how far up bottom-up ones should go (although you should note that this may change as the quality of the evidence base for each causal pathway becomes clearer).
- Determine which questions will require what types of evidence – from monitoring data (see Steps 3 and 4) to types of causality analysis (see Step 5 and 6). Be explicit about which questions are primarily useful for measuring impact, as opposed to those contributing to broader learning.
- Develop and start to fill in your evaluation matrix (see Steps 3–5 to help develop this).

33. For example, a questionnaire that is able to judge if capacity in an institution is improving over time.

IOM – Chapter 5: Measuring Change – what happened (Stage 2a)



Chapter 5, Measuring change, covers Stage 2 of the process of implementing the IOM guidance: This stage is split into two sub Stages –2a, and 2b.

This section focuses on Stage 2a, providing guidance to FSDs on assessing what change has happened. This includes changes that are directly related to the FSD programme and projects, as well as changes in the financial system more generally.

Stage 2a is broken into two steps:

Step 3 – Developing indicators: This outlines a range of indicators FSDs can use to monitor the results of their programme.

Step 4 – Data collection methods and sources: This provides the corresponding data sources, and methods for collecting the information related to these indicators.

5.1 Developing indicators (Step 3)

5.1.1 Overview

1. Impact-oriented indicators are needed to monitor the overall programme ToC and the performance of specific projects. Such indicators will help both to improve programme decisions and to build an evidence base to inform impact assessments.
2. For FSD programmes, this requires moving beyond monitoring direct outputs and outcomes of projects, and augmenting 'traditional' monitoring in a number of ways. We have identified four key issues that FSDs should consider when developing or reviewing their indicators:

   a) **Progress indicators:** Because the clear, discrete outputs and outcomes of interest for the FSDs may take some time to materialise, it is important to also measure intermediate effects, i.e. the steps between these discrete changes. Such indicators can include changes in behaviour on the part of market players and policy-makers, as well as other shifts in the market.
   b) **Market system development:** Monitoring the 'systemic' changes stemming from FSD interventions.
   c) **Combining top-down and bottom-up indicators, including sector tracking:** Tracking how a complex and dynamic market is changing systemically beyond anything resulting from specific FSD projects through top-down monitoring /sector tracking and combining (and triangulating) that with bottom-up monitoring, thereby getting a richer picture of what is happening in a financial sector, and why.
   d) **Monitoring beyond indicators:** FSDs may also need to capture evidence that does not fit into regular monitoring; for example, stakeholder perceptions and views on particular (and unexpected) events, processes and outcomes in the financial sector.
3. **In practice,** once indicators are identified, FSDs need to consider specific issues regarding their use, including:

   i. how to set baselines, given that FSDs manage dynamic programmes;
   ii. how to update indicators based on changing information and changing priorities; and
   iii. how to incorporate impact-oriented indicators into an M&E system, traditionally built around a linear and usually static logframe and largely used as a tool to enforce accountability to funders.

5.1.1.1 Extending current monitoring systems

FSDs need good indicators to monitor the progress of their programme and projects against the ToC and the results chains. Indicators are also important for monitoring assumptions and risks, and for building an evidence base that will inform evaluation, facilitate cross-programme learning and help programmes to adapt current and future investments. It is good practice, already applied by many FSDs, to set out individual intervention logics using a results chain (linked to their overall ToC), and then set corresponding indicators to confirm whether various changes have occurred (see Step 1). These results chains should be complemented by a measurement plan clearly identifying the indicators, how they are defined, how they will be measured and what are the expected changes (targets/milestones). FSDs then aggregate and synthesise the indicators into an overall results matrix – most commonly a logframe.[34] Programme indicators for various FSDs, as recorded in these logframes, are available in a separate paper,[35] although commentary on the similarities and differences amongst logframe indicators used by FSDs is included in Annex C.

Table 11 Typology of indicators for an FSD IOM system – focus of this guidance paper

Types of indicators/ frameworks	Focus of guidance paper
Monitoring direct performance of projects	Limited (focus on building evidence of direct (and immediate) project impact)
Logframes	Limited (the logframe is considered as a sub-set of the evidence required to test a programme's ToC)
Progress indicators	√
FSD facilitated systemic change indicators	√
Sector tracking (and combining with bottom-up)	√
Beyond indicator monitoring	√

34. FSDK synthesise a range of indicators into an annual 'impact synthesis' framework, which is structured similarly to the logframe but is more comprehensive in terms of detail.

35. See FSDA website www.fsdafrica.org/knowledge-hub.

Discussion point: Funders, including DFID, stressed that there is more flexibility in setting indictors (including in logframes), than often thought by implementers. FSDs need to invest in establishing open partnerships with their funders to agree a coherent set of indicators that track progress towards programme objectives, facilitate learning and risk taking as well as ensure accountability. The indicators can be changed in consultation with funders and could distinguish between those used for accountability and others used for learning about broader market change.

This section does not seek to provide a 'best practice' set of indicators for individual projects or programme logframes. Each market and FSD programme operates in different ways, and will have to define indicators accordingly. FSDs will need to agree programme level indicators with their funders. **This guidance paper focuses on extending existing FSD processes to develop a more comprehensive set of impact- oriented indicators, most notably trying to improve the ways FSD measure systemic change** (Table 11). This will provide FSDs with a typologies of indicators/ themes which can influence FSDs' discussions about measuring project results chains, and can help to finalise programme reporting (e.g. logframes, annual reviews, programme evaluations).

5.1.1.2 Moving towards impact-oriented indicators

From an IOM perspective, the main gap in current monitoring is the limited evidence of results relating to the relationship between the direct outputs of the FSD interventions and the larger market change that the FSDs seek to effect[36] (see Figure 13), i.e. the development of inclusive, pro-poor financial markets and other changes in the underlying structures and dynamics of these markets.

Closing the evidence gap between FSD programme outputs and its final outcome of an improved market for financial inclusion presents a number of challenges, including:

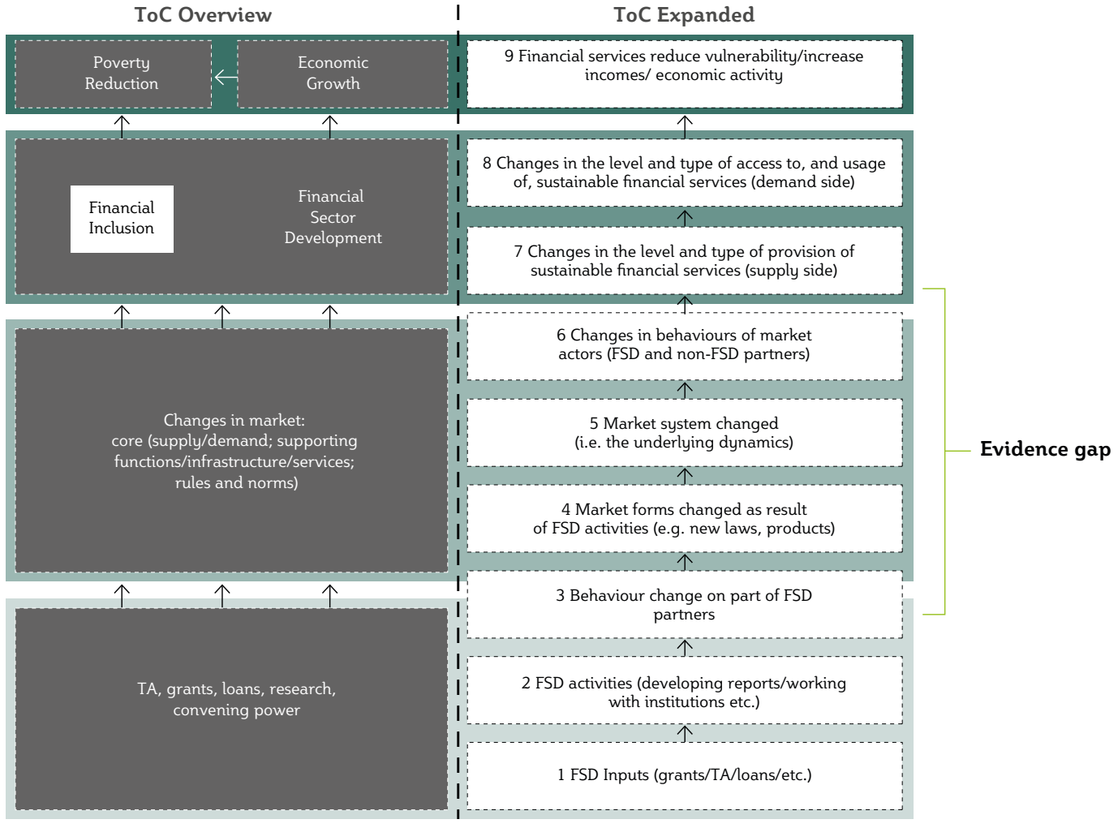
a) The progress indicators/ **intermediate steps** are often difficult to define and measure, especially with traditional quantitative indicators. These intermediate changes may include the easing of market constraints or changes in attitudes or behaviour. FSDs may informally monitor the progress being made, but this evidence is not systemically analysed and documented, and therefore is not easily available. These indicators also tend not to be captured in funders' results frameworks, as they are building blocks towards achieving programme objectives. But given that the time-scale of market change is unpredictable, FSDs need to capture indicators that show the potential for changes to come in the medium to long term.
b) **FSD facilitated systemic change** (i.e. going beyond changing specific forms of the market in the short term to assessing the results of these changes and longer term dynamics) can be difficult to define and measure, especially when using only traditional quantitative indicators. In addition, systemic change results can often be indirect, less attributable, and long term, which leads to the risk that systemic change indicators are deprioritised in place of indicators that focus on programme accountability. For FSDs, these types of market system development indicators are critical for measuring the true value of their interventions.
c) To collect appropriate evidence for a contribution analysis or impact evaluation, it is important to capture both **top-down (a financial sector focus) and bottom-up** measurements. Programmes are often hesitant about focusing on and tracking sector-level indicators (top-down) because they go beyond the direct results that can be attributed to the programme. However, complementing bottom-up indicators with sector-level change indicators is an important component of IOM, moving beyond an exclusive FSD focus and embedding a sectoral perspective in relation to systemic change. In addition to FinScope, sector-level tracking would focus on the size, access, diversity, efficiency, and stability of the financial sector, through annual collection of data from existing sources.
d) Some aspects of the ToC cannot be measured sufficiently well with indicators. By focusing only on simple quantitative measurements, FSDs are not able truly to capture the progress that is being made. As explained earlier, concepts like systemic change may require **measurements beyond traditional indicators,** including qualitative indicators and narrative description, particular where important unanticipated events have occurred.

36. Depending on the design of the ToC, this evidence gap may be classified either as an 'output' or as an 'outcome'. The essential challenge lies in measuring the steps/changes that occur between the FSD outputs and the outcome of improved financial access – for example, at market level.

Figure 13 below is an extract of Figure 4 and illustrates the need to ensure that the ToC is unpacked. Changes in boxes 3, 4, 5 and 6 are carefully tracked to help assess whether expected changes in box 7 and 8 are likely to occur in the next three to five years.

Tip: FSD programmes should focus on whether the chosen indicators for boxes 3,4,5 and 6 (within Figure 13) will help track progress towards boxes 7 and 8. They should not worry too much about whether these are included in the logframe and whether these are labelled as output or outcome indicators. Progress indicators of this type also provide FSDs with a more realistic approach to reporting results from what can be lengthy acts of facilitation, which may not result in concrete market changes in the short term.

Figure 13 Evidence gap and going beyond traditional monitoring



The rest of this section elaborates the IOM approach and illustrates how IOM can help address key issues identified above:

1. ensuring that there are sufficient indicators that measure progress (the intermediate steps) towards the medium and longer-term outputs or outcomes;
2. capturing systemic changes; and
3. capturing indicators of financial sector change beyond FSD interventions (i.e. top-down sector indicators). Some changes in the market are less amenable to specific quantitative indicators but are still important events that can be tracked using qualitative methods and what we call 'narrative reporting' and 'monitoring beyond indicators'.

5.1.2 Indicators of progress

Defining and measuring indicators of intermediate steps between discrete outputs or outcomes (as currently recorded in the logframes) is important. These indicators of progress capture the intermediate steps that are essential to achieving the financial inclusion outcome. Depending on the specific FSD ToC, three broad areas that an FSD may want to measure include: changes in the attitudes or behaviour of FSD partners; easing of market constraints; and process-based indicators of progress made by FSD partners.

5.1.2.1 Attitudes, knowledge, behaviour, skills

The FSD partner (financial services providers (FSPs), policy-makers, meso-level service providers) will go through a series of steps before market change actually happens. Indicators can be established to track the steps that the market or market players (including the FSD partner) are likely to take to improve their capacity and steps that contribute to change in the market, though not yet at a system-wide level. FSDs want to identify a **change in the attitude, knowledge, behaviour and skills** of the partners that FSDs are trying to influence, be it a private sector provider, a non-governmental organisation (NGO), a regulator or policy-maker. Indicators to measure such changes will be qualitative and will rely on partners being open about what is happening in their organisations. The data source may be a conversation/interview with a senior regulator or financial institution executive confirming that a particular change is underway, or, possibly a survey or FGD among FSPs. Proxy indicators can also be useful here. See Table 12 below on how to define progress indicators, as changes in attitude, knowledge, behaviour and skills are certainly key steps towards desired results. Also, see the Section 5.1.5 below on qualitative indicators and narrative reporting for additional ideas about how to measure these types of changes.

Box 10: Progress indicators

Progress indicators are akin to marking an ascent up a high mountain by establishing camps at different altitudes. No one is capable of climbing Everest or Kilimanjaro without breaks along the way to rest and take stock of the journey so far. Thus, progress indicators are just what they say they are: a way of taking stock along a route that leads to the summit, i.e. in the case of IOM for most FSDs, the outcome of greater financial inclusion. To achieve this outcome FSDs often need to engender and support new thinking, attitudes and behaviour changes, possibly with further re-thinking and attitude changes stimulated by evidence emerging from previous actions.

One illustration of this process involving progress indicators could be the development of new, less strict know your customer (KYC) rules for small-sized mobile money transactions. This would not be an end in itself. Rather, succeeding in ensuring there are lighter KYC requirements would be one key step (or 'camp') along the route to achieving greater access for poorer people by reducing the friction for small transactions. First, evidence from mobile network operators (MNOs) and banks might need to be gathered to show how mobile money transactions were not increasing as fast as expected and the reasons for this, along with comparable evidence from other countries. Then the regulator would have to accept that evidence and arguments, and change its thinking about the issue. This may be followed by new, less onerous KYC requirements. This in turn should encourage MNOs and banks to increase access and their customer bases, and change their behaviour in relation to such target groups.

However, it should be noted that not all results chains are going to be as linear as this example. Some progress and associated indicators might be achieved, but then other factors, mostly outside of an FSD's control, might come to bear on the overall process. So progress towards financial inclusion is made or may get stalled, or a different path towards the summit may need to be taken. Progress indicators help to identify key intermediate outcomes and track progress towards these.

Table 12 Example of indicators of progress for a micro-insurance project

	Progress indicators for the FSD partner	Progress indicator for the wider market
Expected result	Micro-insurance provider (partner) improves its capacity to serve low-income people	Product for the micro-insurance sector targeting poor people is established
Indicator/ change of interest	1. Partner attends a training session that is specific to the topic 2. Partner conducts a diagnosis of the sector 3. Partner conducts market research/ segmentation on low-income households 4. Partner designs suitable micro-insurance product 5. Partner take out a licence to provide micro-insurance products 6. Training budget spent on micro-insurance 7. Partner sets up new dept. to focus on micro-insurance 8. Strategy is developed/ adopted by partner for low-income households 9. Number of staff (in partner organisation) trained and certified in pro-poor product development 10. Board of FSD partner approves strategies/ business plan with micro-insurance elements	a. Partner designs suitable micro-insurance product b. Partner takes out a licence to provide micro-insurance products c. New micro-insurance providers registered d. Non-FSD-supported partners recruiting staff for micro-insurance team e. New product has reached market (increase in number of low-income households/ clients reached/ served; increase in number of micro-insurance policies)

5.1.2.2 Easing market constraints

Given that actual market change is often a long-term process of easing identified market constraints (for example, those constraints that are identified in the initial FSD market diagnosis), an intermediate indicator can help identify progress towards the expected market change. Some examples have been provided in the table below.

Tip: It is useful for an FSD to periodically revisit its original market diagnosis, which informs its ToC, to assess if there are signs of 'progress' that show that the important constraints in the market are changing. This may go beyond only identifying constraints that a specific project is working on. It may also happen that the removal of one market constraint reveals the existence of another constraint, not previously identified or clearly seen.

5.1.3 Capturing systemic change

5.1.3.1 Overview

As discussed in Chapter 3, systemic change moves the focus of a monitoring system away from direct acts of facilitation or direct numbers and seeks to understand the broader transformations taking place in the sector. It is difficult to define systemic change precisely, but characteristics used by other market development programmes appear to be well suited to FSD objectives:[37] trying to promote sustainability, scale-in that acts of facilitation benefit an increasing proportion of the FSD's target group(s), and resilience (ability to respond to shocks and adapt models/processes to changes in the market).

5.1.3.2 Measuring systemic change for FSDs

Measuring systemic change for FSDs Figure 13 shows that the core of the measurement process relates to a direct results pathway, from an FSD intervention creating some concrete change in the market (output) and leading to improved outreach of financial sector (outcome). For some FSDs, there are direct indicators that already capture such market system development in their logframes, offering evidence that they will promote sustainability, scale and resilience; but for others, either because they have not explicitly noted this aspect in their ToC or because the market system developments are for some reason not considered to be the main outcomes of interest or are too difficult to measure, there can often be an evidence gap as regards a particular act of market change and how this has improved the underlying dynamics of the system.

Discussion point: During the consultation there was discussion as to whether systemic change always happened at the national level, or if it could also be more localised. The IOM accepts both, as long as such change reflect changes to the underlying structures. While localized change may have less scale in terms of absolute numbers, it may have changed a low base (for example, marginalised farmers) significantly, and therefore in relative/ percentage terms it can viewed as having scale.

37. See DCED (2014b).

Table 13 Easing of market constraints

Level	Examples of constraints	Examples of indicators of progress towards constraints being eased
Macro	1. Restrictive KYC requirements 2. Lack of regulation for micro-insurance	a. Regulator agrees to new or amendment regulation that addresses the constraint identified b. Central bank and/ or ministry of finance, plus private sector FSPs, agree to and then participate in developing a national financial inclusion strategy which includes reviewing KYC requirements and micro-insurance c. Central bank and/ or ministry of finance form a task force to review relevant issues
Meso	3. Lack of reliable credit history information in the market	d. A study on creating a credit information bureau is conducted e. The central bank recommends a new law or issues directives on establishing a credit information bureau f. Number of unique clients added to the credit bureau g. Number of FSPs that report to the credit bureau
Micro	4. Market actors do not have the skills to leverage new delivery mechanisms 5. Market actors have not prioritised reaching out to new market segments	h. Number of market actors that receive TA in relation to developing products for the new delivery mechanisms i. Senior executives of the FSPs demonstrate increasing interest in financial inclusion topics j. The FSP has designated a person or unit to lead outreach initiatives or establish new branches/ agents for financial services delivery k. Market actors develop and roll out new products

There are two overlapping aspects of market change that FSDs can bring about. Although these aspects may be treated differently in results frameworks – the FSD has direct control over the first type (described below) whereas the second type will depend on many other factors – **they both should be acknowledged as important outcomes of the programme:**

1. Market system development indicators that are directly caused and identified by the programme, and that for accountability purposes can be included in the logframe (often at output level). In this case, FSDs need to be quite specific about the types of change they anticipate bringing about in the market, and therefore indicators are quite focused on their action. These may include indicators such as 'improving the enabling environment' or 'strengthening the capacity of financial institutions to offer pro-poor financial services' that build on and may overlap with the 'easing market constraint' progress indicators noted above. Such indicators, often focused quite narrowly on an FSD (i.e. the rules and policies they have worked on, or the financial institutions they have worked with), offer an important, albeit quite narrow, part of a market development narrative that tends to correspond to the initial rationale for the setup of an FSD – i.e. removing market constraints to pro-poor financial services.

Discussion point: Some FSDs questioned IOM's emphasis on systemic change, given that they spend considerable time identifying the 'systemic market constraints' in their ex-ante analysis. This has led some FSDs to ask – "isn't all our work systemic?" The IOM recognises this argument to a degree, and section 5.1.2.2 that addresses 'easing market constraints' guides this type of measurement. However, the IOM argues that FSDs need to go further than this, and understand the broader system changes that result from removing these constraints, both from a FSD's perspective (see 5.1.3), and across the financial system as a whole (see 5.1.4).

2. For IOM purposes we are also interested in indicators that signal a more medium-term and long-term shift in the market, which may go beyond the immediate actions of the programme. These can be considered indirect outcomes facilitated by the programme and may take longer to achieve. They are still important to capture, but because of the longer time-frame, it will also make sense to measure intermediate indicators that herald the changes to come.

Tip: Even if an FSD focuses on direct market change indicators for the outputs in its logframe, it should be aware of how these relate to the systemic constraints in the overall market. It should then monitor beyond the logframe to capture such systemic changes, or incorporate indicators into the logframe (see Box 13).

Box 11 Why we need specific FSD systemic change indicators

Having concrete indicators for assessing these changes provides a number of important functions (both for measurement and implementation) because they:

1. build an evidence base for a programme to assess how systems and the market actors within them are changing over time. For some market development programmes (including FSD programmes) the rhetoric regarding influencing markets in this manner is not backed by their measurement systems;
2. provide an evidence base for FSD programmes to show their impact beyond the financial outreach caused directly by their interventions (providing FSDs with a basis for claiming results arising from the full range of their interventions); and
3. guide an FSD programme as to where and with whom to intervene next. For example, if an innovative business model that is helping an FSD target group is not spreading throughout the market, an FSD may consider additional intervention to promote this demonstration effect (see Figure 9, above).

Buy-in indicators, measure the degree to which FSD partners have taken ownership over new ideas, whereas the broader **market indicators** look at changes across a financial sector or a sub-sector.[38] This latter type of indicator is particularly important for FSDs: it represents a departure from other non-FSD market development programmes in that it can capture the broader effects of work involving directly improving the structures (e.g. rules, infrastructure) of the market. Whilst changes in partners beyond a project are likely to be relevant to all FSD interventions, the other indicators in Table 14 below are likely to be more suited to particular interventions (for example, replication is less likely to apply to policy work).

Tip: It is also important for FSDs to try to define the parameter of the 'system' in which they are operating, otherwise measurement can quickly become unwieldy. To that end using results chains that extend horizontally (as per Figure 9) provides a useful tool for thinking through what types of systemic change may be relevant to a particular project. To recap, the IOM will need to track three types of changes:

To recap, the IOM will need to track three types of changes:

1. **Partner changes supported by an initial project:** such changes will be tracked within each project ToC and/or logframe
2. **Partner changes beyond an initial project:** FSDs are interested in assessing if partners have adapted and expanded their practices beyond the specific project goals (an indication of systemic change). It is useful to think of three types of FSD partner change that can occur beyond a project:[39]
   a) **Adopt:** This looks at the potential for the partner to continue the project after the FSD support has concluded. For example, an indicator may assess if the partner is financing work to continue the project beyond the scope originally agreed.
   b) **Scale:** Related to the above, and particularly relevant to micro projects, partners may scale up pilots.
   c) **Adapt:** The partner tailors the project in some manner, for instance to add more functions to a particular service provided, especially if adapting a product or service to better suit low-income customers.

38. Another framework, the adopt–adapt–expand–respond model provides a similar set of indicators FSDs can think through when trying to identify potential systemic change effects of their interventions. This is presented in Annex E.

39. Note, these do not have to be mutually exclusive categories; they can potentially all apply.

Tip: These systemic change indicators are not meant to suggest a linear progression of change occurring. The time when to measure these changes will vary depending on the nature of the market and intensity of the intervention, with adopt, and scale potentially likely to occur before adaption and replication. Different indicators will also be more relevant to certain types of projects (see Table 14). Moreover, there may be feedback loops: for example, a proliferation of new business models may lead to changes in regulation, which lead to further changes.

Table 14 Generic systemic change indicator typologies

Type of intervention	1. Changes within the partner beyond initial project			2. Broader market changes		
	Adopt	**Scale**	**Adapt**	**Replication, demonstration, crowding in effects**	**Incentive / structural change**	**Resilience / responsiveness**
MACRO: New/ improved regulation established	Change in attitude, knowledge, behaviour, skills of central bank Change in the internal organisation within the central bank, e.g. creation of financial inclusion unit Central bank's ability to continue to support and implement the new regulation	n/a	Policy-makers are better able to respond to changes and opportunities in the market	Not applicable in the context of an individual FSD programme; however, there are definite opportunities for FSDs and even FSD 'macro' and 'meso' partners to share experiences, and in this way replicate successful approaches.	New/ improved regulations introduced triggering new ways of working by market actors	Not always applicable but over time policymakers or market norms may respond to FSD interventions at meso and micro levels
MESO: Banking association strengthened	Change in attitude, knowledge, behaviour, skills of association Change in internal management of the association	% increase in the number of active members of the association Association continues to offer new services to its members	Association is better able to respond to new challenges and expand into new service areas for its members	While it is unlikely for a banking association to be replicated within a country, it is possible for other meso-level interventions such as supporting a financial sector business service provider to replicate	Not applicable for this example, but other meso interventions such as collateral laws may change incentives for a broad group of actors	Not always applicable but over time market infrastructure may respond to FSD interventions at macro and micro levels

Type of intervention	1. Changes within the partner beyond initial project			2. Broader market changes		
	Adopt	**Scale**	**Adapt**	**Replication, demonstration, crowding in effects**	**Incentive / structural change**	**Resilience / responsiveness**
MICRO: New business model created New delivery system piloted New product launched	Change in attitude, knowledge, behaviour, skills of FSP Change in the internal management systems used to develop and provide products and services Partner's ability to continue the project once FSD funding stops (e.g. financial, human resources etc.)	Percentage increase in the number of customers that use products or services provided by the FSD partner (based on trajectories) Partner is scaling up, with innovation becoming mainstream and/ or new business practices pushing innovation to scale[41]	Change in the relationship (collaboration) between the partner and others within its value/supply chain FSD partner organisation moves into new areas/market segment	**Replication:** Number of subsequent partners/ providers that take-up the business model as a result of pilot **Demonstration:** Leads other market participants to change their behaviour, without FSD involvement[42] **Crowding in:** Extent to which other market actors (not the same as FSD partner) respond to FSD-supported approach	New market strategies/ business models (based on change in structure) Change in attitudes to enabling environment Respect for rules/regulations/ standards (e.g. adhering to voluntary/ industry codes of conduct and compacts) Sector growth rates (pre/post change in structure)	Market actors reorganising, assuming new/ improved roles or repositioning to take advantage of opportunities/ mitigate challenges that have been created[43] Banks setting up new divisions to mainstream mobile money,[44] or policy-makers developing new rules to manage mobile money

3. **Broader market changes:** The third type of change is beyond the immediate sphere of the FSD partners and relates to signs that the broader sector is adapting and changing, with the scale and breadth of change being important characteristics. This is where what has been a fairly FSD-centric monitoring perspective (bottom-up) also connects with our broader sector tracking perspective (see below). There are three main categories of change here, which micro and meso/ macro level projects influence differently:
   a) **Replication, demonstration, crowding in effects:** particularly relevant for micro projects, these assess how an FSD project has triggered changes among non-FSD partners (either to do something similar, or to adapt and build on what the FSD partner is doing). This has close links to the types of change the DFIs look for (Box 12).

Box 12 IFC approach to systemic change[45]

The IFC have identified a number of characteristics of systemic change but recognise that there is no precise approach to measure if these outcomes are occurring and a yes/no answer in relation to their presence is largely provided by Programs, backed up with evidence, where available. These characteristics include:

- Demonstration effect
- Attracting foreign direct investment
- New structures/ instruments (e.g. new laws, or type of providers)
- Viability of (new) financing instrument
- Viability of lending to (new) sector
- Replication of financing instrument
- Corporate governance:
  - Improving board structure and function
  - Systems improvements (e.g. risk management, information technology)

41. Adapted from AAER framework in Annex E
42. Adapted from IFC
43. Adapted from Springfield Centre (2014)
44. Measurement methods might include case studies undertaken post significant events - see Step 4
45. Demonstration: Spread of new behaviours and activities. Demonstration of replicable products and processes new to the economy; new investments stimulated by the project; demonstration of ways of successfully restructuring companies and institutions; demonstration of new ways and instruments to finance private sector activity. See http://www.ifc.org/wps/wcm/connect/39a9c900488773bd8de1fd299ede9589/FIG+DOTS+Indicators-Final.pdf?MOD=AJPERES

   b) **Changing incentives of market actors resulting from structural change:** particularly relevant to policy/regulations, infrastructure or even the advent of new market information (e.g. when FinScope was first introduced), this captures the effects of changing risks, costs and incentives of market actors (e.g. showing how lower-income market segments with strong demand might represent new opportunities for FSPs).
   c) **Resilience and responsiveness:** is difficult to measure but relatively easier to assess ex-post rather than ex-ante while finalising project results chains. This applies to all market actors, from policy-makers to informal savings groups. Measuring resilience and responsiveness has the aim of assessing how well actors and market structures (e.g. rules / infrastructure) respond to shocks in the market – either beneficial ones, such as the introduction of new technologies, or negative ones, such as drought or sudden changes in interest rates. The AAER model (see Annex E, with further FSD examples) provides a seful question to consider here. *If you left now, would the system be supportive of the changes introduced, allowing them to be upheld, grow and evolve?*

Tip: When to measure systemic change? During consultation with FSDs there was a worry that systemic change requires measuring many years after an intervention given its focus on sustainability. The IOM guidance argues that the change processes need to be tracked on an continuous basis as trend data will provide evidence of the type and pace of change (useful if an FSD needs to adjust its intervention). This can then be triangulated with FSD actions and timelines (i.e. to measure an FSD's nature and timing of contribution). Some changes, particularly those at the macro/ meso level, may take some time to have an impact on financial sector outcomes, and may need to be measured one or two years after the intervention. This in turn has implications for the need to be flexible in terms of establishing/ updating baselines and agreeing that partners will share data on a long term basis (see Box 18).

FSDs can use these generic types of indicators to formulate and prioritise indicators that are specific and time-bound. Examples are provided in Annex D. The typologies of indicators set out in Table 14 above provide some examples which FSDs can use to adapt and develop context-specific indicators at project and programme level. This will help to ensure that FSDs consider systemic change mechanisms apart from delivery indicators (e.g. the number or value of SME loans). Which indicators are considered important by an FSD programme will depend on both the type of project and at what level of the market the FSD is operating.

Tip: In market development programmes some of the logframe indicators can and should change over time. For example, for each output there could be a mix of process and results indicators. Apart from measuring whether outreach has expanded, both the implementers and funders need to track progress in underlying structure and processes, which will be different over the course of the programme life cycle. Change in partner behaviour in a particular country may be critical in year one, but by year three the measurement focus may need to shift to demonstration effects in that country. Also, initially FSDs may place greater effort on providing support to FSD partners to test business models, gain market knowledge and credibility, while for more mature markets a greater proportion of effort is likely on improving the policy environment and supporting infrastructure.

Box 13: Linking systemic change indicators to the logframe

The monitoring of the above indicators will provide useful evidence for IOM purposes, but it can be challenging to incorporate such indicators into logframes, especially if the latter largely focus on direct results from FSD programmes (as an accountability tool) and assume linear progression from outputs to outcomes (See Tip above). However, FSD programmes can consider a few options to ensure that, for this snapshot of programme-level progress from a bottom-up perspective, they are able to report their overall impact to funders.

1. **Incorporate indirect outreach numbers resulting from systemic change:** for example, this would require aggregating financial access outcomes from those market actors with whom the FSD programmes have not been directly involved but who have copied FSD partner business models, or responded to changes in risk/costs and incentives resulting from FSDs improving the enabling environment. These can be recorded as 'indirect' outreach numbers in the logframe, with the evidence to demonstrate the contribution of FSD to these numbers being provided as part of the broader annual review.
2. **Provide an outcome proxy for the changes in the underlying structure of the market:** it is very difficult to provide just one or two indicators that can summarise if a market has developed (for example, is it sustainable, able to reach a greater number of people, and is it better able to deal with shocks?). However, using the set of indicators listed in Section 5.1.4 below to show that a financial sector has developed over time (e.g. efficiency, diversity, size, outreach, stability), it may be possible to provide a weighted average indicator (at outcome level)[46] for a few categories that are most relevant to an FSD.
3. **Incorporate systemic change indicators directly in logframe:** this would shift an FSD not only to focus on concrete outputs (e.g. laws passed, infrastructure strengthened, business models developed), but to attempt to provide a summary as to the sustainability, resilience and scale of these projects. These systemic change indicators would be included as output-level indicators in the logframe, and could be both programme focused (e.g. no. of projects showing sustainability) or focused on specific areas of the market (e.g. assessing if the enabling environment was more effective). Caution should be applied when adopting this approach as there is a risk that FSDs will be held to account for changes that are only partly in their control. FSDs can only use such indicators if funders allow variation in the use of indicators relevant for the next reporting period and these change over time as progress is achieved and new market development aspirations are agreed.

5.1.4 Top-down and bottom-up impact-oriented indicators

The sections above focused largely on extending FSD output monitoring, in order to think more about measuring FSD facilitated systemic change as an intermediate outcome (as shown in the ToC diagram in Figure 4), and in order to be able to confirm how FSD interventions have promoted such change. This section focuses more on the sector level (top-down) – that of the supply and demand for financial services, and how these dimensions can help FSDs to augment their evidence as to how the underlying dynamics of the market have changed (from a non FSD perspective). Top-down analysis can start with sector-wide analysis and data already available, e.g. annual reports of the central bank or other regulators, a special study by bankers or micro-insurers association, or relevant analysis from a Financial Sector Assessment Program (FSAP) report or an International Monetary Fund (IMF) Article IV consultation. A special study should be considered by an FSD only in regard to a specific evidence gap.

Incorporating top-down monitoring to complement the more traditional bottom-up monitoring is critical for bridging the evidence gap described in Section 5.1.1.2 above. Figure 14 shows different and complementary methods applicable for bottom-up and top-down monitoring and Table 15 summarises indicators for both bottom-up and top-down measurement of programme outcomes and impact. These indicators can be collected at a sector level and, in combination, help build a contribution story. For example, increases in the provision of financial services through FSD-supported institutions can be compared with overall increases in the sector.

46. Further work will be needed by FSDs and FSDA if this is considered to be a priority area.

Figure 14 Tracking changes in the ToC

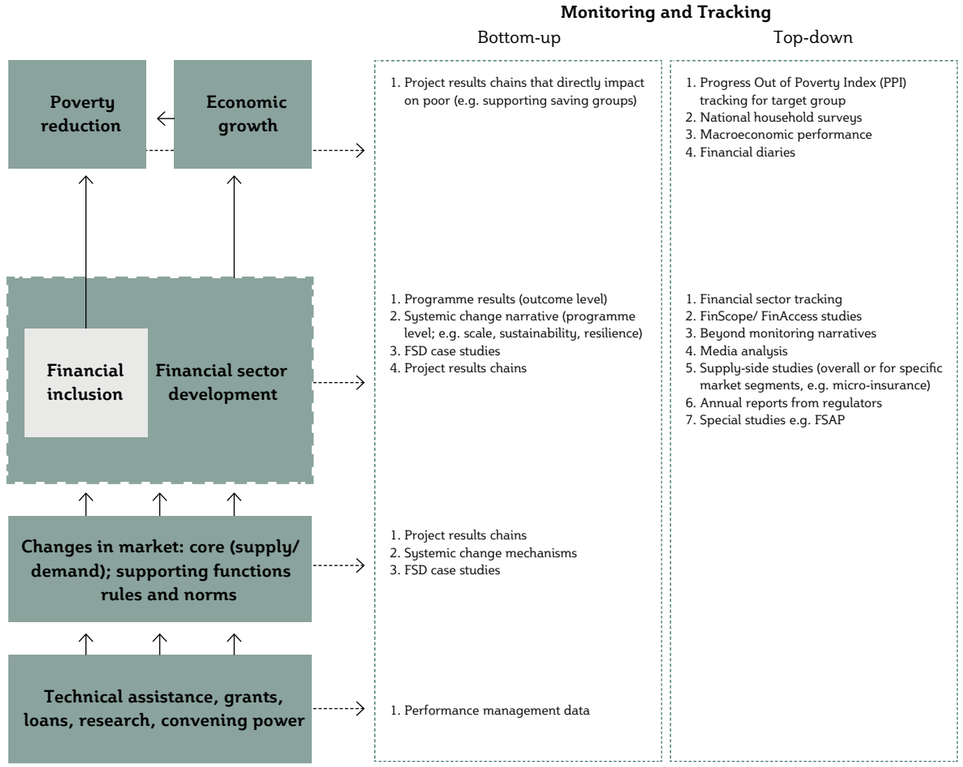


Table 15 Types of indicators – programme outcome and impact (top-down and bottom-up)[47]

PROGRAMME OUTCOME		IMPACT	
Financial Inclusion	**Financial Sector Development**	**Livelihood/ Poverty Reduction (for more details see Box 14 below).**	**Pro-poor Growth**
Examples of sector focused indicators – top-down			
FinScope/ FinAccess surveys complemented with: i. Frequency of financial services use ii. Reduced cost of use iii. Reduced account inactivity/ drop-out rates iv. Financial diaries and/ or ethnographic research to provide qualitative data on how people and small businesses view specific financial products, services and providers	i. Reduced use of cash transactions ii. Increased use of mobile money payments and remittances iii. Deposits/ GDP per capita increasing iv. Increases/ growth in credit to SMEs as a ratio to total credit outstanding v. Percentage of small businesses identifying access/ cost of credit as a major constraint (see also separate paper on Tracking financial sector development)	i. PPI ii. Reduced number of occasions when household needed to borrow from money-lenders to cover daily expenses or emergencies	i. Increase in private sector investment ii. Average growth rate for last three years iii. Additional jobs created in SMEs Possible tracking through: efflcient exchange of goods and services; mobilise/ pool savings; allocate capital (at financial sector and firm level); risk diversification and management. See Section 5.1.6
Examples of FSD-focused – bottom-up – aggregated indicators (including outcomes as a result of partners scaling up, replicating and responding to changing risks and costs resulting from structural changes)[48]			
(i) Number of individuals/ enterprises using financial services as a result of FSD interventions (ii) Volume of credit / deposits provided by a cross-section of FSPs supported by FSD programmes	(i) No. of banks and other financial institutions using FSD-supported credit reference bureaus (ii) No. of new entrants to sector supported by FSD programme (iii) Efflciency improved amongst FSPs supported by FSD programmes	(i) No. of people provided with income opportunities (ii) No. of people reporting reduction in losses when savings Many projects and factors (outside the control of FSD programmes) contribute to outcomes and impact, and FSD programme focused indicators are not expected at the impact level. However, project-specific evaluations may, for example, attempt to assess how income changes for a specific group of end-users benefited from the project	(i) No. of jobs created by enterprises funded by FSD-supported partners

Note: See separate papers on The relationship between financial sector development, economic growth and poverty reduction; Tracking financial sector development; and Assessing the quality of access.

47. See Annex C for a discussion of how FSD programmes are currently tracking these indicators in their logframes.

48. It may be useful to keep indirect results (i.e. those from non-FSD partners copying an FSD-funded project, or from macro interventions stimulating a broader market responses) separate in the IOM system and the logframe.

Box 14 How far should M&E focus on the final impact of poverty reduction?

In general we advise FSD programmes to focus on financial sector outcomes, rather than expending significant resources on trying to understand the final impact on poverty reduction. Evidence should show that links to poverty reduction are present and, for a number of interventions, it should be possible to use the emerging evidence (linked to the type and context of your project, e.g. are poor people and small businesses getting financial services?) to show that these links exist. However, there may be times when FSD will be able to go further and collect specific data on poverty impacts of individual projects. Indeed many of the reasons listed in Box 25 in Step 5 on when to undertake in-depth evaluations of specific projects apply. The process may, for example, include collecting evidence on changes in livelihoods/poverty reduction when implementing a particular innovative project or a particularly expensive project (where VfM analysis may also be applied). These can be done to convince realtively informed critics of the link between market change and peoples' lives. Again, there is a spectrum as regards the types of studies/methods that can be used to collect and analyse such data. For example, light-touch methods for collecting poverty data may include FGDs, compared with more extensive (and expensive) representative sampling and before-and-after surveys.

5.1.4.1 Measuring and tracking financial sector development (Top-down monitoring)

Tracking financial sector development is crucial for FSDs to meet several different objectives:

i. to augment the evidence base for impact evaluation (e.g. it captures how interventions are changing the financial sector in ways that go beyond FSD partners) as well as to pick up unexpected impacts;
ii. to improve understanding of how complex and dynamic markets are changing and place the FSD interventions in the overall market context[49] so that FSD managers can assess the role and progress of their own programmes in order to develop an overall credible narrative around impact;
iii. to identify priority areas/gaps for future FSD work; and to strengthen dialogue and advocacy.[50]

FSDs are familiar with some of these objectives and have used FinScope and other studies to pursue similar objectives in the past. However, very few FSDs have systematically collected and used the financial sector data so far, especially on the supply side – and even on the demand side few have mined the FinScope data as deeply as they might. Of course, some financial sector data are at a much higher level than FSD contributions; but without an overview it is quite hard to map where FSDs are, or to prioritise in which sub-sector/market there is an opportunity for an FSD intervention to make a major difference.[51]

Discussion point: During the consultation FSDs commented that top-down measurement is difficult as it is always not clear where to start. However, there was agreement that a clearly articulated ToC (and theme based ToCs) can help provide some limits on the type of sector tracking an FSD undertakes. For example, an FSD focusing on finance for growth may want to focus on longer term loans whereas this need not be the case if the FSD focus is mainly on finance for all. Moreover, this type of tracking is intrinsic to what FSDs have already been doing, for instance with their funding of FinScopes/ FinAccess. This not only provides an essential source of data for measurement, but it also allows for informed discussions with policymakers and other stakeholders based on non-FSD specific set of credible evidence.

Tip: Given their lack of direct relevance to FSD interventions, sector-wide data need not be used by an FSD to show accountability of the FSD's performance through annual reporting processes, but they can and should be used for broader IOM measurement purposes.

As noted at the start of this chapter, a particular challenge for FSDs is measuring how the underlying structures of the market have changed. We have showed how to assess this from the perspective of FSD projects, but being able

49. For example, what are the implications for FSD-supported financial institutions showing x for a particular indicator while the sector is showing y for the same indicator? (x and y could be absolute numbers, growth rates or percentages, depending upon the specific indicator being analysed).
50. Sector-level data and insights can also provide a public good function (used by national policy-makers, market actors and other market observers). The process of identifying/prioritising sectorlevel indicators and discussions about how these market data should be compiled/ funded and disseminated in itself can be an important part of the FSD market facilitation function. These data also provide evidence for an active dialogue with financial institutions, policy-makers, bankers' association and others as to why the needle is shifting in some indicators (and not others) and what more can/ should be done.
51. There is something of a parallel here to a share market operator who needs to track the overall changes in the capital market even though the trader may only be interested in a few shares, bonds or other financing instruments.

to track how the sector is developing and changing over time is also important. Most of us could point to the structures (an enabling environment, innovative business models, developed infrastructure and skills, high levels of financial capability etc.) that we would want to see in a more inclusive market, with market players that were sustainable, had achieved scale and were resilient. But this is hard to measure in practice. Indicators that relate to financial sector development, along with descriptions of how forms in the market have changed (e.g. new rules/ credit registries and the entry of new players), provide useful proxies for how underlying dynamics are operating.

Financial sector development incorporates (at least) certain types of change in a range of indicators (size, depth, etc.). Indicators that can be used to measure these changes, and which can be collected at relatively low cost by all FSDs, are identified in Table 16.[52] The trade-offs in selecting these indicators are further elaborated in a separate paper.[53] When considered together, the proposed indicators can provide a useful overview of the state of a country's financial sector and its evolution over a number of years.

Box 15 Do all interventions need top-down sector tracking?

Some form of understanding of how the sector or sub-sector is changing is important for most interventions. However, the sector/sub-sector indicators provided in the Section 5.1.4.1 are likely to be useful across a number of interventions, meaning that there is no need for additional sector analysis for each new intervention. Over time it may be found that these indicators are not related closely enough to an FSD project (to build up a contribution narrative), and will therefore need to evolve accordingly, with new indicators added and others dropped.

Discussion point: Recognizing that the total number of indicators the paper identified exceeded 30, there was a risk that if, for the sake of ease of data gathering, an FSD selected to track two or three of the five categories of indicators, these might paint a misleading picture. The example of Myanmar was given where, judged by some limited number of measures, the country might be considered to have a well-developed financial sector, which is clearly not the case. So it was probably important to track all five categories, even if the total number of indicators followed was reduced to a more manageable number.

Table 16 Tracking financial sector development – some common indicators

Financial sector change	Indicator
Size	Ratio of bank deposits to GDP Ratio of private credit to GDP Ratio of cash holding to deposits (given interest in 'cash-lite' economies)
Depth/diversity	A listing of the country's main non-bank financial institutions with a simple summary of their size based on total assets A composite indicator[54] comprising the following (or sub-set of): insurance company assets to GDP; life insurance premium volume to GDP; non-life insurance premium to GDP; pension fund assets to GDP; stock market capitalisation to GDP; stock market total value-traded to GDP
Access/inclusion	ATMs per 100,000 adults Bank accounts per 1,000 adults Bank branches per 100,000 adults Received loan from a financial institution in last year (% of adults) Saved at a financial institution in last year (% of adults) Small firm with a bank loan or credit line (%)
Efficiency	Bank net interest margin as a percentage of earning assets Bank overhead costs as a percentage of total assets Bank return on assets before tax (%), and return on assets after tax (%)
Safety and soundness	Ratio of bank regulatory capital to risk-weighted assets Bank non-performing loans as a percentage of gross loans Narrative descriptions from financial stability studies

52. Some individual FSDs may also be able to collect more granular information from their respective central banks at low cost.
53. See FSDA website, at www.fsdafrica.org/knowledge-hub.
54. More details are available in a separate paper on tracking financial sector development. See, www.fsdafrica.org/knowledge-hub

In total some 30 indicators have been identified. In many cases the data for use in the monitoring process can be assembled from readily available international databases. However, these sources would need to be supplemented by some limited new research efforts to extract relevant 'new' data from national data sources that will be available in most countries. They can also be supplemented by country-specific indicators if an FSD has ready access to more bespoke data.

For a stronger narrative, it will also be helpful[55] to analyse these indicators on a disaggregated basis: i.e. are financial inclusion-focused financial institutions as sustainable as those who serve the entire market, and/ or is efficiency and profitability of financial institutions improving over time? Similarly, while the bottom-up analysis may look at growth of FSD-supported FSPs, placing these data alongside market data can strengthen the narrative, e.g. absolute size and number of SMEs financed by FSD-supported FSPs and the overall market, and how this has changed from year x1 (when FSD support started) to period x2. Or if an FSD focuses on a particular sub-sector, say agriculture, additional indicators to the above can be collected, as shown in Table 17.

Analysing such sub-sector data over time may reveal surprising or unexpected trends. Whilst useful in itself for FSD programming, the data can also be compared back to FSD interventions, to see if the interventions were having unexpected impacts revealed by such trends.

5.1.4.2 Tracking financial sector development contribution to economic growth

Possible indicators for tracking FSDs' contributions to economic growth can be identified by considering functions such as: efficient exchange of goods and services; mobilising/ pooling savings; allocating capital (at financial sector and firm level); and risk diversification and management. Table 18 below provides some suggestions and comments in this area: a separate, more detailed paper (see Tracking Financial Sector Development) is available to explain why these indicators have been shortlisted.

Table 17 Focused sector tracking – illustrative example (agriculture)

Agri-financial sub-sector change	Indicator
Size	Ratio of bank (and any non-bank financial institution (NBFI)) deposits from agri-businesses and farmers to GDP Ratio of private credit to agri-businesses and farmers, to GDP
Depth/diversity	A list of the country's main NBFIs with a strategic focus on agri/rural-finance (e.g. insurance companies, savings and credit co-operatives (SACCOs) and microfinance institutions (MFIs), plus a simple summary of their size based on total assets[56] The proportion of adults sourcing credit from NBFIs and savings and loans and other informal groups (FinScope)[57]
Access/ inclusion[58]	Agri-finance loan from a financial institution in last year (% of adults) Agri-businesses and farmers who have saved at a financial institution in last year (% of adults) Agri-businesses with a bank loan or credit line (%) Purchased agriculture insurance (% working in agriculture, aged 15+) Received payments for agricultural products: deposited in an account at a financial institution (% recipients, aged 15+)
Efficiency	Bank net interest margin on agri-credit as a percentage of earning assets Bank return on agri-assets before tax (%) Top banks engaging in agricultural sector – net interest margins on all credit as a percentage of all earning assets[59]
Safety and soundness	Bank non-performing agri-loans as a percentage of gross agri-loans Narrative descriptions from any financial stability studies

55. We still encourage analysis of the above indicators in order to obtain an overall picture.
56. Use cut-off for minimum size, possibly based on number of members, e.g. 25,000.
57. FinScope also provides access indicators.
58. Agri-insurance indicators available from Findex and the World Bank can also provide disaggregation by gender.
59. This would be a proxy to be used in the event that central bank data were not available specifically for agri-finance.

Table 18 Indicators to track FSD contribution to economic growth

	Theme	Possible indicators	Comments
1	Efficient exchange of goods and services	Volume of transactions performed through the banking system	Need to check with the central bank as to whether such data is easily available. The payment system survey by the World Bank may have these data, but it may not be available on a regular basis
		Transaction costs for payment services	Some countries have started collecting such data at bank level, but there is a need to check broad data collection efforts in this area
		Ratio of cash in the economy to deposits in the banking system	This might be the best and most readily available indicator
2	Mobilise/pool savings	Ratio of total deposits to GDP	Readily available
		Ratio of total credit to GDP	Readily available
		Loan–deposit ratio (a gauge of intermediation efficiency)	Can be easily calculated
3	Allocate capital (financial sector level)	Percentages of MSMEs noting constrained access to loans and other financial products (e.g. restrictive collateral requirements, high application fees and lengthy processes)	Assess percentage of MSMEs that obtain a loan compared to those who say they need a loan (i.e. exclude firms that do not apply because they do not need a loan)[60]
4	Allocate capital (firm level)	Ratio of assets of NBFIs to GDP	Ideally defined as the combined total assets of insurance companies, mutual funds and pension funds as a percentage of GDP. Initially data may be available for only some NBFIs; this that should be ok as long as similar metrics are used for comparison across countries and over time
		Doing Business indicator of property registration or efficiency of credit information-sharing	Important to note that these indicators are actually institution/ policy variables and – unlike the others – not 'output' measures of the banking system
5	Manage liquidity risk	Share of loans with a maturity above one year relative to demand deposits	Might be available for individual countries, more difficult at cross-country level
		Actual number of firms listed on the stock exchange	Readily available
		The availability of a long-term yield curve in the economy	May be available for some countries
6	Risk diversification and management	Life insurance penetration	
		Market capitalisation of listed companies (% of GDP)	Readily available
		Stock traded, total value (% of GDP)	Readily available
		Ratio of cash in the economy to deposits in the banking system[61]	This might be the best and most readily available indicator

60. This is a difficult indicator as all the MSMEs who aspire to obtain a loan but may not be worthy of a loan: e.g. entrepreneurs may gloss over their own weaknesses and just blame the banker even though bankers may be right to refuse a loan in some circumstances.

61. This has been listed for efficient exchange of goods and services. The data are collected only once but can also be used to track evidence of changes in risk diversification and management at the level of the economy.

5.1.5 The use of qualitative indicators and monitoring beyond indicators

Some indicators are a straightforward quantitative measurement of what happened – for example, 'number of policies changed'. These indicators are well suited for inclusion in the donors' logframe or results framework because they are clear cut and easy to measure. On the other hand, other indicators address more complex outcomes. Often such indicators signify a mechanism of change rather than the change itself, let alone the desired end results of enhanced financial inclusion, financial sector development or impact on livelihoods.

Such data are critical for capturing and providing valuable insights for attribution or contribution analysis and impact evaluations. Examples might include changes in policy-makers' and regulators' attitudes (e.g. regarding the importance of informal savings groups for financial inclusion), changes in FSPs' strategies towards serving lower-income segments (also perhaps resulting from changes in attitudes – to the enabling environment, for example), and willingness of MNOs and FSPs to support inter-operability and non-exclusivity of mobile money agents. As we noted earlier, these types of indicators could be substantiated by memos recording conversations with senior regulators or FSP executives, or surveys or FGDs.

Table 19 provides examples of qualitative indicators and how they can be measured. The table also includes possible quantitative proxy indicators. These do not measure the qualitative indicator adequately, but they may be easier to track with greater frequency than qualitative measures; quantitative proxies are also helpful for triangulating the findings from the qualitative data collected.

Table 19 Examples of qualitative indicators

Qualitative indicator	Qualitative source of data/ means of verification	Possible quantitative proxy	How to report findings and results
Improved enabling policy environment	Meeting notes or recording from interviews, FGDs with regulators and/or with FSPs Special studies	Number of regulatory reforms Score on EIU Microscope[62]	Synthesise findings and report key findings and evidence in quarterly/annual report
Changing attitude of policy-makers towards financial inclusion issues	Meeting notes or recordings of interviews with regulators	Number of international meetings or events on financial inclusion topics attended by representatives of the central bank Perception surveys of selected market actors and/ or policy-makers (calculating average score on a few indicators)	Synthesise findings and report key findings and evidence in quarterly/annual report
Improved capacity of FSPs to reach 'down market'	Meeting notes or recordings of interviews with FSPs. Surveys of FSPs and/or FSP staff	Number of training sessions or TA received by FSP Proxy indicators for capacity agreed with FSPs and scored	Synthesise findings and report key findings and evidence in quarterly/annual report
Changing behaviour of FSPs to improve consumer protection	Meeting notes or recordings of interviews or focus groups with FSPs and clients. Surveys of FSPs and/or FSP staff	Non-performing loans Percentage of active account users Number of consumer complaints received by FSP or by regulator Percentage of complaints addressed	Synthesise findings and report key findings and evidence in quarterly/annual report. Case study

62. http://www.eiu.com/public/topical_report.aspx?campaignid=microscope2014.

Taking the concept of qualitative indicators a step further, it should be noted that even quantitative and qualitative indicators alone rarely present a full picture of the results that are being tracked. Results chains and logframes place an emphasis on the FSD as the influencing factor in the financial sector, but given the size and complexity of the financial sector, this may miss important relationships, different perspectives or influences outside FSD contributions, or new approaches/ perspectives/ providers that are much more important but are not being captured through indicators currently being tracked. For instance, as noted above, qualitative and quantitative indicators may not capture unplanned or unforeseen outcomes. This is a particular issue with complex programmes like FSDs, where results and their indicators are often difficult to define precisely in advance. In addition, such indicators only provide insights perhaps once or twice a year when they are updated. It is therefore important that we capture evidence of outcomes beyond specified indicators.

Box 16 Example of measuring beyond indicators: financial protection

An FSD seeks a long-term outcome: firms take concrete steps towards the adoption of new policy and industry practices that better protect base-of-pyramid financial consumers. Thus, the FSD could commission a comprehensive study to track firms that are adopting these practices. Alternatively, the FSD could ask their stakeholders (e.g. FSPs, trade associations or central banks) to refer cases that they come across, and FSD team members could also capture examples themselves, through field visits or through desk research. An FSD programme may not be able to establish how representative these cases are, but they can learn from them, as case studies, and feed that knowledge back to the relevant market actors with a view to improving the outcome(s) sought.

For sector-level tracking, there are also indicators – often broad in nature and more conducive to narrative description than direct trend analysis – that might be used to look at the systemic changes occurring in the sector from a broader perspective than those provided above in Table 14. There is no set list of things that can be identified, but Miehlbradt and McVay (2006) provide a number of examples of signs that a market system is functioning which can be picked up by such narrative reporting; including if attitudes are changing (Box 17), if a major event has taken place (such as a new financial institution has entered the market, or different organisations have started to create partnerships (e.g. between MNOs and banks)).

Box 17 Attitudinal change

Because some of the changes that FSDs aim to effect address fundamental market behaviour and are long-term, it is important to identify indicators that progress is being made towards those ultimate goals.

Changes in knowledge and/or 'mental models' (including attitudes) by policy-makers and other important market actors may indicate future changes in behaviour by these individuals as well as the market system. Such actors can include large bank and insurance company CEOs, central bankers and policy-makers.

Attitude changes of interest could include, for example, the attitude of policy-makers towards the relaxation of KYC requirements in favour of a more accessible formal financial sector, or the attitude of MFIs towards predatory lending practices. These changes can be tracked through surveys, FGDs, key informant interviews, media monitoring, case studies, and even through observations by FSD staff.

Tip: One way to explore these measurements beyond indicators is to define the specific assumptions or hypotheses based on the ToC that it is important to track and that are not sufficiently captured in the quantitative and qualitative indicators.

5.1.6 Impact-oriented indicators in practice

As market development programmes, FSDs also need to answer certain questions, as regards how they use indicators: **Are they useful?** FSD programmes should select indicators based on whether they are useful in relation to managing interventions, providing accountability to funders annually (through the logframe) and demonstrating impact. For IOM purposes the key is that indicators can help answer the measurement questions that are of interest to the FSD.[63]

- **Are they realistic/ feasible?** FSD programmes should be realistic and pragmatic in selecting indicators, and should choose a manageable number of such indicators, for which data can be collected. FSDs need to consider the people, time and money required to track these indicators.
- **Are they varied?** FSD programmes should select a variety of indicators that monitor shorter and longer term changes, in order to assess progress along the pathway. For example, FSDs should not only select indicators that establish longer-term systemic change, but should also choose indicators that measure short-term behavioural changes that are necessary for the longer-term changes to occur (and that help FSDs anticipate such longer-term changes).
- **Are similar indicators being used by others?** FSD should discuss the proposed indicators with the implementing partners and see which indicators could meet the needs of both the FSD and their partners. The FSD programme team can also draw on high-level indicators already being promoted by CGAP, DFID, the Global Partnership for Financial Inclusion (GPFA), World Bank, the Alliance for Financial Inclusion (AFI) and others.
- **Are they measurable?** The indicator should specify the **qualitative** or **quantitative** unit of measurement that will be used. Baselines, a key measurement tool for traditional monitoring, may also need to be adapted for dynamic FSD programmes (see Box 18). Given the wide range of indicators discussed as part of building an IOM system (on top of those core measurement processes FSDs are already using), it can be helpful for FSD programmes to develop detailed technical notes on each indicator – referred to as 'indicator profiles'. Indicator profiles can help to ensure that indicators are well understood by all potential users (stakeholders) and that there is a plan in place to capture information related to them. The indicator profile describes the indicator and its rationale, and clarifies definitions where needed. It identifies:

  i. the data source(s) and method(s) for data collection and any cost implications that they might have;
  ii. the baseline and targets, as well as the rationale for both; and
  iii. the responsibilities for data collection, analysis and reporting.

Tip: Where possible indicators should be prioritised in consultation with sector stakeholders. This prioritisation should recognise the fact that indicators are used not only for accountability to funders, but also for influencing decision-makers and market actors.

Box 18 The role of baselines

Baseline information is important for both the programme as a whole and for individual projects as it allows an FSD to compare the situation at the beginning of the intervention(s) with that at the end, to establish what change has occurred. Baseline information can be collected for all indicators, at all levels of the ToC, so that predicted change can be assessed against actual change, and adaptation can be made. **But given the nature of FSDs, some caution is required before significant effort is expended on gathering baseline data.**

At the programme level, it is difficult to design a detailed baseline for an FSD, although sector indicators such as level of overall inclusion are still useful. More bespoke baselines, focusing on specific households and enterprises, risk becoming obsolete as FSDs' plans (e.g. type of interventions, geographical areas, objectives) change or are adjusted.*

At the project level baselines are easier to construct. However, FSDs should still be pragmatic given the dynamism of FSD interventions: some of their interventions are pilots, and may not be scaled up; the formation of partnerships, for example, takes considerable time, so baselines should not be too fixed in case these relationships evolve in unexpected ways; and as FSDs are intervening in dynamic contexts/ trajectories of growth, crucial factors may be missed in baseline information that is obtained ahead of the full-scale implementation (see Figure 15)

Tips to mitigate these risks, include the following:

i. Effort can be made to identify pre-existing data that can be used for baselines. For example, this may draw on information gathered during the FSD diagnostic processes or information collected by FSD partner(s) as implementation proceeds.
ii. It may be easier and far less costly to start with supply-side data, which will be (or can be) more easily collected by the implementing partner, as against demand-side data for customer behaviour and usage.
iii. FSDs should not devote all their M&E resources to a programme baseline; instead, they should use a range of techniques, existing sector data and project monitoring and baseline information, as for different programme priorities, different data may need to be collected at different times.
iv. For complex interventions, it may be necessary

63. Over time it is also useful to include checks as to 'Who is using (or going to use) the information?' and 'can it be used in a practical way?' If the answer to either of these questions is 'no-one' and 'no' for a specific indicator, then the indicator can be dropped or changed. Also, if FSD programmes notice important gaps, they can fill them by identifying what additional information or analyses are needed.

to establish baselines retrospectively once FSD understand how change is proceeding (e.g. through recall methods, documentation and other sector data). This is more useful than comparing an intervention against an inappropriate baseline.

v. FSDs should ensure that the baseline is appropriate given the trajectory of change in the area of intervention. This may require including trend growth rates in a baseline rather than only relying on performance at a particular point in time. It may also require updating baselines as the programme approach (and understanding) evolves.

For further reading, see Kessler and Sen (2013); Springfield Centre (2014)
**See OPM (2014).*

Table 20 Types of indicators and change trajectory

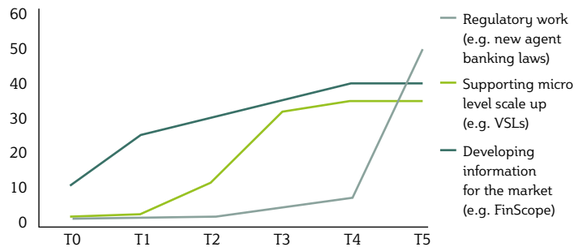
Type of indicator	Description
Leading	Provides information before the result takes place
Coincident	Yields information at about the same time as the result
Lagging	Provides data after the result takes place, often with considerable time lag, either due to data collection routines and/or long results chains

Source: Britt (2013)

– **Has the time-dimension been carefully considered (and have the projected results)?** Given the process of change is often non-linear, the indicator should consider the time-frame over which the change is expected. Table 20 highlights the fact that indicators can be used to provide information at different points of the change process. For FSDs, lagging indicators are likely to be less useful for programme management but are more useful in terms of evaluating how a system has changed over time. Where results may take considerable time to emerge (for example, in the case of policy change), an FSD can rely on indicators that would be expected to change in the shorter term, leading towards the expected longer-term change.

Furthermore, Figure 15 shows how different types of intervention can have different trajectories, in terms of when they produce results. Thus, the dates by which they can be expected to achieve their results will differ.[64] However, given the unpredictability of the pace of change, targets should be realistic, and should not be pursued at all costs, if that undermines the change process (for example, through moving away from facilitation towards more direct forms of intervention/ delivery).

Figure 15 Possible trajectories of impact by different interventions



– **Do the indicators need to be updated?** As with all aspects of a measurement framework, FSDs must also review and revise their indicators as their programmes and environments evolve. It is important to note that any revisions affect the ability to compare indicators with baselines, and to conduct trend analysis. Box 19 below describes FSDK's recent experience with its outcome and output indicators.
– **What incentives do the indicators provide?** Measurement systems can by themselves create incentives and distortions. Implementers are encouraged to focus more on what is being measured and may miss other important, but difficult, reform processes. For example, short-term support to an FSP may be justified to kick-start financial services delivery, but if the indicators being tracked only focus on outreach while failing to pick up issues around costs of delivery/ efficiency and customer response/ drop-out rates, this may not lead to sustainable business models beyond the period of an FSD's support, and other FSPs may be put off rather than encouraged to focus on this market segment.

64. Please note that these charts are hypothetical. The numbers on the y-axis are purely illustrative and the times shown on the x-axis could be months, quarters or years.

Box 19 FSDK's experience of updating indicators

In November 2014, after a mid-term review, FSDK made a number of adjustments to its outcome and output indicators. The main reason for changing the programme-level (outcome) indicators was practical: some of the indicators were ill-defined or there were simply no data to track them. Every year DFID and FSDK struggled to report against these indicators, and reporting to the donors on these indicators was not very effective. The idea was therefore not to revolutionise the logframe, which would have required a lot of negotiations with DFID and the PIC, but rather to maintain the core meaning of the indicators and to change them so as to make them more feasible and more meaningful as a means of assessing FSDK's performance.

The changes to the indicators were spread across the four themes (Formal financial services, direct poverty impact, inclusive growth and knowledge), based on the quality of the individual indicators rather than how they collectively measured the impact of the theme as a whole. For example, FSDK changed the indicator 'Average cost of a single retail transaction through formal financial providers' to 'Cost of a KSh 500 electronic transfer across most widely used retail payment platform'. The problem is that a 'single financial transaction' can be defined in very many different ways. By contrast, the revised indicator, though very broad, is feasible and not open to interpretation. At present, M-PESA is the 'most widely used retail payment platform', but this could change over the next few years. The KSh 500 figure might sound like an arbitrary amount, but it was chosen based on findings from financial diaries, showing that among low-income household, KSh 500 is the most common size of payments considered 'large-scale'.

Source – discussions with FSDK, 2015

Box 20 Step 3 checklist

– Indicators are aligned with the ToC and results chains, and the overall reporting is agreed with the funders
– Have you considered the different types of indicators suggested;
  – progress indicators;
  – market system development indicators;
  – top down sector tracking; and
  – 'beyond indicators'?
– Ensure your indicators distinguish between indicators used for accountability and those indicators which will help track and test the impact measurement questions (some overlaps can be expected)
– In the final selection of indicators prioritised, do you have clear indicators for systemic change and sector tracking
– Does the set of indicators adequately fill the gap between programme outputs and the final desired market change?
– Indicators capture key quantitative and qualitative data (especially, in the case of the latter, for sector tracking). Are you capturing both at different steps of your ToC/results chains?
– Have you prepared indicator profiles for each selected indicator – definition, rationale for use, the data source(s), frequency and method(s) for data collection, cost implications, and who will be responsible for data collection, analysis and reporting?
– Have you collected baseline information where possible? Have you set realistic and transparent targets (i.e. based on evidence and explicit assumptions) for those indicators?
– Have you established processes to:
  – periodically check if the indicators being measured miss a focus on key drivers for expected change and create distortions in the behaviour of FSD staff and/or implementers?
  – Are there unintended and/or negative impacts happening? Do you have processes to measure these?

5.2 Data collection methods and sources (Step 4)

5.2.1 Overview

– This step sets out examples of various types of data sources, mapped back to the indicator framework presented in Step 3. It includes:
  – guidance on relevant methods and sources for collecting information on systemic change;
  – sector tracking; and
  – monitoring beyond indicators.
– It argues that a mixed method (quantitative and qualitative) approach to evidence collection is appropriate for FSDs.
– It suggests tips for assessing data quality — in particular, what to watch out for with regard to supply- and demand-side data.

5.2.2 A mixed methods approach

As noted above, the use of both quantitative and qualitative data is important in order to understand all the changes and related processes FSDs are assessing. **A mixed methods approach is therefore recommended, no matter whether data collection and analysis is under-**

taken at the output, outcome or impact levels. Key arguments for using mixed methods include (Lund, 2014):

- Mixed methods can answer complex research questions related to both describing causal paths and explaining how they work.
- Mixed methods research may provide robust inferences regarding causal paths.
- Qualitative and quantitative results may sometimes be contradictory and can generate new insights.

A mixed methods implies the use of a combination of quantitative and qualitative methods, with data that can be captured from primary and/or secondary data sources (see Table 21). Data can be drawn from intervention-specific engagements from direct partners, industry associations, regulators, policy-makers, academic research or other donors' or programmes' reports and studies, as well as from global data sources.

Data collection methods should be determined by the type of question that is being asked (see Step 2). However, normally outcome and impact analysis should contain some quantitative measurement of the target group. In most cases, qualitative methods can be used to complement quantitative methods, either to define what to measure and/or to understand why the quantitative data reveal a particular trend. This helps triangulate evidence of impact and tests the plausibility of intervention pathways as laid out in the programme ToC and/or the nested project results chains. Qualitative methods also provide a mechanism for identifying unforeseen results by addressing why something did or did not happen. This is especially important when impacts are dependent on complex pathways of change.[65] Examples of common FSD data sources mapped against qualitative and quantitative, as well as primary and secondary, sources are illustrated in the table below.

Table 21 Data collection methods

	Qualitative	Quantitative
Primary *(collected or funded by FSD programmes)*	– Formal or informal interviews with key informants, market players and partners – FGDs – Case studies (can also draw on quantitative data sources) – FSD staff's professional experiences, educated judgements and opinions – Observations from the field (i.e. discussions with financial services consumers and field staff of FSPs)	– Surveys.[66] These include FinScope/ FinAccess and market studies (which can also draw on qualitative data) – Geospatial data on financial access[67]
Secondary	– Minutes of meetings – Memorandums – Policies/ laws enacted – Observations/ specific analysis in reports (e.g. central bank annual reports/ banks' annual reports, special studies, FSAP studies) – Press releases (can also draw on quantitative data) – Studies by FSD network	Data in annual reports of central bank and other regulators, FSAP, special studies, global databases

Other data sources for measuring supply- and demand-side aspects of financial inclusion can be found on CGAP's website: http://www.cgap.org/blog/10-useful-data-sources-measuring-financial-inclusion and http://www.cgap.org/blog/making-sense-financial-inclusion-data-sources.

5.2.3 Sources of data and indicators – systemic change and sector tracking

5.2.3.1 Systemic change

To illustrate this point, Table 22 takes a specific example related to systemic change and provides examples of the types of data sources that could be used for the micro-insurance example presented earlier in this step. Annex D provides a more comprehensive list of data sources for all types of projects.

65. ITAD (2012).
66. These can also collect qualitative data, to analyse and present in a quantitative form. An example of converting qualitative data into quantitative indicators is the Economist Intelligence Unit (EIU) report on the national regulatory environment and institutional support in the provision of financial products and services to low-income populations. See http://www.eiu.com/public/topical_report.aspx?campaignid=microscope2014
67. These may sometimes be from secondary sources.

Table 22 Data sources: Illustrated example for a micro project and systemic change

Categories of change	Change in attitude, skills, knowledge and behaviour of partner	Market change occurred (often presented as an output in a FSD logframe)	Changes within the partner beyond the initial project	Broader market changes
Detailed measurement question	Has the micro-insurance provider (partner) improved its capacity to serve low-income people?	Has an innovative product for the micro-insurance sector targeting poor people been established?	Will the micro-insurance product continue to be provided to low-income people by the FSD programme partner in the absence of subsidies?	Has there been an increase in the number of providers of micro-insurance products in the last two years? How has the FSD partner and those influenced by the partner contributed to this change?
Indicator/ change of interest	– Partner's attendance at a training session specific to the topic – Partner conducts a diagnosis of the sector – Partner conducts market research/ segmentation on low-income households – Partner designs suitable micro-insurance product – Partner takes out a licence to provide micro-insurance products – Training budget spent on micro-insurance – Partner sets up new department to focus on micro-insurance – Strategy is developed/ adopted for low-income households – Number of staff (in partner organisation) trained and certified in pro-poor product development – Board approval of strategies/ business plan with micro-insurance elements	– Partner designs suitable micro-insurance product – Partner takes out a licence to provide micro-insurance products – Partner sets up distribution channels (branches, agents, arrangements with mobile phone operators) to deliver insurance product(s) to customers – New product has reached market (increase in number of low-income households/ clients reached/ served; increase in number of insurance policies)	– Partner's business model is viable (i.e. likely to make money) [adopt/ scale] – Increase in number of insurance policies [scale] – Increasing number of policy renewals as % of total policies in a partner [adopt] – Partner adapts product to respond to demand [adapt] – Partner commits their own funds to scale up [adopt] – Partner continues to offer the product two to four years after pilot completion [adopt]	– Partner's market share increases [scale across the market] – Number of additional insurance providers serving micro-insurance market [scale] – Total uptake of micro-insurance across the market [scale across the market] – New types of micro-insurance products available (health, agriculture, insurance as a part of product package[68] etc.) [crowding in] – Decrease in the average market price for the insurance product [respond/ scaling across the market] Changing focus on micro-insurance amongst regulators [respond]
	QUANTITATIVE	QUALITATIVE	QUANTITATIVE	QUALITATIVE	QUANTITATIVE	QUALITATIVE	QUANTITATIVE	QUALITATIVE
Data sources	FSD partner-level information: – Financial reports – Human resources (HR) data/ management reports from partner	Interviews with senior partner executives HR data/ management reports Partner annual reports and press releases	Market surveys Information from regulator Quarterly/ annual reports from partners Reports from industry association	Partner announcements Interviews with senior partner executives Information from industry association/ regulator	Partner-level information: client data analysis, annual reports, financial reports, business plan, management information system (MIS) data FSD survey of partners Market assessments Regulator information FinScope/ Findex	Interviews with senior partner executives/ industry association/ regulators/ competitors	Regulator reports Industry association reports FinScope/ FinAccess Market research/ product scan reports Findex annual reports of new market entrants	Client satisfaction interviews/ FGDs Qualitative research with partners and other micro-insurance providers Annual reports of new market entrants

68. Insurance sold as part of another purchase by customer.

Discussion point: During the consultation FSDs had different views on the time dimensions for examining these different types of indicators. For example, crowding in of other market actors could happen within a few months or a number of years. A results chain for an intervention can help provide some indication to an FSD of likely timing, but they may often be unpredictable, with both on-going project and top-down market tracking important to pick up key changes.

The above example sets out a number of indicators that can be monitored to assess if an FSD programme is contributing to systemic change. As noted above, indicators can be identified as part of setting out an intervention results chain and then tracked accordingly, using the data sources shown in the table, together with other indicators and data sources that FSD programmes may identify or prioritise.

However, there may be times when an FSD wants to go beyond this type of monitoring and use more in-depth techniques for tracking and assessing systemic change. Examples of such a case are set out in Table 23. Some of these tools go further than assessing what has changed – also looking at what has caused the changes to these underlying dynamics (discussed further in Step 5).

Table 23 Data collection methods for capturing systemic changes (qualitative insights and quantitative data)

Type of systemic measurement tool	Summary	Application
Monitoring beyond indicators (picking up narratives/ external stakeholders' insights)	– Pro-actively looking for, enquiring about and capturing observations in back-to-offlce reports, evidence of outcomes and the quality of outputs during FSD field visits, or those of colleagues and consultants (see Annex F) – Having sensitive antennae for this type of information remotely – keeping a log[69] of relevant references FSDs find in media, correspondence etc. – Convincing people in the field – local delivery partners, targeted institutions, etc. – of the value of looking out for changes that may be traced to the intervention. This could be in the form of logs/diaries or through regular processes of group reflection (amongst FSD staff and/or with implementation partners), especially while reviewing specific projects/ programmes[70]	– Leveraging tacit knowledge of FSD programme staff and perceptions of others – Picking up unexpected changes – Identifying changes that are not easily defined by indicators – Confirming that indicators that are being regularly tracked will help in tracking progress (i.e. the right indicators are being tracked)
Most significant change	– Unprompted, FSD partners choose most significant change caused by intervention (i.e. not necessarily what results chain said) – Highly participatory – FSD partners provide narrative and feedback on which stories/ changes the partners feel are most important	– Bringing in a range of market actors' perspectives – Pick up unanticipated change
Outcome harvesting	– Works backwards after change in outcomes has occurred – Use range of common tools to identify changes – Places project's contribution in context with other contributions	– Can be used to assess the causes of change – When causes of change are unclear
Outcome mapping	– Focus on behavioural and attitude change on the part of FSD partners – Use programme journals to capture behavioural change	– Picks up FSD direct partner perspectives (less useful beyond these) – Focuses on attitudes and behaviour – Used as ongoing monitoring

Source: adapted from SEEP (2015)

69. It is important to do this systematically so information can be readily retrieved, in order to assess progress towards systemic change, sector tracking and impact targets.
70. It is important that this knowledge be noted down as soon as possible after you have obtained it, to capture the flavour of the points being made and to minimise the risk of memory lapses. Simple diary entries or action logs can be used for this – see Hovland (2007) for guidance on this. Over time, these can be captured and analysed in specific FSD reports and studies.

Discussion point: It was argued by those working with FSDs that participatory methods can connect FSDs with market actors (including households) which is particularly important in re-examining our mental models of the poor and how they use and value financial services.

Tip: Check with other FSD network colleagues. They may have already identified similar information requirements, and how best to address these.

5.2.3.2 Sector tracking sources

Table 24 below outlines illustrative data sources at the programme level – financial inclusion and financial sector, as well as the two different levels of impact, poverty and growth. FinScope/ FinAcccess – perhaps the most common source – is discussed further in Box 21. A new FMT/ Centre for Financial Regulation and Inclusion (Cenfri) programme being set up in South Africa is summarised in Box 22.

Table 24 Data sources – programme: financial sector, poverty and growth (examples)

	PROGRAMME OUTCOME		IMPACT	
	Financial inclusion	**Financial sector development**	**Livelihood/ poverty reduction**	**Pro-poor growth**
	FinScope/ FinAccess (representative at national and, in some cases, at sub-national levels) Other local surveys Global Findex (national level) World Bank/ IFC Enterprise Survey (for SME credit) FinScope/ FinAccess – perception data Central bank, and other regulators supply-side data Trade association supply-side data (banks, insurers and microfinance providers) IMF dataset for cross-country comparisons GIS mapping of access points	Central bank data World Bank/ IMF IMF dataset for cross-country comparisons IFC Doing Business Index Perceptions surveys using SurveyMonkey or similar tools Financial sector development indicators shown in Table 5 are mostly available from Central Banks. Also from FSAP and World Bank	PPI results National/ regional poverty surveys National household budget and living standard surveys Census data FinScope can provide some information on proxy livelihood factors (e.g. LSM and PPI modules) Note – research underway in Kenya and Zambia may demonstrate the feasibility of developing correlations between FinScope and consumption module data in household budget surveys	Country statistics World Bank/ IMF/ GFS Labour statistics MSME employment surveys
Qualitative	Financial diaries, FGDs and in-depth interviews as part, for instance, of financial landscape studies – largely qualitative	Interviews with key policy-makers, regulators, FSP executives and civil society organisations	FGDs and in-depth interviews at household level	FGDs and in-depth interviews with entrepreneurs

Box 21 FinScope/ FinAccess – status, use and challenges

FinScope is a survey of individuals in a country that looks at demand for financial services and barriers to access (specifically referred to as FinAccess in Kenya and Access to Financial Services Survey in Nigeria, but the generic name is used here). FinScope provides an overall understanding of how individuals generate an income and how they manage their financial lives. It identifies the factors that drive financial behaviour and those that prevent individuals from using financial products and services. It also includes some psychographic questions looking, for instance, at people's attitudes, and issues such as trust. Implementing the FinScope survey over time provides the opportunity to assess whether, and how, a country's situation is changing. FinScope is designed to be at least nationally representative and, in several countries, the sample is enlarged to be representative at a regional (or lower administrative) level as well.

FinScope has been used in South Africa, where it was first developed and deployed by FMT in 2002, and in all the countries where FSDs have now been established, as well as several others in Africa (such as Ghana) and outside the continent (e.g. Myanmar and India). For all the FSD countries, FinScope has become an important yardstick for measuring financial access and changes over time. It is used as such not only by the FSD programmes, but by governments, central banks and the private sector, as well as by academics and other researchers.

FinScope has been tailored primarily to meet national requirements; cross-border comparisons requiring standardisation of at least a number of core questions have been of secondary importance. This is the main limitation when comparing FinScope with the World Bank's Global Findex, another household survey that looks at demand for financial services and barriers to access.

In 2011 the four FSDs (operational at that time) commissioned a study by OPM to take stock of where the various FinScope surveys stood, along with the strengths, weaknesses and challenges FSDs and other stakeholders faced in applying this tool more effectively. The FSDs recognised that, in most cases, the questionnaires had become too long (often in an attempt to meet a wide range of increasing demands from different stakeholders), the analytical framework linking specific indicators to relevant questions was not as strong as it should be, there was a lack of clear definitions, and there were conflicting demands between standardisation and customisation needs.

The FSDs have subsequently absorbed many of the study's recommendations in the recent rounds of FinScope surveys. For instance, questionnaires are shorter and are developed around much tighter and more coherent analytical frameworks. However, there is still work to be done, for instance on agreeing a common set of core questions and thereby balancing local customisation with cross-border standardisation. Also, to some extent FinScope still risks being a victim of its own success: too many stakeholders still want it to provide an increasing range of information and analyses. It will be up to the FSDs in each country as to how they address this. Ensuring clarity on FinScope's objectives in each FSD country remains an important issue, along with whether these objectives should be the same across all FSDs. However, lengthening questionnaires beyond those that take much more than an hour and a quarter to administer is unlikely to be the way forward. Developing complementary (and often qualitative) research tools would probably be more fruitful.

Box 22 New data programme being established by Cenfri and FMT

Cenfri and FMT are jointly establishing a new programme funded by the Bill & Melinda Gates Foundation and the MasterCard Foundation. The programme has secured funding for five years to achieve two broad objectives:

- improve the quality, relevance and comparability of indicators of financial inclusion and the data needed to design effective programmes, products and policies; and
- increase the use and quality of client-centric data, research and methodologies by FSPs to inform business decisions in a way that will lead to the design of a greater range of more relevant and impactful financial products and services for financially underserved individuals. This will largely include developing data solutions for the private sector to promote financial inclusion.

5.2.3.3 Beyond indicator sources/methods

Some of the narrative descriptions of change discussed in Step 3 are challenging to track and there is no single method that is applicable to all cases. At one end of the spectrum FSD staff can provide a narrative, from quarter to quarter, about what changes they are seeing in different characteristics of the system; for example, has a new player moved into the market, or has the governor of the central bank focused more on financial inclusion than previously? **This tacit and explicit knowledge that FSD programme staff gain from their day to day work, which goes beyond their immediate projects, is important and needs to be captured.** This places a greater emphasis on participatory data collection methods in an attempt to observe (from as many perspectives as realistically possible) the types of changes in the behaviours of policy-makers, market actors and customers that are likely to shape future outcomes. It will also require FSD staff and their partners to pro-actively look for, or have antennae to capture, interesting changes in the market, as well as to look out for changes that may be traced to the intervention (see Annex F for an example template that could be of help in this regard).

At the other end of the spectrum specific monitoring methods can be used to assess if items of interest have changed. For example:

- surveys of perceptions of the regulatory environment (e.g. using SurveyMonkey);
- FGDs with important policy-makers;
- network analysis of core actors; and
- announcements in the press and other media tracked by FSD programmes.

A separate paper on research methods provides further guidance in this area.

5.2.4 Data issues for FSDs

The types of indicators and sources that are analysed above give rise to some implications for FSDs' measurement plans, including:

- **Measuring sustainability:** it is necessary to be able to rely on a few data sources for a longer time-frame, rather than on many data sources over a short period. This can mean that there is a need to build into FSDs' partnership agreements conditions that there should be progress reporting to the FSD during, and even after, the end of the direct engagement (in cases where a partner might suspend reporting on progress, it may be possible to build in audit requirements, as well as including sanctions, such as shutting off funding or refusing any future funding).
- **Discussions with market actors:** Some actors, particularly private sector players, are less interested in results such as replication and demonstration, so FSDs will need to look for other data to report these.
- **Leveraging tacit/ informal knowledge:** FSD staff and their relationships are an important data source; they should be considered as core data sources.
- **Using existing supply-side data:** Some data, particularly regarding sector trends, can be leveraged from existing sources. There have been some discussions as to whether FSDA should facilitate/ host a financial sector dashboard for FSD countries.
- **Using a market development approach to develop data sources:** where gaps exist FSDs can help market actors to provide data on a sustainable basis (Box 23).

Box 23 The role of FSDs as data advocates and supporting data initiatives

The combination of good data and the right champions can move markets, and FSDs as market facilitators can play a crucial role in bringing about this combination. Even though the primary driver for this need not be impact-oriented monitoring for FSDs, overlapping interests with other stakeholders can be easily identified. FSDs can encourage, say, a central bank or credit bureau to provide quality data to industry providers if/when there is sound analysis that such data can help inform and/or provide incentives to financial providers through, for example, healthy competition. The starting point for this is to identify which stakeholders will be interested in which data and how it will be useful for them. Once the potential demand and sources of data are identified, challenges around skills, costs, aggregation, confidentiality, prioritisation and dissemination need to be addressed.

FSDs can also provide funding or direct technical support to, for instance, banks and/or industry associations in relation to collecting, analysing and disseminating market data. **Direct funding by FSDs of new data sources should be carefully assessed in terms of sustainability (who pays and for how long) and incentives (why?).** FSDs should carefully assess the purpose of the information dissemination: e.g. does it provide one-off data to kick-start a dialogue and collaboration, does it drive strategic clarity, or are long-term data deemed necessary for market actors and policy-makers? Similarly, there may be a lot of data that are already in the public domain, are not easily usable (e.g. the prices charged may be available for each financial institution) but need further work (e.g. cost for a typical bundle of services or a pricing index) before they can provide useful insights for consumers and other potential users. For example, in South Africa an inflation-adjusted index of the cost-to-user of banking services for low-income users, averaged over the cheapest offerings, declined from 100 in 2010 to 80.6 in 2012. See http://www.afi-global.org/library/publications/use-financial-inclusion-data-country-.case-study-south-africa.

5.2.5 Data quality

It is important to be aware of the strengths and limitations of each of the demand- and supply-side data sources, depending respectively on the evaluation questions and indicators. Demand-side data (data that originates from the users of financial services – individuals, households and enterprises) vary in their indicators, the frequency of data collection, the sampling approach and the extent to which data are representative at the national or sub-national level. Supply-side data (data that originates from FSPs) similarly can vary in their frequency of collection and sub-national representation.[72]

Data quality issues can be mitigated by following the recommended tips in Table 25. However, given the resources and availability of data, FSDs may simply have to find ways to work around these issues. Regarding issues of accuracy/validity, reliability, integrity and completeness – we cannot always be sure that we can trust what the data tell us. It is important that when interpreting the data these limitations are taken into account. It is also critical that when reporting their findings, the FSD acknowledges the possible limitations of the data and the findings and, if possible, the implications. For example, surveys can be too long and this may risk compromising the validity of answers in later questions (both respondents and questioners may become bored), or sampling may go wrong. With regard to issues of precision and timeliness, we have to work around insufficient data, unavailability of some desired data points or data not being available at the time it is needed. These issues are usually mitigated by proper planning and, to the extent possible, by diversifying data sources.

72. Note that the demand-side and supply-side data definition depends upon the specific context, e.g. for credit bureau, usage of credit bureau services by financial institutions is demand-side data whereas financial institution records of the number of loans given is supply-side data as regards an expansion of credit programme.

Table 25 Dimensions of data quality

Dimension	Definition	Tips
Good planning	Analytical plans are the starting point for all research	– Develop an analytical plan to show how data will be analysed, used and reported (responding to research questions) before designing data collection tools to ensure that all necessary data are collected
Accuracy/ validity	Valid data are considered to be accurate. They measure what is intended to be measured	– FSDs can support data collection by market actors to fill measurement gaps (see Box 23) – Data quality checks built into data collection and analysis processes – For demand-side data, sampling and survey methodologies follow industry standards and, typically, are based on master sample frames prepared by national statistics offlces – If using electronic data capture, review and test the scripting thoroughly to ensure skip routines and consistency checks are automated correctly
Reliability	The data are measured and collected consistently (i.e. in the same way and using the same data collection instruments) over time	– Ensure adequate training for enumerators and others in charge of collecting data – Put in place proper documentation and checks to ensure that data are measured and collected in the same way, regardless of who carries out the measurement and collection
Completeness	An information system provides the complete list of data sources and organisations	– Collaborate with partner organisations that can contribute to ensuring the completeness of the data – Review the data for completeness with key informants and stakeholders – Complement quantitative surveys with qualitative research, such as FGDs, as considered necessary
Precision	The data have sufflcient detail (e.g. are collected by gender, urban/ rural location, etc.)	– This should be an output from the planning process (see 'Good planning' above)
Timeliness	Data are up to date (current) and data are available on time	– Plan data collection based on when the data are needed. Allow for adequate time for training of relevant staff, data collection, data entry, data cleaning etc.
Integrity	The data are protected from deliberate bias or manipulation for political or personal reasons	hould be held to the same standards of integrity as the FSD programme – Data practices should follow industry standards, be documented and made transparent – Oversight can be given by external experts and others. If appropriate, form a politically neutral technical team to oversee the data process
Confidentiality	Clients are assured that their data will be maintained according to national/ international standards for data gathering and management	– Enumerators are trained to explain clearly to survey respondents that their names and information provided will be kept confidential. If necessary include in client contracts provisions relating to maintaining confidentiality – Also make data handling policy transparent, possibly also including this policy in the client contract

Source: Adapted from Duvendack (2013)

Box 24: Step 4 checklist

- Are you using mixed methods to take advantage of the strengths of both quantitative and qualitative methods in answering measurement questions?
- Review all data sources listed under indicator profiles and cluster these to identify the most important data sources and data gaps
- Identify data sources for selected IOM indicators, e.g. progress indicators; market system development; top down sector tracking; and 'beyond indicators'
- Assess data quality using standard criteria; FSDs should, in particular, be aware of the known strengths and limitations of demand – and supply-side data.
- Review data sources for frequency and time lag of data availability (i.e. to assess if the data will be available in time for reporting), and whether the data are already in the public domain
- Once data gaps have been identified, consider what additional data collection methods and analyses will be needed for indicator tracking
- Periodically review and refresh selected indicators and data sources
- Be data advocates — provide support to other organisations in relation to collecting, analysing and disseminating data, and in relation to doing so more effectively. Consider what is the rationale for the FSD directly collecting these data, as opposed to supporting a national stakeholder to do so?

IOM – Chapter 5: Measuring Change – Why it happened (Stage 2b)

Stage 1: Clarity of purpose

Step 1: Setting out an evalaution Programme ToC

Step 2: Developing impact measurement questions

Stage 2a: Measuring change – what happened?

Step 3: Developing indicators

Step 4: Data collection methods and sources

Stage 2b: Measuring change – why it happened?

Step 5: Assessing causality and contribution

Step 6: The research agenda

Stage 3: Bringing it all together

Step 7: Developing a credible narrative

Implementing the IOM (Chapter 7)

Chapter 5, Measuring change, covers Stage 2 of the process of implementing the IOM guidance: This stage is split into two sub Stages –2a, and 2b.

This section focuses on Stage 2b, providing guidance to FSDs on assessing as to why the changes they are observing (see Stage 2a) have occurred, and to what extent FSD programmes have contributed to these changes.

Stage 2b is broken into two steps:

Step 5 – Assessing causality and contribution: This provides the tools for an FSD to interrogate and build an evidence base for how and why changes have occurred.

Step 6 – The research agenda: This outlines certain activities exploring causal relationships in the financial sector that are likely to be beyond an FSD's core measurement system, and may also require partnerships with other FSDs, and global institutions.

5.3 Assessing causality (Step 5)

5.3.1 Overview

- The previous sections have focused on *what happened*. But for impact evaluation a description of what has happened is not enough. There is a need to explain *how and why* the changes have happened, and to what extent the FSD played a causal or contributory role. Step 5 therefore focuses on how to build an FSD's evidence base for causality – that is, exploring the mechanisms by which FSD interventions affect change.
- Causality can be established by assessing the linkages between FSD interventions and the observed change (bottom-up) and/or by assessing other pathways linking changes in the financial sector to a range of influencing factors (top-down).
- This section discusses different approaches to demonstrating causality. Methods range in terms of the rigour and investment required, internal and external validity, use of qualitative and quantitative methods, etc.
- Causality methods available to FSDs range from using existing results chains and FSD monitoring data to carrying out additional stand-alone studies. These methods do not have to be undertaken for every intervention and/or for all impact pathways but they do provide an opportunity for an FSD to step back and assess the level of their contribution.
- This paper does not contain in-depth information about how to use each of these methods, but the information provided should be sufficient for FSDs to understand their options and to make an informed decision as to a) the direction they want to pursue for measuring causality within the programme and projects; and b) where external help may be useful.
- A separate technical note is provided for further information on specific methods.

5.3.2 Causality in FSD programmes and projects

There are a number of challenges in assessing causality for an FSD programme, and its various interventions. At the programme level, FSDs are attempting to influence a system that has many interdependent parts, as well as numerous non-FSD players/factors that contribute to change. As previously described (Table 2), due to their function as market facilitators, identifying linear relationships between an FSD intervention and an observed change is challenging.

It is relatively easier to focus on causality between specific links in the programme, as underpinned by an FSD's projects or a group of projects. However, there are challenges even with this. FSD programmes work through partners rather than delivering direct impacts themselves, and thus cannot take full credit for the observed change. Projects also tend to work in combination with other factors (e.g. other policies, technological changes, stakeholder behaviour etc.), and as they purposely seek to facilitate spill-overs (e.g. demonstration effects) in the market, the distinction between what the project has influenced ('treated) and what it has not ('untreated') is often not clear.

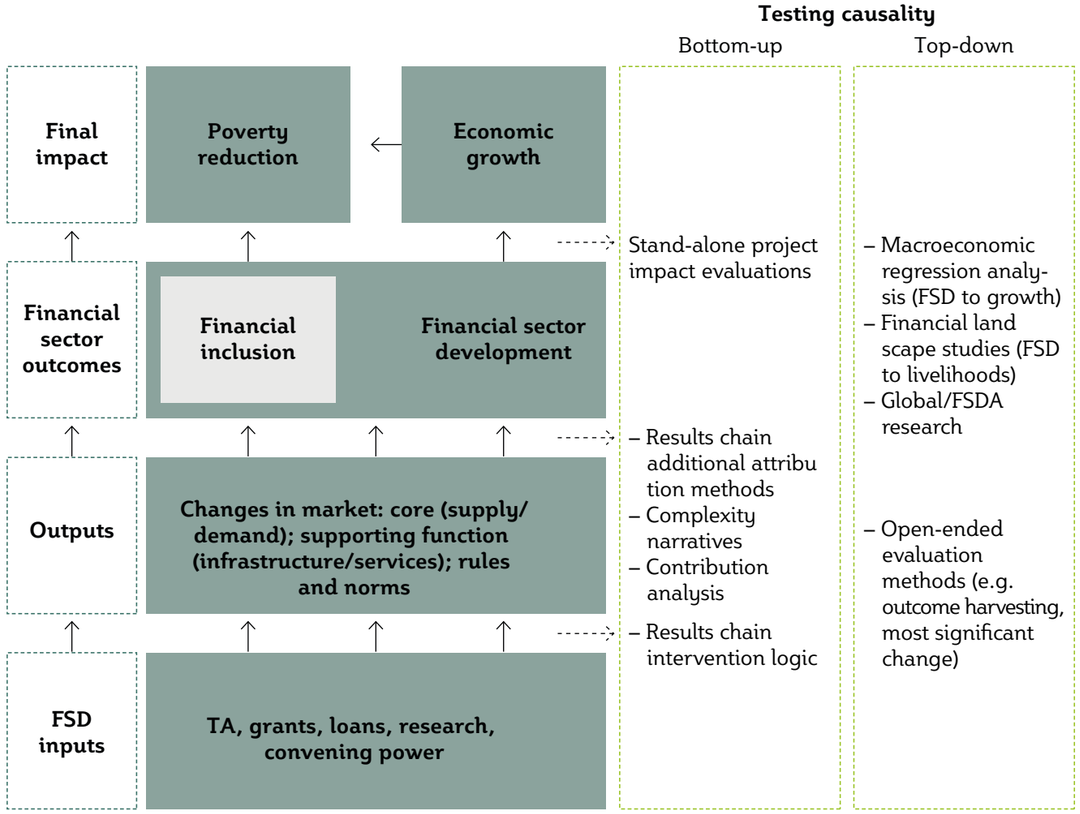
Given these difficulties, there needs to a realistic approach to assessing how far an FSD can show attribution to a particular causal pathway (see Box 3). This framework therefore adopts the DCED Standard's approach to this issue. We advocate applying pragmatic evidence-based judgement, and – in order to understand causality – applying those resources that are appropriate to the pathway being tested. The aim for the FSD should be to ***'convince a reasonable but sceptical observer'.***[73]

Measurement of causality should be driven by the **impact measurement questions (see step 2, Section 4.2),** which call for testing the cause-and-effect relationships within the ToC, i.e. not all of the causal links necessarily need to be explicitly tested. **Causality analysis will occur at different times:**

- Causality of the FSD programme as a whole should be assessed systematically during the planned impact evaluation – most likely at the end of each strategy phase, and possibly also at the mid-point of the phase (see Step 7). Taking a theory-based approach, the impact evaluation for the FSD programme will be informed by the evidence that is collected through specific assessments of causality.
- Evidence of causality for projects, and specific links in the ToC, can be captured throughout the programme implementation period. This includes assessing the causality within each of the projects the FSD supports – with light-touch methods being used regularly and more robust methods at a mid-point or at the end of the project.

73 DCED Standard

Figure 16 Analysing causality in FSDs



5.3.3 Principles of measuring causality for FSDs

The fundamental challenge of measuring causality is moving beyond simply showing change (in the market or in the lives of end-users for example) from the start to the end of the programme, and actually substantiating the extent to which the programme contributed and/or was solely responsible for these observed changes. Here we set out some general principles that FSDs can use to determine when and how to assess their interventions.[74] The following section then briefly outlines some core methods to use for specific FSD interventions and the overall FSD programme, as well as for testing pathways towards the 'higher-levels' of their programme ToC (outcomes and impact).

By taking the following steps, FSDs will be better positioned to measure causality within their overall programme (and supported projects):

- Develop a **robust causal model** that underpins the intervention, and that can be tested. In other words, is there an evaluable ToC/results chain (see Step 1) and has this been updated over the course of the programme, to ensure it remains relevant?

Tip: Facilitated discussions with informed observers about the results chain can be a powerful tool for causality assessment and often the first point of analysis (before any study or research is initiated) to check if the logic is working in practice.

- Develop a **clear implementation plan to measure causal effects** to a sufficient and appropriate level of validity. Given that not all interventions can be measured the same way, different types of validity may be considered.
- **Triangulate sources and methods.** No one source of data or even method is likely to be sufficient to establish causality for complex pathways. Triangulation allows an FSD to build up 'enough' evidence to make a plausible judgement for if, and why, an intervention(s) achieved (or did not achieve) its impact.
- **Be open to findings and look for the unexpected.** Reflecting the complexity and unpredictability of markets, FSD programme teams need to be humble and aware of their limits in terms of their understanding and influence. Therefore FSDs need to pay special attention to assessing if there are any surprising patterns in the data, or if there are unexpected factors at work.

Tip: To deal with unpredictability ask open-ended questions and seek multiple perspectives from well informed observers (e.g. market actors) or review reports/ studies commissioned by others.

- **Be open to failure.** This includes both deciding to research interventions that have failed (to understand why), as well as reporting honestly on those that have failed.[75]

Discussion point: FSDs noted that failure is part of the market facilitation process – programmes have to take risks, and therefore sometimes fail – but they struggle to report on this. It was agreed that detailing examples of failure in annual reports can often form an important source of learning, as well as credibly showing how the programme is seeking additionality, and to be taking risks.

- Ensure **quality of data collection.** Most of this has been covered above, but one additional point that can be made is that when FSDs use survey approaches to test specific pathways, there is a particular need to have a robust sampling strategy, strong research instruments, and close supervision of field personnel.[76]
- **Be transparent** about the strengths and weaknesses of methods used, including any threats to validity and any trade-offs that were made (e.g. sampling strategy, resources used, types of approach etc.)
- Allocate **resources appropriately** in order to understand the causal pathway of interest (see Box 25).

74. Creevey et al. (2010).
75. This may also include examining why certain investments were not made, or investigating the users/customers that did not choose to take-up a product, etc.
76. For help thinking about sampling sizes, see, http://www.enterprise-development.org/page/calculator.

Box 25 How to allocate resources to measure causality?

As the 2014 DFID Evaluation Policy states, 'there is potentially a boundless need for evidence to support decision-making'. **Therefore, where further evidence is needed, there is a need to prioritise.**

Some form of causality analysis should be undertaken for all pathways in the programme ToC, and to assess if individual intervention results chains are working as expected. **However, how resource intensive this analysis is will vary in terms of what methods are being used.** For example, using existing monitoring data to test a results chain for a relatively simple intervention (underpinned by a simple cause and effect theory) might be adequate. But for other cases you may need to go beyond this – from an interview to test important assumptions, to large scale surveys. These present an important opportunity for the programme to step back and reflect in more depth on how change processes have occurred.

There are a number of criteria that can be used, often in parallel, when analysing the level of resources to be devoted to exploring causal pathways. These include:

- the complexity of the causal pathway, in particular how many non-FSD contributory factors are likely to be present (the more complex the pathway, the more advisable it is to undertake analysis beyond simply monitoring the results chain);
- the importance of the pathway to the overall programme ToC;
- the size, cost and significance of the FSD intervention (some FSDs have 'flagship' projects);
- the gap in understanding regarding the causal pathway. Some pathways will be well known, with significant past experience or global evidence suggesting it will be operating (thus no need for additional FSD analysis); and
- the potential for impact (at outcome level on the financial sector, or on livelihoods).

The table below highlights the different implications for measurement of 'important' and 'less important' interventions.

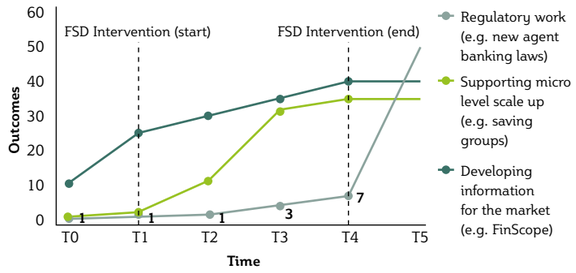
Intervention Category	Measurement Requirements
Category 1: Interventions deemed 'less important' based on an *ad hoc* criteria	Small baseline conducted before any changes have occurred Partners self-report results to validate results chain Few (if any) additional methods are used to assess causality. Methods may include rapid data collection methods, such as interviews with key actors and FSD staff.
Category 2: Interventions deemed 'more important' based on an *ad hoc* criteria	'Category 1' plus: Additional baseline data collected Pathways verified through a triangulation of data sources Causality (particularly at outcome level) measured using rigorous methods

Source: Table adapted from FSDMoç results management handbook

- **Identify the appropriate timing of measurement.** As we noted in section 5.1.6 above and illustrated in Figure 15, it is difficult to identify the most appropriate time for FSDs to undertake measurements. This will depend on the type of intervention (and pathway) in question. On the one hand, the longer the measurement period after the intervention, the more difficult it is to isolate the intervention's impact. On the other hand, take-off trajectories may be very different. Therefore, understanding pre-intervention trends (not just a static baseline) may also be important (see Box 18). As shown in Figure 17 an FSD needs to try to assess if the intervention is 'riding the crest of a wave' (and thus there is a need to be careful not to overestimate impact), or 'putting in building blocks for future' (and thus there is a risk of underestimating the overall impact).[77]

77. See Boulton and Johnson (2013).

Figure 17 Timing of impact



5.3.4 Methods for examining causality

5.3.4.1 Examining the causality of FSD interventions (bottom-up)

A bottom-up approach to examining causality focuses on developing evidence to test how an FSD's programme ToC (or results chain) is operating. Examining how FSD interventions are causing observed changes in the market and leading to outcomes related to financial sector development and financial inclusion will be the main focus of any FSD programme-level impact evaluation.

5.3.4.2 Examining causality at sectoral level (top-down approach)

The bottom-up view should be triangulated with a top-down view. This sectoral-level/ top-down perspective focuses on the more 'removed' pathways that affect the structures, dynamics and changes of the financial system, and subsequent impact on livelihoods, rather than the direct impact of FSD interventions. The 'removed' pathways, for example, could include indirect influences, or other market players or forces.

Tip: Top-down analysis is likely to be undertaken more infrequently than bottom-up analysis, and involves focusing on the impact/outcome part of the ToC.

Many of the causal methods described in this chapter (and in the technical note) can apply to both the top-down and bottom-up approaches. For a top-down approach, rather than focus on FSD inputs, the methodologies should be used to focus on the change that is of interest – for example, the changes in usage of financial services, and what causes have led to that change. For FSDs, these top-down analyses will be at a market or end-user level (Box 26). As noted earlier, some of these studies could be commissioned and coordinated by FSDA, particularly if a cross-country approach is taken. There is also global evidence, for example, of financial sector development leading to economic growth (through cross-country regression analysis) that can provide some comfort to FSDs that important pathways are present. Some of these pathways, particularly those related to economic growth and livelihoods, may form part of a broader research agenda (see Step 6).

Box 26 Top-down pathways of interest and evaluation approaches

Financial sector development and institutions: Financial sector development in one country can be compared with other peer countries (i.e. with similar structural characteristics), with the difference in performance being attributed to the types of policies and institutions (using regression analysis). Rather than compare across countries, Beck (2014b) recommends using a synthetic benchmark that compares a country in a given year to a benchmark derived from multi-dimensional cross-country comparisons. This synthetic benchmark (financial depth frontier) is determined by country variables, such as (i) the structural characteristics of the socio-economic environment in which financial institutions and markets operate and which impose a limit on their development, and (ii) long-term policy variables that either foster or limit financial deepening. The gap between the actual level of financial development and the structural depth line can be related to different policies. The structural depth line is defined as the level of financial development predicted by structural country characteristics that are not directly related to policies and/or the financial sector.

Financial sector development and economic growth. Many studies have looked at this link through large cross-country regression analysis (Beck 2014a). Individual country studies could also be undertaken using various econometric techniques, but this is likely to form part of an overall research agenda rather than being a common FSD measurement tool.

Financial sector development and livelihoods: One method that is being used by FSDK takes changes in livelihoods (tracked through quantitative and qualitative surveys) in four regions in Kenya and then focuses on the linkages between these and local financial sector development (through supply- and demand-side surveys of these local financial markets). This analysis is not primarily focused on capturing the direct impact of FSD programme but can explore questions that stem from key knowledge gaps – for example around people's financial behaviour and use of financial services to enhance (or not) their livelihoods. See FSDK website, at www.fsdkenya.org (forthcoming).

Financial inclusion and livelihoods: One option is to use existing surveys, such as FinScope/ FinAccess, that are designed to measure financial access and some other elements of financial inclusion and explore whether these could be: (a) adapted or have added to them asset- or consumption-based modules; (b) linked to national surveys, such as household budget surveys, that provide much more comprehensive asset- and consumption-based data related to poverty; or (c) some combination of these approaches. FSDK, for example, is also aiming to look at using supply-side research and data to improve insights into the influence that the financial sector might have on poverty. See additional technical note on linking FinScope/FinAccess to poverty surveys.

Longitudinal studies (with households/enterprises). This would involve tracking households and/or enterprises over time using a range of methods, to assess how their usage of financial services has changed, and how this was affected by changes in the financial sector.

5.3.5 Methodologies for measuring causality

There are a range of methodologies for demonstrating causality (Table 26). FSD programmes can select the methodologies to use based on the characteristics or combination of characteristics that best respond to the causal link to be tested and that are feasible to implement with the resources allocated. The following characteristics can help determine which approach the FSD should choose:

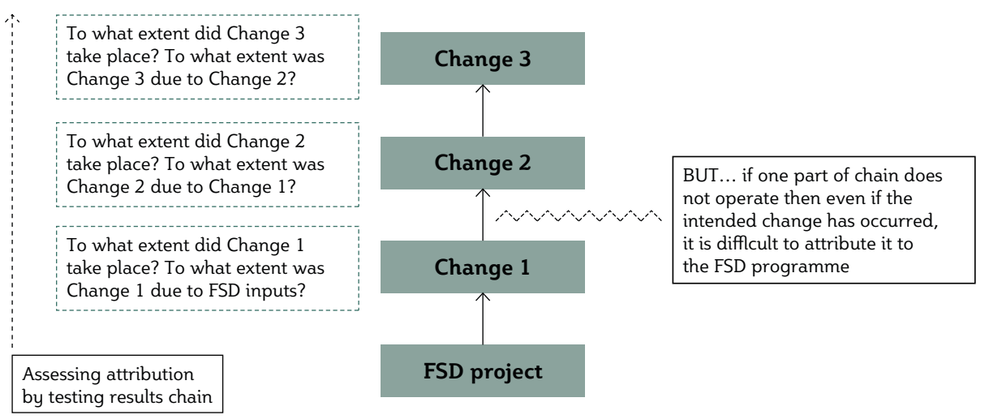
1. **Primary focus: causal demonstration** – the methodology helps collect data that support the causal link(s) articulated by the ToC; **constructing a counterfactual** – this involves striving to prove the causal link by demonstrating what would have happened in the absence of the intervention; **causal explanation** – the methodology collects information on how and why the intervention worked (or did not work) the way it did; **contextual description** – the methodology collects information not just on the intervention but on the contextual factors in which the intervention was implemented.
2. **Type of data: quantitative** approaches to measuring causality measure changes numerically and in some cases they measure the extent to which these changes are attributable to the intervention. **Qualitative** approaches capture qualitative evidence that cannot be presented simply with numbers. These approaches can capture insights into causality and explain how and why changes occurred, but they are generally not considered to be as rigorous as quantitative methods. Whichever approach is used, all types of intervention assessments are likely to rely on a mix of both qualitative and quantitative data – *there is no story without numbers and no numbers without a story. (Jim Tanburn of DCED)*

3. **Timing:** Some methods require that data collection and study design begins before the intervention is implemented (**prospective**). On the other hand, **retrospective** studies collect data after the intervention.
4. **Impact measurement questions (top-down and bottom-up):** consider what research questions you want to answer. Keep in mind that most of these methodologies can explore both top-down and bottom-up impact measurement questions. Table 27 provides a few example questions which each of the methodologies would be well-positioned to address.

A full list of methods is set out in a separate note. These are briefly summarised in Table 26 and Table 27. There are basically two main approaches that can be pursued – theory-based or quantitative (statistical) approaches. These both seek to measure the counterfactual of what would have happened without the FSD intervention, although they take different approaches to testing for causality:

Theory-based approaches: The rigour of these approaches as regards determining causality is derived from the use of quality evidence and logical thinking and testing rather than other rigorous approaches such as comparing actual results against a counterfactual. This can include **testing the results chains with monitoring data** – see Figure 18 – to provide an indication of causality. This indication can be strengthened using relatively light-touch methods to assess if the pathways linking these observed changes are still operating as predicted or as intended. This can include key informant interviews with partners, and other market actors (and other observers), examining relevant trends in sector data, and small-scale questionnaires. Effort should be made in the above methods to assess if (and which) other non-FSD factors have contributed to these changes. The key point is that even if there is a significant change, but the mechanism put in place by the FSD intervention linking these changes is not plausible, then these changes cannot be attributed to the programme.[78] Other more systematic approaches to testing the theory, such as outcome mapping/contribution analysis/most significant change, can also be used for more in-depth exploration, as explained in the additional guidance note on this subject.

Figure 18 Using a results chain for attribution



Source: DCED (2013)

Tip: It may make sense for both analytic and practical reasons to attempt to measure the causality of a group of interventions. There are likely to be points in the programme ToC which are critical for changing market systems but which occur across a series of interventions. For example, FSDs are likely to have a number of interventions focused on developing market information (through studies, research, forums etc.), and it may be possible to apply causality methods to measure how these have informed the enabling environment, and through which mechanisms (e.g. new information, change of attitude etc.).

78. More specifically, if the pathway(s) linking the intervention to the change is/are not in its/their intended state, it becomes evident that other influencing (external to the FSD intervention) factors have affected the change as well. This is not to say that the FSD intervention is no longer valid or that it can no longer be credited, but it does indicate that the intervention is nested within a broader system of influence that needs to be acknowledged in order to identify opportunities for FSD impact and improvements.

Quantitative (statistical) approaches: in all countries where FSDs operate the financial sector has been relatively undeveloped prior to an FSD's establishment (hence the rationale for their existence), and therefore we would expect the sector to expand over time, even without FSD interventions (i.e. the counterfactual). Separating this counterfactual trend requires additional analysis, beyond that provided by the monitoring data.

There are a number of potential approaches to assessing a counterfactual. Whilst the theory-based methods presented above attempt to provide some description of the counterfactual, undertaking traditional methods will require significant surveys and/or statistical expertise and thus are likely to only be used for testing particular important pathways of interest, or when conducting one-off evaluation activities on a particular project or intervention (usually done by a contracted third party). Experimental and quasi-experimental design constructs a counterfactual using quantitative data, usually survey data. Non-experimental methods, while based on actual data, establish statistically significant relationships among variables, but do not statistically prove a causal relationship. Again, fuller descriptions are provided in an additional technical note which outlines methods for undertaking causality analysis (www.fsdafrica.org/knowledge-hub).

Table 26 Approaches to examining causality with example questions

Approach	Description	Example Bottom-up Question	Example Top-down Question	Data Sources, Tools
Experimental	Statistical analysis based on a randomly assigned treatment and control group to fully attribute change to intervention.	Did the intervention cause the observed change?	Was the observed change caused by the intervention or other causal variables?	Survey data (longitudinal).
Quasi-experimental	Statistical analysis to construct a plausible counterfactual without assigning treatment and control groups. Aims to fully attribute change to intervention.	Did the intervention cause the observed change?	n/a.	Survey data, supply-side data (longitudinal, cross-section).
Quantitative non-experimental	Using data to demonstrate a change or difference among segments of a population or between/among points in time. They may show statistically significant correlations, but not causation.	Does financial depth increase after the introduction of pro-financial inclusion policies?	What are statistically significant predictors of reduction in poverty (or other outcome)?	Survey data, supply-side data (longitudinal, cross-section).
Results chain/ monitoring data/ light-touch	Triangulating project/ programme monitoring data that show change with 'light-touch' qualitative information that support a causal relationship.	Was there a change in outcomes from the beginning to the end of programme implementation? Is there evidence that these changes were because of the programme?	n/a.	Monitoring data, key informant interviews, observations, FGDs.
Case studies	A narrative that explores and explains what happened and why with regard to an intervention.	How has the agent banking model unfolded in the country since the introduction of the new agent banking regulation?	What were the factors that contributed to the success of agent banking in the country and how did they contribute?	Surveys, key informant interviews, observations, focus groups, expert panel, document review.

Approach	Description	Example Bottom-up Question	Example Top-down Question	Data Sources, Tools
Contribution analysis	Puts together evidence for why the observed results have occurred and the role played by the intervention and other internal and external factors.	To what extent did the new agent banking regulation contribute to the observed increases in financial inclusion? Why did it work (or not work)?	What were the main causes of changes in financial inclusion in the last X years? How did the different factors interact with each other to generate the observed change?	Surveys, key informant interviews, observations, focus groups, expert panel, document review.
Process tracing	Tests specific hypotheses about causal links. Traces the evolution of given cases over time within the context(s) in which they occur, documenting and explaining the processes by which, and the conditions under which, certain outcomes are obtained.	Did the creation of new village savings and loan associations lead to improved use of formal financial services and improved livelihoods? How did it work? What were the key success factors?	Did the multiple factors we believe contributed to improved use of formal financial services actually contribute in the way we think they did?	Key informant interviews, observations, focus groups, expert panel, document review.
Outcome mapping	Similar to a ToC, outcome mapping produces a 'map' of the main changes (or outcomes) that were achieved and the relationships among them.	What intermediate changes did the programme generate (for example policy change, and increased capacity of financial service providers and improved market information)?	What were all of the pre-conditions that made improved financial inclusion possible? How were those pre-conditions (outcomes) generated by the programme and others (outside of the programme)?	Key informant interviews, observations, focus groups, expert panel, document review.
Outcome harvesting	A participatory approach to identifying the main outcomes of the intervention and working backwards to understand the intervention's contribution to them.	n/a.[79] (Used to explore changes without focusing on FSD interventions.)	What were the main causes of changes in financial inclusion in the last X years?	Key informant interviews, observations, focus groups, expert panel, document review.
Most significant change	A participatory approach to documenting, from the beneficiaries' perspective, the most important outcomes of the intervention.	n/a. (Used to explore changes without focusing on FSD interventions.)	What were the most important outcomes for beneficiaries? What were the main causes of these outcomes?	Key informant interviews, observations, focus groups, expert panel, document review.
Complexity narratives	A structured approach to investigating the 'back-story' of an intervention and the contribution of the intervention.	To what extent did the intervention contribute to the observed outcomes?	What were the main outcomes and what were the causes?	Key informant interviews, observations, focus groups, expert panel, document review.

79 These can be used to explore interventions but their use for an FSD is in providing a non-FSD intervention focus.

Table 27 Summary of approaches to examining causality

Approach	Approach to causality				Timing		Level of FSD testing			Type of data used		Type of change		
	Causal demonstration	Construct a counterfactual	Causal explanation	Contextual description	Retrospective	Prospective	Individual project	Group of projects	Overall programme	Quantitative	Qualitative	Systemic	From micro intervention	From macro/meso intervention
Experimental	**XX**	**XX**				**XX**	X			**XX**			**XX**	
Quasi-experimental	**XX**	**XX**			X	X	X			**XX**			**XX**	
Quantitative non-experimental	X	X	X	X	X	X	X	X	X	X	X		X	X
Results chain/ monitoring data/ light-touch	X		X		X	X	X	X	X	X	X	X	X	**XX**
Case studies	X		**XX**	**XX**	X	X	X	X	X	X	X	**XX**	X	**XX**
Contribution analysis	**XX**		**XX**	**XX**	**XX**	X		X	X	X	X	**XX**	X	**XX**
Process tracing	**XX**		**XX**	**XX**	**XX**	X	X	X	X	X	**XX**	**XX**	X	**XX**
Outcome mapping	X		X	**XX**	X	**XX**	X	X	X	X	**XX**	**XX**	X	**XX**
Outcome harvesting	X		**XX**	X	**XX**	X	X	X	X	X	**XX**	**XX**	X	**XX**
Most significant change	X		X	X	**XX**	X	X	X	X	X	**XX**	**XX**	X	**XX**
Complexity narratives	X		**XX**	**XX**	**XX**	X	X	X	X	X	**XX**	**XX**	X	**XX**

XX indicates a relatively stronger focus

Box 27 Step 5 Checklist

- Decide what criteria to use while choosing when (and how often) to undertake deeper analysis on causality for projects/ specific causal pathways
- Not all causal analysis needs to involve complex and expensive studies
- Have you considered key bottom-up and top-down impact pathways which might need extra analysis?
- Consider the principles for applying causality methods. Are these being applied when considering the FSD's impact?
- Note how different causality methods meet different types of evidence needs; which one fits your objectives? (see technical note on methods for undertaking causality analysis)
- Consider if you need external expertise for specific studies/ research. Can project managers explain the type of causality technique being used, and transparently present how the FSD is claiming its contribution?
- Do not leave all causal analysis till the end of the programme. Taking up (at least) one or more themes/ links for causal analysis on an annual basis can: a) build internal capacity; b) help check its usefulness with funders and others; and c) strengthen analysis in the annual report

5.4 The research agenda (Step 6)

5.4.1 Overview

The research agenda is not strictly a linear IOM step like those discussed above. The consultative process also included the development of a research agenda (for FSDs, and FSDA) **that will create a better understanding of the causal relationships between certain kinds of financial sector interventions and the results or impacts that they are expected to generate.** This is also important for generating evidence for top-down measurement. This section therefore covers:

- why FSDs are involved with research;
- what research FSDs might undertake;
- how FSDs can contribute to the global research agenda on financial sector development; and
- the respective contributions of FSDs and FSDA.

5.4.2 Why FSDs are involved with research

FSDs' primary focus is on facilitating market development so that financial markets work better for poor households and small businesses. FSD programmes conduct research as well as use research conducted by others to improve their own performance. **Rooted in a practical context and seen as an independent voice (by the public and private sector), research that is conducted or facilitated by FSDs is also valued by other stakeholders within and outside the country.** FSD research efforts start by asking three basic questions: (a) What is the knowledge gap that this research is trying to address? (b) Who will potentially use this knowledge? and (c) How will they use it? FSDs' long-term engagement with FinScope research confirms that the same research can be used for multiple stakeholders and objectives. However, FinScope research also confirms that almost as much effort is needed to distil and disseminate tailored messages from the research to the appropriate stakeholders, as is needed for the primary research. In addition, FSDs have come to appreciate that many users in both public and private sectors do not fully understand how to use research to improve decision-making. It is also evident that FSDs face competing demands for research and have to find effective ways to establish their research priorities.

Discussion point: During the consultation several people questioned whether FSDs should be "doing" research or should be facilitating local researchers. As noted, FSDs have an 'independent' function that makes their research credible and their research focus on poverty reduction differs from that of private sector actors. However, where possible, FSDs have an important role in helping market actors unlock their existing data as well as improving the skills and capabilities of the private and public sectors to analyse and exploit existing data.

5.4.3 What research FSDs might undertake

The research agenda can meet very different and overlapping requirements of the FSD programmes in relation to improving their own effectiveness, as well as contributing to global knowledge and learning:

- research for FSD project/ programme evaluation;
- contributing to global knowledge and learning;
- research that informs FSD's strategy and design of specific interventions/ projects; and
- research for market facilitation and development.

Research for supporting FSD project/ programme evaluation

As described in Section 5.3 (measuring causal relationships) research can play a significant role in helping to understand the impact of the programme from a top-down perspective. **Research is another source of evidence for assessing an FSD's impact, providing in-depth exploration of specific issues, using robust methodologies.** Monitoring systems and bottom-up analysis will be insufficient to develop the contribution narrative: research can therefore help to fill specific gaps in the IOM. It can also contribute to better understanding of causal chains, enabling FSDs to improve their effectiveness and communications in the future. Given the extensive literature on the relationship between financial sector development, growth and poverty reduction **FSDs can largely rely on this to obtain comfort regarding the links between financial sector development and the final impact (or goal) of poverty reduction.** However, individual FSDs may also want to undertake impact evaluations of specific links, as per Step 5, when they deem it appropriate. As noted in Box 26 above, the use of financial landscape studies and/or the connecting of FinScope data to national poverty data are two FSD-specific approaches that are currently being explored to assess these links.

Tip: Top-down analysis does not necessarily need to take the form of in-depth research. As described in Step 3, FSDs can track the evolution of the sector from secondary data. Furthermore, while research is well suited to examining complex links in the financial sector, and within households and enterprises, it could also be used to augment bottom-up analysis – for example, through comparative research comparing how similar FSD projects (for example, with saving groups) have caused change.

Some FSDs are also supporting country-specific research: e.g. using FinScope to assess the poverty profiles of those who are financially included and excluded (see below) and assessing the poverty impacts of specific FSD interventions/ projects, such as savings groups.

Contribute to global knowledge and learning

There are areas within the input to outcome parts of the causal chain in which FSD research aimed at improving effectiveness can contribute significantly to the global research agenda. This is because even today the literature is not always very clear about the mechanisms through which financial inclusion/ financial sector development directly contributes to lower poverty and inequality. Thus, FSD research at the output and outcome stages of the causal chain can be useful in clarifying these mechanisms.

By the same token, FSDs often undertake or commission research to confirm contribution to the ToC at the outcome to impact for end-user level (at the end of the causal chain). This research could be for a variety of reasons, from supporting the 'proving' impact agenda with national stakeholders to 'improving' their understanding of the mechanisms though which outputs lead to impacts for end-users. Such work may in itself be useful for the global research agenda, or it may need only a small adjustment to fulfil that wider purpose. Box 26 provides two examples of research into the links between: (i) financial sector development and economic growth; and (ii) financial inclusion and poverty.

Research that informs FSDs' strategies and design of specific interventions/ projects

FSDs need to know what does and does not work. What are the areas of high potential impact, such as mobile money or insurance? Why are customers not opening accounts even when they are in physical proximity to FSPs? Why are they not using these accounts, despite having them? Over time it should be a core function of FSPs to segment markets, assess customer behaviour and customer take-up, and/or to design and pilot-test new financial products. However, in the early stages of the market development, FSPs may lack the skills and/or the resources to undertake high quality market research and analyse all the customer data that they have already collected but have not really analysed or applied. FSD programmes are also interested in research to design interventions/ projects at meso or macro levels (e.g. how the current arrangements for deposit insurance and credit registry are working) before considering what kind of technical and financial support they should provide.

Research for market facilitation and development

This includes data and research that FSDs produce for the market. For example, research to improve market information or to understand contexts and markets, using evidence to influence policy-makers and market actors, and move markets, such as market-wide research on levels of activity/ inactivity of mobile money agents (Box 23).

Tip: It is important to not only think about the type and quality of the research but also about the investments an FSD is willing to make in high quality communications and presentation (e.g. dashboarding, web-design) to ensure effective dissemination/ use of the research.

During the IOM consultation, it was asked whether FSDs should be 'doing' research, or if they should be facilitating local researchers. There is significant existing data and information already. However, on the other hand, FSDs are uniquely positioned to act as 'independent' researchers in the financial sector (e.g. FinScope). Therefore, each FSD's research agenda has to be context- specific, on a country-by-country basis. **Each FSD needs to decide what areas of research are required within their overall country strategy and how it should be conducted; just as they decide, for example, what kind of capacity building they need to do for financial institutions, or what projects they should fund in the area of digital finance.**

Tip: In making decisions about their research agenda, FSDs can discuss with their funders the possible benefits of the research from a global perspective, especially where two possible pieces of research have equal merit from a national point of view.

5.4.4 The role of FSDs and FSD Africa

FSDs will normally be the appropriate organisations to undertake or commission research that is specific to the countries in which they are working, but **FSDA can play an important role in at least three ways:**

- by undertaking or commissioning studies that have to cover more than one country (including cross-country studies) and focus on gaps in the measurement of FSDs' ToCs;
- alternatively, by helping to liaise with two or more FSDs that wish to work together to undertake or commission such studies themselves (e.g. on links between financial sector outcomes and poverty reduction); and
- by supporting the FSDs in knowledge management, as an information exchange as well as a disseminator of the results of research (within the FSD community and in the wider world).

It will be useful for FSDA and the FSDs to come to a specific understanding of how these roles will be fulfilled, through periodic discussion at network meetings and/or around discussions of specific research opportunities.

5.4.5 Examples of possible research areas

For illustrative purposes, some examples of possible areas for research that came up during consultation are given in Table 28 below. Each FSD will decide their own priorities based on ongoing research and the key evidence gaps that they have identified. Further discussions are needed before any of these topics are selected for joint research by the wider FSD network.

Table 28 Examples of possible research topics

Research category	Research topics
Financial sector development and impact on economic growth	– Changes in indicators for financial sector development and its potential contribution to economic growth – Net income changes of people and households in different market segments that are attributable to changes in the financial sector
Contribution of financial inclusion to poverty reduction	– Changes in demand for larger volumes of and more sophisticated financial products (primarily from businesses) and the FSPs' corresponding suite of product offerings (this could include watching how businesses grow, including whether net job creation results, and, if so, how much) – Changes in quality of access to financial services and whether – and if so, how – that contributes to the improved ability of individuals and/or households to achieve socio-economic goals (country-specific) – Longitudinal studies to track how different types of poor households are managing changes in the availability of financial services (country-specific) – Changes in financial behaviours, including household economics and resource allocation – Basic research documenting whether what is going on in specific products or market segment (e.g. M-Shwari) also has value for a global audience (although there may be less value in FSDs doing this)
Improving the take-up of financial services	– To what extent and how does financial inclusion help poor people live the lives that they value? – How payments, mobile usage, savings and wage payments can facilitate credit – Why people do not use accounts, or use them in very limited ways (analysis of account inactivity and drop outs) – Why FSPs do not undertake big data mining to understand and nudge customer behaviour – Issues around disclosure of information and how these alter customer behaviour – Research into how FSPs apply research and other information to develop their own products – How can remittance payments be used as collateral for credit – Why consumers do not activate insurance policies that are bundled with seed purchases
Improving the effectiveness of FSDs	– Changes in the financial sector that are attributable to FSD programmes – Level and quality of adaptability of FSDs to (a) changes in market conditions, and (b) evidence of flaws in the ToC – What causes some change management processes to succeed and others to fail? – The merits of different ways of delivering financial sector training – Changes in perceptions of the role of the FSD as a key market facilitator, and the level of its contribution – Procedures, frameworks and tools in place that facilitate real-time learning to ensure that FSD programming remains relevant and on target to contribute to its desired impact

Box 28 Step 6 checklist

- FSD programme-led research may have multiple uses. It is useful to confirm 'what is the knowledge gap that this research is trying to address, who will primarily use this knowledge, and do they have the capacity to use it effectively?' before initiating any research
- FSDs should set out which causal links in their ToC they want to explore with in-depth research, and if this will be undertaken by the FSD, or rely on global research
- With growing demand for research and evaluation, FSD programme should develop clear criteria for prioritising research and evaluation efforts
- Is it clear who has the responsibility for knowledge management, learning and communications within the FSD programme and how they can leverage this for its core function of market facilitation?
- Do you have a clear communication strategy for using research as a market facilitator?
- It may be useful to set aside a budget for research, evaluation and learning in each theme/ project, and across the overall programme
- Different stakeholders have very different needs and may find short tailored notes more useful than lengthy research reports. Leverage the FSD's own website and other channels (other websites such as a central bank's, and industry events) to disseminate key research findings

IOM – Chapter 6: Bringing it all together (Stage 3)

Stage 1: Clarity of purpose

Step 1: Setting out an evalaution Programme ToC

Step 2: Developing impact measurement questions

Stage 2a: Measuring change – what happened?

Step 3: Developing indicators

Step 4: Data collection methods and sources

Stage 2b: Measuring change – why it happened?

Step 5: Assessing causality and contribution

Step 6: The research agenda

Stage 3: Bringing it all together

Step 7: Developing a credible narrative

Implementing the IOM (Chapter 7)

Chapter 6, Bringing it all together, covers Stage 3 of the process of implementing the IOM guidance: bringing together the evidence that the IOM system has generated in order to answer programme-level impact evaluation measurement questions. This stage takes place primarily at the end of the FSD's strategy or at particular strategic points such as annual and mid-strategy reviews.

Stage 3 incorporates the final step in the process:

Step 7 – Developing a credible narrative: Guidance is provided on how FSDs can analyse evidence and tell a robust story of their programme's impact.

6.1 Developing a credible narrative (Step 7)

6.1.1 Overview

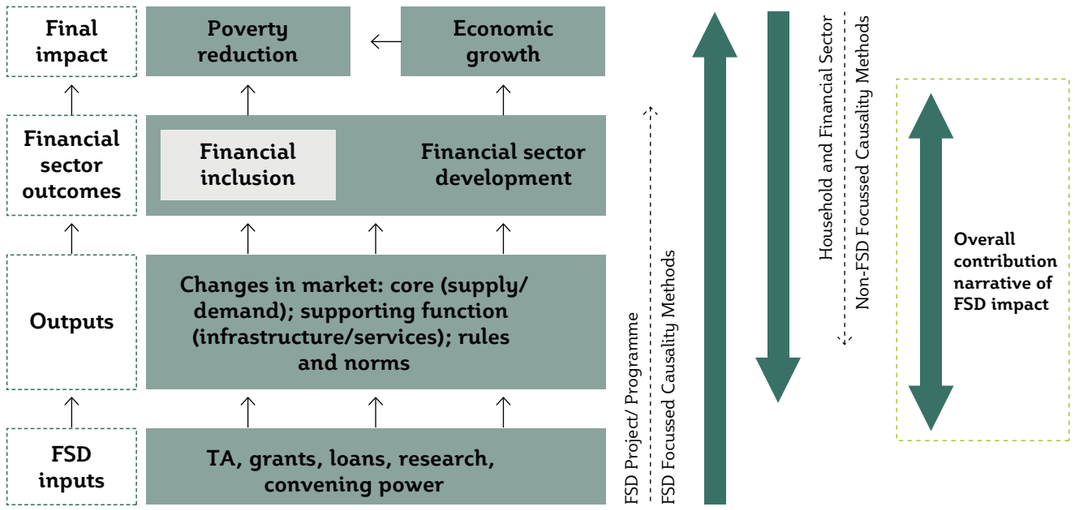
- This section provides an overall approach for bringing together the evidence that the IOM system has generated in order to answer programme-level impact evaluation measurement questions prioritised by an FSD programme and its funders.
- There are different points in an FSD programme cycle at which an FSD seeks to develop a credible narrative to assess and report on its programme impact.
- Considering the importance the IOM places on measuring systemic change, FSD programmes should try to report on this explicitly when building up a credible narrative.
- Contribution analysis, a method for testing the ToC, is presented as an option for evaluating the overall programme's impact at these reflection/ review points at different points in a programme cycle.
- A number of steps are set out for undertaking contribution analysis, many of which rely on processes from earlier steps in the IOM guidance.
- This step also provides guidance on aggregating results, including for VfM analysis.

6.1.2 Contribution analysis of an overall programme impact

The objective at this stage is to bring together all the bottom-up and top-down evidence to articulate a narrative (see Figure 19) that would ***'convince a reasonable but sceptical observer'*** **about an FSD's contribution to a specific change being measured (e.g. access to financial services).** At this stage an FSD will analyse all the data and information collected to answer the impact measurement questions set out at the beginning of the process. In particular, it will also attempt to draw together the impact logic from changes observed in the financial sector, and the programme's contribution to these changes. The approach to answering the programme-level impact measurement questions is a TBE, and, more specifically, a contribution analysis.

Bringing together a credible narrative can occur at different times and can take place at different levels of robustness, ranging from an annual light-touch review of the impact evidence, alongside the updating of the programme logframe and annual report, to full testing and verification by an independent evaluator at the end (and perhaps also mid-point) of the strategy period.[81]

Figure 19 Developing a credible narrative



81. Of course, as evidence emerges as to the effectiveness of interventions, or changes in market trends, FSD monthly and quarterly meetings should identify what analysis needs to be undertaken.

As noted in the figure above, contribution analysis incorporates a number of steps which are set out here, most of which have already been covered above. However, we summarise these here to help FSD programmes visualise what the process entails.

6.1.2.1 Develop a ToC/ results chain (see Step 1)

The main risk to note when using this ToC for an overall impact evaluation conclusion is the risk of bias in terms of how the ToC is tested. The risk is two-fold: first, there is the risk of self-attribution bias, where there is too much focus on gathering evidence to confirm or refute the anticipated casual mechanisms, without exploring alternative theories. To some extent, the top-down approach helps mitigate this risk by recognising what has happened more broadly and the contributions of other factors. But a second risk is that even if we do look for alternative theories, and confirm or refute our ToC, the whole evaluation is still shaped by our original logic. This problem can be mitigated to some extent by including an independent evaluator (see the next chapter), but there also needs to be a systematic search for potential alternative explanations throughout the evidence gathering process. This means that FSDs in their M&E processes need to be continually open to, and honest about, the various non-FSD contributions to impacts, and they need to document these throughout.

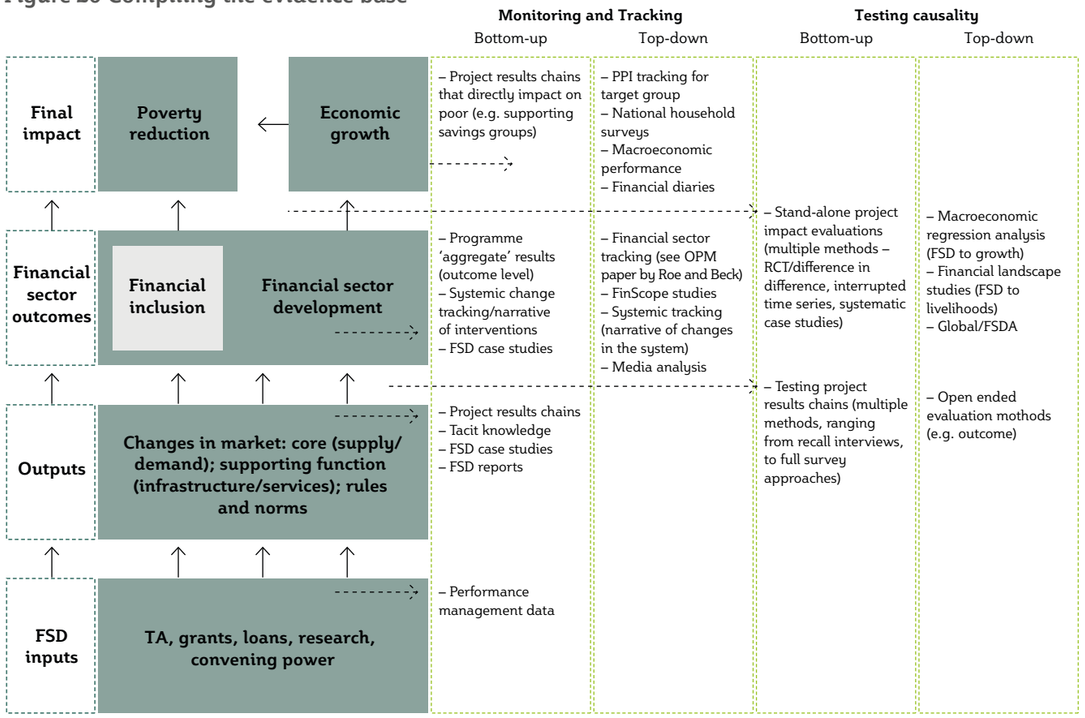
6.1.2.2 Set out the impact measurement questions to be addressed (see Step 2)

The main impact measurement questions identify the likely areas of enquiry for an FSD, as well as directing the way the credible narrative is presented. The evaluation questions should also be supplemented with a synthesis question to ensure cross-checking of results: **'given the observed changes in the financial sector (e.g. established through top-down approaches), what were the contribution of FSD interventions?'** Some interventions, particularly those at the micro level, might be able to show impacts right through the results chain (i.e. from FSD inputs to livelihood changes), but most cases will require some bottom-up and top-down triangulation.

6.1.2.3 Gather the existing evidence on the ToC (see Steps 3–6)

As noted earlier, **in most cases FSD programmes should document robust evidence of contribution (rather than attribution).** The overall evidence base that is used to conduct this contribution analysis is built up from FSDs' monitoring and causality analysis over time, as set out in Figure 20.

Figure 20 Compiling the evidence base



6.1.2.4 Assemble and assess the contribution story/ credible narrative

This involves determining whether the ToC and the assumptions underlying it hold true in the light of evidence, and whether it is reasonable to conclude that the activities of the programme have contributed to the impacts of interest, and in what way. Such an assessment will include developing a step-by-step chain of arguments, backed by evidence, asserting that the programme intervention(s) have (or have not) made a contribution to intended impacts, and to assess the strength of the evidence for that contribution. Different sources of evidence are likely to have different strengths in this context, depending on where in the ToC they are applied, and what data sources and methods the evidence relied upon. In part, this comparative analysis helps to ensure that the evidence base has been comprehensively tested.

Tip: It can also be useful to attempt to rank which interventions (or group of interventions) have made more important contributions than others.

Qualitative research and storytelling can play a crucial role, alongside quantitative evidence, in the difficult task of correctly attributing changes in the financial sector, as well as in growth and poverty reduction, to the activities of FSDs. Moreover, this analysis can be highly relevant for FSDs in regard to improving their programmes at periodic points in a strategy cycle. This kind of research is not the same as telling anecdotes: standards of rigour and credibility in qualitative research are needed. As Copestake and Williams (2011) note, 'Smaller and more flexible studies based on careful interpretation of systematically collected self-attributed impact data can provide faster and more context-specific feedback, and hence do more to strengthen learning, experimentation and improved practice in complex and fast changing environments than a smaller number of larger and lengthier studies'.

Triangulation is a core principle when testing each step of the logic. This may not be feasible in all cases, and will depend on the resources available to gather evidence from multiple sources. Figure 21 illustrates the points in the ToC at which the contribution analysis will be applied, and Table 29 shows how the evidence can be triangulated using four (illustrative) sources, bringing together the strengths of each – the four sources being: results chain monitoring; market actor interviews; landscape studies; and research, whether global or undertaken by FSDs. Of course, in practice FSD programmes can use many more potential sources of evidence.

Figure 21 Testing points in the ToC

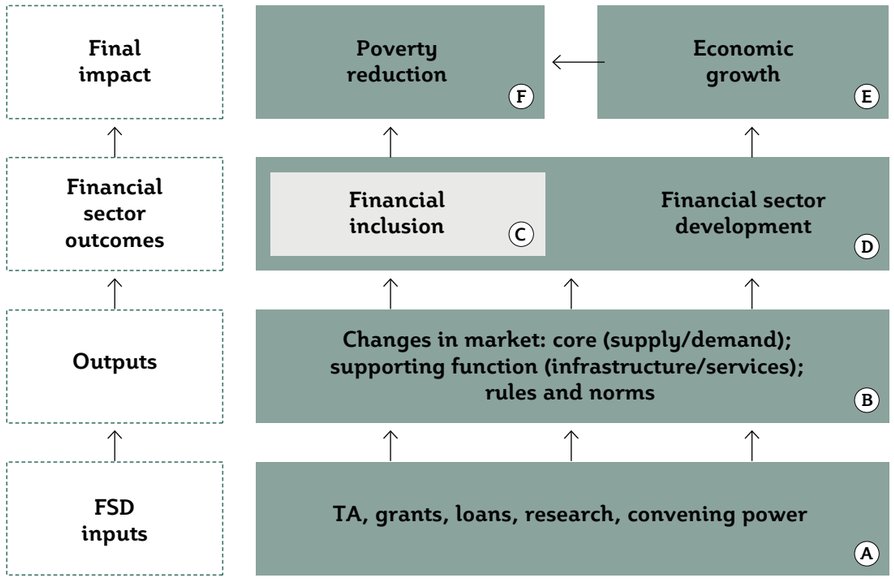


Table 29 Triangulation of the evidence (illustrative example)

Observed changes in ToC

SOURCE OF EVIDENCE \ TESTING POINT	A	B	C	D	E	F
Results chains	***	*	**	*		
Market actor interviews		*		*		
Landscape studies			*		***	*
Global/FSDA research		*		*		

Causality in the ToC

SOURCE OF EVIDENCE \ PATHWAY	A → B	B → C-D	D → C	C → F	D → E	E → F
Results chains	*	***	*			
Market actor interviews		*	*			
Landscape studies				***		
Global/FSDA research		*		*	**	***

Strength of evidence: ***Strong **Medium *Weak
*in practice, there are many more sources of evidence than shown in the above table that can be used

The preceding table may appear daunting, but it presents a summary of what FSD programmes and their evaluators should look at in order to document change processes and the extent of those changes. The evidence table can be used to make judgements about a programme's causal links, with the relative extent of an FSD programme's contribution to those links being articulated (for example, 'a second order contribution'). The triangulation process is illustrated above using an example based on an impact assessment of FSDK undertaken in 2010.[82] The example is taken from the micro level (of the market development programme), based on a range of projects that aimed at building the capacity of retail financial service providers in various different ways.[83]

Please note that Table 28 has two parts – strength of evidence as to whether there is greater financial inclusion or financial sector development (irrespective of what is causing this) and the second part, which then considers the evidence for causality. Thus, some strong evidence of poverty reduction could be possible in a case where there is weak or no evidence that financial inclusion is causing this reduction in poverty. FSDK's 'programme theory', as articulated during the impact assessment, was somewhat differently presented from Figure 21 above, but the structure is similar, as illustrated in Table 30.

82. Stone et al. (2010).
83. The projects investigated for the assessment were: support to Equity Bank to transform itself from a building society to a bank; support to Faulu and the Kenya Women's Finance Trust (KWFT) for transformation under the 2006 MFI Act; and support to Kenya Post Office Savings Bank (KPOSB) to introduce a new business model.

Table 30 Programme theory for FSDK retail capacity building

Programme stage (FSDK definition)	Specification		Testing point (as shown in Figure 22
Final impact	Services used by poor people are reducing vulnerability and increasing incomes		F
↑ **Direct impact**	↑ Increased provision of appropriate and affordable services to poor Kenyans		C
↑ **Outcome**	↑ Increased institutional viability		B
↑ **Output**	↑ Transformed MFIs (Equity Bank, Faulu, KWFT)	↑ New business model in place (KPOSB)	B
↑ **Activity**	↑ Capacity building of retail providers		A

The assessment of the pathway from activity to output to outcome (A to B to C) produced the result illustrated in Table 31 (at the time the assessment was done, the operation of the pathways from B to C could be credibly tested only for Equity Bank, and it was too soon to test the pathway from C to F for any intervention). As noted earlier, FSDs should be interested in impacts at all levels and not just the final impact on end-users.

Table 31 Illustration of triangulation of evidence: Capacity building for service providers: activity to direct impact (A→B→C)

Item of evidence	Type of source	Confirms / refutes /other	Strength of evidence
Equity Bank successfully transforms and increases deposit account numbers	Bottom-up, results chain study	Confirms	***
Faulu and KFWT successfully transform and increase deposit account numbers	Bottom-up, results chain study	Partially confirms	***
FSDK facilitates the transformation process	Top-down, market actor interviews and FSDK research	Confirms	**
Relevant regulations change	Top-down sector tracking	Other contribution	**
FSDK speeds up licensing process	Top-down, market actor interviews	Confirms	*
Successful introduction of new business model by KPOSB	Bottom-up, results chain study	Confirms	***
KPOSB achieves sustainable outreach	Bottom-up, FSDK market research	Inconclusive	*

The judgements involved in this process need to be systematically verified by asking:

- How credible is the story?
- Do reasonable people agree with the story? Does the pattern and timing of results observed validate the results chain?
- Where are the main weaknesses in the story?

As noted in Step 5, the standard for proof is what would convince a 'sceptical observer'. To assess the contribution story, it is therefore useful to ask a number of independent actors to analyse the emerging narrative, to assess its credibility and weaknesses. Ideally this would include both country-based and non-country actors, ensuring both depth and breadth of understanding.

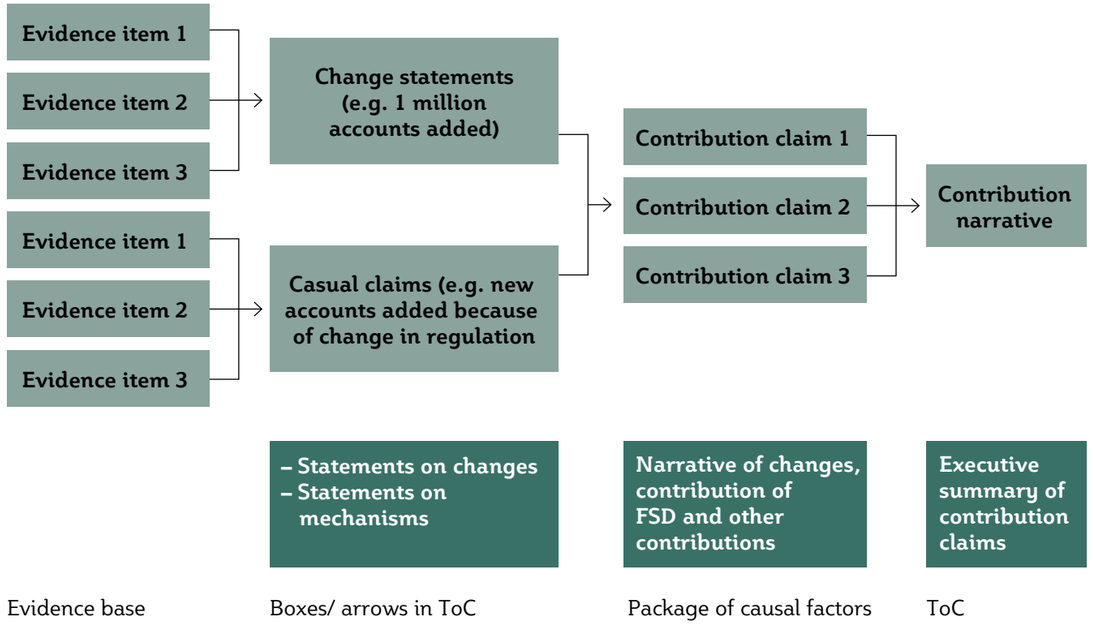
6.1.2.5 Revise and strengthen the contribution narrative

Finally, a contribution narrative for an FSD's programme impact can be developed. A positive contribution story may flow along the following lines:[84]

There is a reasoned ToC for the intervention: the key assumptions behind why the intervention is expected to work make sense, are plausible, are supported by evidence and/or existing research, and are agreed upon by at least some of the main financial sector players.

i. The activities of the intervention were implemented as set out in the ToC.
ii. The ToC —or key elements of it— is supported by and confirmed by evidence regarding observed results and underlying assumptions, and the chain of expected results occurred. The ToC has not been disproved.
iii. Other influencing factors have been assessed and either shown not to have made a significant contribution, or their relative role in contributing to the achieved result has been recognised.

Figure 22 A contribution narrative



Source: adapted from Delahais and Toulemond (2012)

84. See http://betterevaluation.org/plan/approach/contribution_analysis.

All contribution claims are ordered into or recorded in the boxes and arrows of the logic model and assembled into the contribution story (Figure 22). Again, it is useful to challenge this story in an open and critical environment with stakeholders who are distant from the evaluation team but who are well informed about what happened, as well as why or how it happened.

Contribution claims are composed of a series of change statements and causal claims, backed up by evidence as presented in Table 31. A contribution claim asserts that an intended change:

i. did or did not occur;
ii. occurred due or not due to the intended contribution;
iii. in conjunction with a few selected contextual factors;
iv. all considered mechanisms being explained and ranked by order of influence; and
v. other non-selected mechanisms being acknowledged.

The claim is said to confirm the logic model if the intended change occurred and if the intended contribution is highly ranked in comparison with other contributing factors.[85] Contribution claims often include statements about the magnitude of the causal relationship, expressed in terms such as 'major', 'minor', 'marginal', 'important', and so on.

6.1.3 Contribution analysis for aggregating quantifiable results

To provide a comprehensive picture of programme impact and to undertake VfM analysis it is useful to aggregate results where possible. Using the contribution analysis helps build up the evidence base and thus helps us to know what, and how much, to aggregate. Ideally, FSDs want to be able to aggregate outcome indicators across, for instance, thematic areas. For FSDs this tends to focus on 'number of accounts added' or 'customers provided credit/savings'. Aggregating requires two processes: first, knowing how much of an observed change can be claimed as the programme's contribution, and, second, the extent to which changes from different interventions can be aggregated together. This is useful for impact evaluation as it provides a sense of the magnitude of the programme's contribution to observed changes, but it can also be useful for undertaking VfM, given the quantitative approach to aggregation in Box 29.

While aggregating various sources of data, problems such as double counting, inactive accounts, multiple accounts per client, claiming the same account opened via different interventions and poor MIS systems need to be considered. To avoid double counting, programmes can estimate the level of overlaps between interventions and discount accordingly. For example, DCED recommends that for overlaps of less than 5% the programmes can add all end-users, and for overlaps higher than 95%, only the largest number for the largest project should be counted. For overlaps of between 5% and 95%, the projects are encouraged to estimate each overlap and show calculations.

There are different approaches to quantifying a programme's contribution:

- Some FSDs count all (100%) of results that their partners have produced (i.e. change from baseline) attributed to the FSD intervention (ideally backed up using the causality methods set out in Step 5).
- If other donors are also contributing to a project, the FSD programme may only count a percentage of the results achieved from a partner. For example, if an FSD contributes 20% to a project it will claim 20% of any results achieved.[86] This assumes a very general one-to-one ratio in terms of return on investment.
- An FSD can make an estimate based on the level of contribution it feels it has made to results vis-à-vis other factors (i.e. this may be more or less than the proportion of its project funds, but it could also reflect the criticality of the input).[87]

Tip: In some ways it does not matter which approach is taken (unless funders direct otherwise) as long as there is clear evidence that the FSD is causing or contributing to the change, and that the assumptions used to quantify this are transparently set out. But it is still important to try to judge how additional the FSD contribution is to the market forces, or other influences at play, and if the evidence for this would convince a reasonable, but sceptical, observer.

A key risk when aggregating results is that the programme becomes focused on only reporting more easily quantifiable micro-level changes in the market, rather than changes to the overall system. This is a real risk and needs to be recognised and managed, including by ensuring that appropriate and adequate effort is made to report on how the underlying structures of the market have changed (i.e. it is not all about numbers). However, macro and meso projects can use approaches set out in this IOM, along with sector-level tracking, to make estimates of some common indicators, such as 'number of accounts added'. The assumptions used in these calculations will need to be clear and transparent.

85. Ibid.
86. This is the approach DFID recommends.
87. An FSD colleague has argued that sometimes the FSD input is akin to providing a car with an engine, which makes all the difference between a functioning and a non-functioning car. In such a case, should FSD contribution be assessed in terms of proportion of inputs or criticality? Ultimately, this is a matter of judgement and agreement will need to be reached with funders and any external evaluators.

Box 29 VfM assessment

Assessing VfM has become an increasingly important aspect of funders' measurement processes. This tends to assess the three 'Es' (economy, efficiency and effectiveness). An overall – cost effectiveness – calculation is also often provided, which compares the costs of a programme against the monetary benefit it has provided. This is difficult for FSDs, given the challenges of providing a single measure of their impact.

There are two broad approaches to such measurements for FSDs. First, as conducted for FSDK in 2012 and 2014, a set of quantifiable projects can be measured to assess the monetary benefit that they have generated. The second approach takes a sectoral perspective and assesses to what extent sector trends (e.g. an increase in access) provide a monetary benefit to the sector. Both approaches require understanding the types of monetary benefits that may result from changes in the financial sector, the magnitude of these, and, most importantly, the level of contribution an FSD programme has made to them (i.e. what percentage of the monetary benefit can be allocated to the FSD). The IOM assists in this process in two ways: first by setting out and measuring the various impact pathways resulting from the FSD's interventions (and therefore potential impacts that can be monetised), and second by providing a robust evidence base for transparently setting out the assumptions based on which an FSD can claim a proportion of this monetary benefit.

For example, we can take OPM's VfM analysis of FSDK's work in the SACCO sector. The sector suffered many problems in terms of challenges around governance, regulation, supervision, and liquidity, which led to FSDK choosing to work on improving the regulation of the sector. A number of potential benefits (i.e. reducing risk around losses reported through SACCO accounts) from improved regulation in the sector were identified. Based on a number of assumptions regarding risks of transacting with SACCOs (e.g., the differences between the rates of depositor losses in SACCOs and those in commercial banks was used as a broad indicator of the greater risks associated with the poorly regulated SACCOs) and the magnitudes of these risks, based on the data from FinAccess and the regulator, and the potential benefits of regulation (based on the project's ToC), it was estimated that reduction in losses would generate a benefit of around £19.24 million. Together with other FSDK projects this could then be compared with the total costs of the FSDK programme.

Source: Assessing Value for Money – The case of donor support to FSD Kenya, OPM 2012

Box 30 Step 7 checklist

- Identify points in the programme cycle (e.g. annual review, annual reports, mid-strategy review, strategy refresh points, end of strategy) when internally documenting the programme impact is useful
- Each FSD should aim to finalise a plan to bring together the evidence of key changes and FSD's contribution for the overall programme and key projects
- Each annual review can be used to present such an analysis for one or more theme/ project
- Assess how each step of the contribution analysis (or other causality method) will be undertaken, and by whom (see also the next chapter)
- Explicitly set out how the FSD's contribution to quantifiable results is being calculated?

IOM – Chapter 7: Implementing the IOM

Stage 1: Clarity of purpose

Step 1: Setting out an evalaution Programme ToC

Step 2: Developing impact measurement questions

Stage 2a: Measuring change – what happened?

Step 3: Developing indicators

Step 4: Data collection methods and sources

Stage 2b: Measuring change – why it happened?

Step 5: Assessing causality and contribution

Step 6: The research agenda

Stage 3: Bringing it all together

Step 7: Developing a credible narrative

Implementing the IOM (Chapter 7)

This final chapter of this guidance, implementing the IOM covers issues throughout the previous three stages.

In particular, it sets out the key strategic opportunities for embedding IOM in FSD operations, including integrating IOM principles and guidance into existing M&E systems.

7 Implementing the IOM

7.1 Overview

- Monitoring or evaluation on their own (in isolation) cannot measure the changes that result from the interaction between evolving FSD programmes and dynamic market contexts.
- Impact orientation needs to be embedded in existing FSD monitoring systems, as well as in the design, implementation and review of the project and the overall programme.
- Key strategic opportunities for embedding IOM in FSD operations include: the development and revision of the overall strategy and the logframe; investment decisions and reviews of specific projects/ interventions; annual reviews; programme evaluations and specific impact assessment studies.
- Up-front investment in results measurement is valuable as much for improving programmes as for confirming and communicating results, and needs to draw on the technical skills of those who implement the FSD programme, as well as those of the M&E team and the FSD network.
- A broad/concurrent partnership between the FSD and an independent organisation is recommended to assess and confirm the impacts of the programme, and to make changes to programme design and monitoring systems, to make it more evaluable.
- The issue of the independence of the evaluator can be addressed by entering into a pro-active dialogue with the FSD governing body/ PIC (which includes funders) and by setting up a sub-committee of the FSD governing body. This sub-committee should then become responsible for the recruitment and oversight of this evaluator.
- A number of options exist throughout an FSD programme cycle for integrating IOM principles and guidance into existing M&E systems.

7.2 Building on existing M&E systems

Section 2.1 noted that the purpose of IOM is to improve FSDs' (and their funders') ability to measure and evaluate their impact. Section 2.3 provides arguments about how FSDs and other stakeholders can derive benefits from the IOM process. **This guidance paper recognises that the nine FSDs in Africa (including FSDA) are at different stages of implementing a results framework, and only some of the FSDs have already conducted, or are considering commissioning, programme-level evaluation,** even if project-level evaluations are more common. That is why this guidance paper has not been prepared as a manual. Since monitoring or evaluation on their own (in isolation) cannot measure the changes that result from the interaction between evolving FSD programmes (as market facilitators), and dynamic market contexts, it has been agreed that impact orientation needs to be embedded in the monitoring systems as well as in the design, implementation and review of the project/ programmes.[88] This is captured in the 'sweet spot' that is identified as sitting between monitoring and evaluation in Figure 2, and is further elaborated below.

7.3 Strategic opportunities to mainstream IOM

IOM should be fully integrated into the M&E function as well as the overall programme implementation. To 'mainstream' IOM, the first step is to identify the various stages in the project/ programme planning and its implementation, and to identify areas in which IOM can be embedded as a part of the core internal processes of the FSD. The main areas are:

- development and periodic revision of the overall FSD strategy;
  finalisation and periodic review of the logframe;
- investment decisions on specific projects/ interventions, and periodic monitoring of these, and of the wider system;
- annual reviews;[89]
- programme evaluations (e.g. mid-term reviews and end-term reviews);
- specific impact assessment studies; and
- through work with the FSD Academy and FSDA working groups.

Discussion point: FSDs were clear that they viewed the role of FSDA in relation to measurement as creating space for dialogue and exchange, as well as facilitating training and research. There was greater reluctance to see standard indicators imposed across the board. However, there was an appreciation that there were common elements of FSD programmes that could benefit from a harmonised approach, which FSDA could help facilitate. For example, this may include help thinking through the types of systemic change and indicators that are specific to saving groups projects, which a number of FSDs currently facilitate.

88. By FSD programme teams and others participating in the two workshops (October 2014 and March 2015), and a number of other discussions.

89. Traditionally referred to by DFID as output to purpose reviews (OPRs).

7.3.1 Development and revision of the overall FSD strategy

As noted in Section 3.4, to ensure better evaluability of an FSD programme, **it is important to look at both the theory and the practice of the programme.** The conceptual framework that sets the programme ToC is set out in the overall FSD strategy (normally for a three- to five-year period), and at the same time the funding period and scale of funding is also finalised.

Many of the older FSDs have gone through more than one cycle of development and revision of the overall strategy. This is influenced by a number of factors: the experience of FSDs (what works); better understanding of the market context, key constraints to financial sector development and demand for FSD support (e.g. from policy-makers, market actors and support agencies); and key priorities identified by funders at the stage of finalising funding agreements.

How one should go about checking that the FSD ToC is evaluable is explained as a part of the Stage 1 guidance. For an existing FSD programme, there may be limited appetite to review and revise the overall programme logic in the middle of a strategy period. A key strategic opportunity for testing the evaluability of the ToC is when the overall FSD strategy is being developed or revised (Step 1). However, the process of development and agreement of impact measurement questions (Step 2) can be done at any stage. **In fact, without agreement on what levels and types of impact an FSD programme is interested in,[90] it is very hard to integrate the IOM into the existing monitoring system.**

7.3.2 Finalisation and periodic review of measurement frameworks, including the logframe

While the programme logframe can provide the overall outcomes, on its own it may be unable to help prioritise the impact measurement questions on which the programme should focus.

The logframe is primarily used for accountability purposes. DFID's guidance on logframes suggests one outcome and a maximum of 10 outputs, although programme teams are encouraged not to prioritise more than six outputs. Each output can have a maximum of three indicators. Therefore, around 20 indicators may be used by DFID and other funders to track the progress of the programme at output and outcome levels. **As advised in Step 3, the IOM measurement framework will extend beyond this, with the logframe providing a sub-set of the types of evidence being collected.**

However, measurement frameworks can and should be used for more than just reporting to funders. Market facilitators like FSDs need to monitor much more than just accountability indicators in the logframe: they need to build partnerships with policy-makers, business associations, financial institutions and others to identify which data can help track and then move markets.

Market development programmes need faster feedback between intervention and results and should use results measurement to adapt to changing markets, scale up what works, and play down or discard what does not work. **They can also use IOM as a powerful mechanism for influencing market actors, regulators and policy-makers, who are interested in tracking financial inclusion objectives and trends, as well as policy implications.** Even financial institutions have an interest in the measurement of results to identify business opportunities with underserved customers or to better serve financially included customers. Evidence-based advocacy is a critical part of market facilitation and FSDs can help identify, prioritise, collect, analyse and disseminate market data, especially where they fund the generation and initial analysis of those data.

Tip: FSDs should talk to their funders about using logframes flexibly (which is allowed!), alongside other indicators collected.

Measurement indicators can be quantitative as well as qualitative, and can have a short-, medium- or long-term orientation. Step 3 provides examples of tracking the behaviour of directly supported partners in the short term. This may lead to FSD-supported partners increasing access and usage, as well as revising business models and practices in the medium term. In the long term, FSDs are interested in systemic changes in the behaviour, performance, sustainability, scale and resilience of the wider market players, and in the implementation of more enabling rules and norms at macro level, combined with responsive support organisations at the meso level.

7.3.3 Investment decisions regarding projects and other interventions, and periodic FSD monitoring

IOM can and should inform the way FSDs make their investment decisions. In the FSDs that were set up initially, investment decisions were typically driven by the output (and outcome) indicators in the logframes – usually quantitative measures. What IOM calls for is a perspective that looks beyond the logframe to consider

90. E.g. Section 3.6 confirms that FSDs focus on different areas of impact – financial sector development, financial inclusion, livelihood improvements, etc.

not just the prospective effect of an intervention on outcomes, and perhaps impacts, **but also on whether that intervention is likely to contribute to systemic market change – and what those results chains might look like.**

This guidance paper therefore calls for FSD programmes to adapt the documentation, such as project appraisal reviews (PARs), that they present to their decision-making bodies so that they capture not only the direct output and outcome measures, but also the wider systemic changes they expect to see and the indicators proposed to capture these. Sometimes the indicators adopted to measure systemic change will be qualitative. **As this guidance paper has noted before, assessing impact is not just about numbers.**

PARs and similar documents used to present and support investment proposals also usually set out the methods used to collect the data. Typically, this is generated by investees, at least for output measures. However, measurement of outcomes may need additional research and evidence that FSDs will have to collect from FSPs, sector associations and regulators, e.g. some FSD partners working in the private sector, may not be interested in issues such as replication and crowding in. FSDs may also need to commission specific studies.

When it comes to tracking systemic changes, this calls for FSDs to at least modify and adapt their regular monthly and/ or quarterly meetings. **Tracking what is happening in the wider system (i.e. the financial sector as a whole), is something that already happens anyway in most FSDs.** Individual managers will read items in newspapers or come across them in other media. They also have direct conversations with policy-makers, senior executives in financial institutions and other market actors and other observers. So the data is usually generated. **However, there are often two missing steps:**

i. More often than not FSDs do not record these new bits of information in a form that is readily accessible, and that over time can be analysed to look for trends or new factors emerging that might contribute to systemic change and how this is documented to provide evidence for learning and evaluation purposes.
ii. They do not have a means of reviewing such new information systematically and asking themselves what might be the implications for systemic change. This also applies to data gathered as part of regular project monitoring.

This guidance paper therefore recommends that FSDs set up mechanisms to periodically capture the insights of staff and partners, including those set out in Annex F. In addition to adding this as an agenda in quarterly/ six-monthly meetings and changes to the reporting formats, FSDs may consider setting up a small group within their management teams, possibly augmented by a knowledgeable outsider from, say, the financial sector (who may also be a member of an FSD's governing body).

The role of this 'systemic change monitoring group' (a specific sub-category of normal FSD monitoring) could be to ensure (a) that information gathered from both public and private sources is recorded systematically (for example on simple templates), and (b) that every quarter or six months it is used to assess what changes are being seen in the financial markets. The key issues and conclusions reached by the 'systemic change monitoring group' should also be recorded and be readily accessible. Some of this analysis is captured in the annual report. However, the underlying evidence and some of the hypothesis which it may be pre-mature to explicitly share with external stakeholders at the stage of report writing, are not recorded and get lost. As well as monitoring existing investments/ interventions for signs of systemic change, such records should guide future investment decisions and may well provide an important source of data for evaluators several years down the line.

7.3.4 Annual reports and discussions with funders

FSDs normally submit an annual report to funders and other stakeholders. They also usually carry out an annual review and related discussion – the OPR. This is often based on the FSD's internal monthly/quarterly reviews at project level and an annual report prepared for this purpose.

To date, OPRs focus on the achievement (or not) of outputs and outcomes identified in logframes. This, however, rarely captures the full picture of what FSDs have accomplished in the previous 12 months, although recently adapted or developed logframes in some FSDs may seek to capture systemic changes that are underway or that have been achieved.

An annual review process that seeks to provide not just an OPR but also a detailed IOM analysis will provide a much richer source of information and guidance, both for funders and FSD management (see Step 7).

In addition, it can also provide a useful opportunity to take up a specific programme theme for more detailed analysis – for instance, to confirm evaluation questions, to identify existing sources of evidence and build on the insights of FSD staff and partners, to confirm causality, as well as to identify possible follow-up work/ specific studies for the following year to strengthen the impact orientation of the existing monitoring arrangements. Again, records from the 'systemic change monitoring group' should also inform such analysis, as well as provide data that contribute to progress towards, or the achievement of, desired systemic changes initiated by an FSD's interventions.

7.3.5 Programme evaluations (e.g. mid-term and end-term reviews)

Evaluations may be undertaken for a variety of objectives – to prove the results of the programme as well as to improve FSD programme performance, e.g. document programme effectiveness, test specific delivery models, innovations, and implementation choices, as well as to improve performance of FSD partners and their operating models.

As with annual reviews, adopting the IOM approach will mean that the processes of undertaking evaluations/ reviews are likely to be both more efficient and be based on richer, broader sets of indicators and other information. **They provide an opportunity to undertake a robust TBE of the programme impact (see Step 7).** This in turn is likely to generate more insights and lessons to be fed back into the particular FSD (in the case of mid-term reviews) and the wider community of FSDs (in the case of end-term reviews). Further discussion of mid- and end-term reviews is provided below in relation to independent evaluations (see Section 7.5).

7.3.6 Specific impact assessment studies

FSDs may also commission specific studies based on particular research needs and impact questions previously identified, or to confirm the impact of a flagship project. These studies could focus on confirming causality at specific links in the results chains, ToC, demand and/or supply-side analyses of financial services access and use, as well as on understanding and confirming specific impacts.

7.4 Investing in results

All FSDs have agreed logframes and approved business cases which provide the high-level M&E strategy and focus. These are agreed with funders. However, in some cases an FSD has to report against more than one results framework, as some funders may have a different focus and reporting requirement.

Tip: FSDs should work with funders to agree that the IOM reporting system can be a common reporting requirement around programme impact.

This guidance focuses on integrating results measurement in the entire programme cycle – scoping, programme design, implementation and review. This necessitates the use of the technical skills of staff responsible for both programme implementation, as well as those focusing mainly on M&E. Considerable variation also exists across the FSDs in terms of human and financial resources committed to M&E. Up-front investments in results measurement is valuable as much for improving programmes as for confirming and communicating results. Thus, irrespective of the final arrangements around the involvement of an independent evaluator (see below), dedicated results measurement capacity within the FSD is essential in order to fulfil a wide range of expectations.

What should be the balance of the effort expended between implementing programmes (achieving results) and assessing programmes (measuring results)? For example, should an FSD invest 5% or more/less of its annual budget on M&E? Once the impact orientation is embedded in the core programme design and implementation, this becomes less of an issue. **Many of the results measurement functions are indeed part of programme implementation** and so should not strictly be treated as a general overhead cost. Furthermore, many of the public goods functions of FSDs (such as FinScope/ FinAccess and other research, sector-level data tracking) should be part of programme costs, not overhead costs. In any case, once the IOM framework is accepted, it is possible to have a more constructive conversation between FSD management and their investment/ oversight committees and funders around the allocation of human and financial resources for results measurement. Such discussions should also address the question of the budget lines to which impact evaluation-related expenditures should be allocated.

7.5 Role of an independent evaluator

Decisions around final M&E arrangements need to be made in consultation with an FSD's governing body (which includes funder representatives). In addition to programme-level monitoring and reporting, funders often seek independent evaluations of a programme.

Dedicated results measurement capacity within the FSD is key to fulfilling various expectations. In addition to internal capacity, the IOM guidance provides three potential scenarios for how FSDs can inject greater expertise and independence into their IOM approach. These scenarios, which are not mutually exclusive and have potential overlaps, are:

a. rely on the IOM system to produce evidence, with the process of implementing the IOM tested by an independent evaluator;
b. a broad/concurrent partnership between the FSD and an independent organisation(s) to assess the impacts of the programme, and make changes to programme design and monitoring systems to make it more evaluable, using both monitoring data and specifically commissioned impact research; and
c. periodic external impact evaluations to assess if the programme is delivering the expected results, e.g. at the mid-point and end-point of the strategy, but

using the evidence collected through the IOM.

The implications of each of these choices for FSDs and funders are noted in Annex G. Before finalising arrangements, it is important to note why independent evaluation is such a big issue for the funders and what role an independent evaluator can play. As a DFID-appointed *Independent Advisory Committee for Development Impact* noted in 2008: independence is central to the credibility of evaluation. The committee noted that accurate and fair evaluations combine intellectual detachment with empathy and understanding.[91] As is clear from the core principle of bringing monitoring and evaluation closer together, external evaluators often lack an appreciation of the operating context.[92] Independence combined with disengagement increases information asymmetry, ruptures contacts with decision-makers and restricts access to relevant sources of information. Good evaluation, just like good science, calls for a frame of mind that is characterised by curiosity, scepticism and a hunger for evidence.

Independent evaluators can make use of the work undertaken to implement an IOM in several ways:[93]

The reliability and quality of the monitoring data produced by programmes that are implementing the framework proposed in this guidance are likely to be of a higher quality than those of other programmes.

- A focus on defining indicators of change, baselines and ways of measuring change should lead to this improvement, along with an overall focus on reporting and the results measurement system.
- Clear articulation of the intervention results chains, their evidence bases and associated indicators provide the basis for understanding the programme's ToC. The evaluator may supplement and validate the results chain and incorporate this when confirming whether the ToC holds up in practice.
- The ToC may be used to determine evaluation questions (jointly with funders) and to agree the evaluation approach. The use of evaluation questions that have already been agreed with funders can help inform the scope of work for the evaluator and reduce the risk that the evaluator will assess programme performance against newly created impact evaluation questions, for which the evidence may be much less readily available.
- The attempts made to estimate and justify the attribution of impact to the programme and the measurement of systemic change can be validated by the evaluator. These may be sufficiently rigorous to be included in the evaluation approach, or may be subject to further analysis and data gathering by the evaluator.
- Cost data tracked by the programme will support any VfM and cost efficiency related evaluation work undertaken by the evaluator.

Each of the three options recommended for FSD has strengths and weaknesses in relation to specific interests, such as: the ability to support real-time learning; the degree of focus on impact; independence; the human and financial resources required from the FSD; and additional data collection/analysis required. **This guidance recommends option b) above for most FSDs.** On the differing interests listed above, this option scores lower on independence, but more than compensates through stronger ownership and usefulness of the results measurement process for FSDs and funders. Fortunately, this option is increasingly being implemented by many funders in order to facilitate learning from the market development programmes.

Moreover, the issue of independence can also be addressed. FSDs should initiate a pro-active dialogue with their governing body (including funders) to finalise these arrangements.

Tip: To ensure independence, a sub-committee could be formed by the FSD governing body, which will be responsible for the recruitment and oversight of the evaluator.

The FSD team should work closely with this committee to define the scope of the external evaluation (impact evaluation questions and terms of reference) and the possible nature of arrangements (long-term, periodic engagement) so that the evaluators can guide and quality assure the IOM process and the impact results/ communication. We suggest that the FSD(s) develop a long-term relationship with an evaluator who can also play the role of a learning partner, and should contract a research firm (rather than an individual) on a long-term call-down contract.

91. Drawing on the good practice standards of various agencies, the committee highlighted four inter-related criteria for ensuring independence: (i) organisational independence; (ii) behavioural independence; (iii) protection from external influence; and (iv) avoidance of conflicts of interest. This means that an evaluation is independent when it is 'carried out by entities and persons free of the control of those responsible for the design and implementation of the development intervention' and enjoys 'freedom from political influence and organizational pressure', 'full access to information' and 'full autonomy in carrying out investigations and reporting findings'.
92. The ability to engage with diverse stakeholders and secure their trust while maintaining the integrity of the evaluation process is the acid test of evaluation professionalism and diminishing returns can arise when evaluation independence assumes extreme forms of disengagement and distance.
93. This analysis draws on the discussions of Calvert (2014).

Tips for appointing an independent evaluator

FSD programmes should consider the following tips when appointing an independent evaluator:*

- On its own, accountability to funders has not been successful in driving better measurement. What creates incentives for better measurement is the drive, on the part of the implementation and evaluation teams, to be effective.
- Early dialogue, and good chemistry, between implementers and evaluators are important; the relationship is that of a 'critical friend'. The funder, implementer and the evaluator should discuss and agree clear role definitions for those implementing the FSD and the independent evaluator, as well as key principles that will guide the relationship, and possible mechanisms to address any issues that may emerge.
- Consider who will be responsible for the collection of different data: evaluator or implementer? There is a variety of possibilities here, from evaluators collecting all their own data, through to validation of monitoring data. Some implementers are sometimes concerned that the evaluators could disrupt their relationships with private sector partners, ask inappropriate questions, raise expectations or make excessive demands on the implementation partners.
- Evaluators might focus on the counterfactual and on the collection of qualitative data. However, it is important to note that the qualitative data and quantitative data have to be linked to clear lines of enquiry in order to be able to help articulate a credible story.
- The relationship between the implementing and evaluation team could be damaged if evaluators are also given responsibility for conducting annual reviews.
- Appointing an independent evaluator at the start of the programme means that the evaluation units do not have to establish their credibility in midstream, when catching up with the implementation team.
- Based on when the evaluator is appointed and the balance of focus between accountability and learning, the evaluation team can engage with the implementation team on the design/ review of the ToC/ results chain and the logframe.
- Consider issues of access to, use and dissemination of confidential data from private sector players and central banks (and issues such as non-disclosure agreements).

** This section draws on a DFID/DCED-facilitated discussion with different evaluation agencies that took place on 13 January 2015, as well as practical experience gained by the OPM team while conducting multiple evaluation assignments.*

7.6 Summary

Table 32 Summary of implementing IOM

	FSD timeline	Applicable IOM guidance
1	FSD strategy development/ review	Step 1 and Step 2
3	Finalisation and periodic review of measurement frameworks, including the logframe	Step 1, Step 2, Step 3 and Step 4
4	Investment decisions about specific projects/ interventions, and periodic monitoring of these and the wider system	Step 3, Step 4, Step 5, Step 6
5	Annual review	Step 7
7	Specific impact assessment study	Step 5
6	Programme evaluations (e.g. mid-term review and end-term review)	All, especially Step 7 and 'Implementing framework' (for independent evaluator guidance)

Box 31 Chapter 7 checklist (implementing the IOM)

- The FSD should discuss the overall results measurement approach with its governing body/investment committee and funders
- The governing body/investment committee should also agree on budgets and other resources (i.e. human resources) to be allocated to IOM, as well as M&E
- The FSD team should review the guidance in this chapter and should take advantage of various possible strategic opportunities to integrate IOM within its existing programme design, review and reporting processes
- The FSD should have a time period for reviewing and updating current internal M&E processes/ guidance in the context of IOM guidance
- The FSD should have a documented plan, which is understood by all staff, on how it intends to implement IOM (or parts of). Are responsibilities and tasks well known by staff?
- The FSD should distinguish between measurement for accountability and that which can aid market facilitation. While some overlaps of measurement processes will emerge, a good analysis for market facilitation should also help in gathering evidence for causality and an overall impact narrative
- The FSD should work with funders to ensure there is one common results reporting framework
- The FSD should work with funders to prioritise timeline, process and key questions for the external evaluation of the programme
- The FSD should work with FSDA to ensure that over time relevant research is accessible in one place, so that evaluation/ research priorities for the FSD can be established

References

Arora, S, Roe, A. and Stone, R. (2012), 'FSD Impact Research and Value for Money Assessment: A Suggested Approach', OPM.

Beck, T. (2008), 'The econometrics of finance and growth', Policy Research Working Paper 4608, World Bank.

Beck, T. and Cull, R. (2013), 'Banking in Africa', Policy Research Working Paper 6684, World Bank.

Beck, T. and Levine, R. (2004), 'Finance, Inequality and Poverty: Cross-Country Evidence', World Bank Policy Research Working Paper 3338.

Beck, T., Demirgüç-Kunt, A., Laeven, L. and Levine, R. (2004), 'Finance, Firm Size and Growth', World Bank.

Beegle, K., Dehejia, R. and Gatti, R. (2003), 'Child Labour, Income Shocks and Access to Credit', World Bank Policy Research Working Paper 3075, June 2003.

Boulton, J. and Johnson, S. (2013), 'Impact assessment of financial market development through the lens of complexity theory', University of Bath.

Britt, H. (2013) 'Complexity-Aware Monitoring', USAID Discussion Note, Version 2.0, December 2013

Calvert, S. (2014), 'Evaluation and the DCED Standard for Results Measurement', mimeo.

Calvert, S. (2012), 'Targeted SME Programmes: Evaluating Market System Projects', Evaluation Department, DFID, 18 June 2013.

CGAP-DFID (2013), 'Impact Evaluation for Financial Inclusion', workshop held on 10–11 January 2013.

Claessens, S. and Feijen, E. (2006), 'Finance and Hunger: Empirical Evidence of the Agricultural Productivity Channel', World Bank Policy Research Paper 4080, December 2006.

Clarke, G., Xu, L. and Zou, H. (2003), 'Finance and Income Inequality, Test of Alternative Theories', World Bank Policy Research Paper No. 2984.

Collins, D., Morduch, J., Rutherford, S. and Ruthven, O. (2009), 'Portfolios of the Poor: How the World's Poor Live on $2 a Day', Princeton: Princeton University Press.

Copestake, J. and Williams, R. (2011), 'What is the impact of microfinance, and what does this imply for microfinance policy and for future impact studies?', paper presented at MicroNed Conference, 'Taking stock of the evidence on impact, the way forward', Utrecht, Netherlands, 28 June 2011, available at http://www.opml.co.uk/sites/opml/files/impact_study.pdf.

Cortes, L. and Jenal, M. (2013), 'Monitoring and measuring change in market systems –rethinking the current paradigm', the SEEP Network.

Coryn, C., Noakes, L., Westine, C. and Schroter, D. (2011), 'A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009', American Journal of Evaluation 32(2). 199-226

Creevey, L., Dunn, E., Northrip Z., Snodgrass, D., and Cogan Ware, A. (2010), 'Degree Of Evidence in Assessing the Effectiveness of Economic Growth Programs', Private Sector Development Impact Assessment Initiative.

Cull, R., Ehrbeck, T. and Holle, N. (2014), 'Financial Inclusion and Development: Recent Impact Evidence,' CGAP Focus Note No.92, April 2014.

Davies, R. (2012), 'Planning Evaluability Assessments: A Synthesis of the Literature with Recommendations', DFID Working Paper 40.

Davis, R. (2012), 'Criteria for assessing the evaluability of a Theory of Change', Posted on 'Rick on the Road', 5 April 2012: http://mandenews.blogspot.co.uk/2012/04/criteria-for-assessing-evaluablity-of.html.

DCED (2011), 'Why have a Standard for Measuring results? Progress and plans of the Donor Committee for Enterprise Development, by Jim Tanburn and Nabanita Sen'.

DCED (2013), 'Guidelines to the DCED Standard for Results Measurement: Estimating Attributable Changes', DCED Standard, by Nabanita Sen, March 2013.

DCED (2014a), 'Why Evaluations Fail: The Importance of Good Monitoring', Adam Kessler and Jim Tanburn, August 2014.

DCED (2014b), 'Assessing Systemic Change: Implementation guidelines for the DCED Standard', Adam Kessler, August 2014.

DCED (2014c), 'The 2014 Reader on Results Measurement: Current thinking on the DCED Standard', edited by Adam Kessler. 2014

Delahais, T. and Toulemond, J. (2012), 'Applying contribution analysis: Lessons from five years of practice'. Evaluation July 2012 vol. 18 no. 3 281-293

Demirgüç-Kunt, A., Beck, T. and Honohan, P. (2008), 'Finance For All', World Bank.

DFID (2004), 'Financial Sector Development: A Pre-Requisite for Growth and Poverty Reduction?', Briefing Note, June 2004.

DFID (2014), 'DFID Evaluation Strategy: June 2014 to 2019'. DFID

Duvendack M., Palmer-Jones, R., Copestake, J., Hooper, L., Loke, Y. and Rao, N. (2011), 'What is the Evidence of the Impact of Microfinance on the Well-being of Poor People?', DFID Systematic Review, University of East Anglia.

Duvendack, M. (2013), 'How can monitoring data support impact evaluations?', Blog: 52 weeks of Better Evaluation: Week 44, at http://betterevaluation.org/blog/me-data-impact-evaluation.

Honohan, P. (2004a), 'Financial Development, "Growth and Poverty: How Close are the Links?"', World Bank Policy Research Working Paper 3203.

Honohan, P. (2004b), 'Financial Sector Policy and the Poor', World Bank Working Paper No.43, World Bank.

Hovland, I. (2007), 'Making a difference: M&E of policy research', Working Paper 281, Overseas Development Institute.

IFC/GTZ/DFID (2008), 'The Monitoring and Evaluation Handbook For Business Environment Reform', World Bank.

ICAI (2011), 'ICAI's Approach to Effectiveness and Value for Money', Report No. 1, November 2011.

King, R., and Levine, R. (1993), 'Finance and growth: Schumpeter might be right', Quarterly Journal of Economics 108(3).

Levine, R., Loayza, N., and Beck, T. (2000), 'Financial Intermediation and growth: causality and causes', Journal of Monetary Economics 46: pp. 31–77.

Li, H., Squire, L. and Zou, H. (1998), 'Explaining International and Intertemporal Variations in Income Inequality', Economic Journal, 1998, 108(1): 26–43.

Miehlbradt, A. and McVay, M. (2006), 'The 2006 Reader – Implementing Sustainable Private Sector Development: Striving for Tangible Results for the Poor', International Labour Organization.

Morcrette, A. and Pennotti, C. (2011), 'Know What You Know: Harnessing Tacit Knowledge in Value Chain Monitoring', the Groove Learning Network.

Nippard, D., Hitchins, R. and David Elliott (2014). 'Adopt-Adapt-Expand-Respond: a framework for managing and measuring systemic change processes', Briefing Paper, the Springfield Centre for Business in Development, March 2014.

OPM (2012) Assessing Value for Money – The case of donor support to FSD Kenya, OPM 2012

OPM (2014) 'Consultative process on FSD impact evaluation in Africa: Background Paper FSDA'.

Osorio-Cortes, L. and Jenal, M. (2013), 'Monitoring and measuring change in market systems–rethinking the current paradigm', the SEEP Network.

Patton, M. (2010), Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use, Guilford Press.

Ripley, M. (2014), 'Monitoring and measuring change in market systems in Nepal: Using the DCED Standard to help make sense of "messiness", Samarth-NMDP.

Rogers, P. (2009), 'Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation', Journal of Development Effectiveness, 1(3).

Ruffer, T. and Wach, E. (2013), 'Review of M4P Evaluation Methods and Approaches', DFID and Itad, April 2013.

Rutherford, S. and Arora, S. (2009), The Poor and their Money: Microfinance from a Twenty-first Century Customer's Perspective, Practical Action Publishing, UK.

ITAD (2012), GEMS Results Measurement Handbook, Version 1.0, December 2012.

World Bank/GPFI (2012), 'Impact Assessment Framework', SME Finance, World Bank.

OECD-DAC (2012), 'Evaluating Budget Support Methodological Approach', OECD.

SEEP (2015), 'Measuring Systemic Change – Building the Toolbox', SEEP Network, video presentation available at https://vimeo.com/121831416.

Springfield Centre (2014), Chapter 5: Measurement, in M4P Operational Guide, available at http://beamexchange.org/en/resource-detail/resource/134/.

Stern et al. (2012), 'Broadening the Range of designs and methods for impact evaluations', DFID April 2012.

Stone, R., Johnson, S, and Hayes, J. (2010), 'Impact Assessment of the Financial Sector Deepening Trust, Kenya', FSDK.

Ton, G., Vellema, S. and DeRuyter DeWildt, M. (2011), 'Credible evidence on complex change processes: key challenges in impact evaluation on agricultural value chains', "Munich Personal RePEc Archive Paper from University Library of Munich, Germany".

USAID (2008), 'Program Evaluation Methodology: The Degrees of Evidence Framework'. Breakfast Seminar #35, October 2nd 2008

Vogel, I. (2012), 'Review of the use of "Theory of Change" in international development: Review Report', DFID, April 2012.

White, H. (2009),'Theory Based Impact Evaluation: Principles and Practice', June 2009, Working Paper 3, International Initiative for Impact Evaluation, available at http://www.3ieimpact.org/media/filer/2012/05/07/Working_Paper_3.pdf.

White, H. and Phillips, D. (2012), 'Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework', International Initiative for Impact Evaluation.

Yin, R (2003), Applications of Case Study Research, Sage Publications.

Accompanying papers to the IOM (below) can be found at the FSDA website, www.fsdafrica.org/knowledge-hub/. These include:

Technical guidance notes

1. Impact-oriented measurement for FSD macro interventions
2. Linking FinScope/ FinAccess and poverty surveys
3. Methods for undertaking causality analysis
4. Logframes in the context of FSD and market development programmes
5. Tracking financial sector development: A practical note
6. FSD logframe indicators
7. Benchmarking financial sector development
8. Impact-oriented measurement: Frequently asked questions

Analytical background papers

1. Tracking financial sector development
2. The relationship between financial sector development, economic growth and poverty reduction
3. Assessing the quality of access

Annex A Definitions

Before studying the main elements of the approach to FSD impact evaluation it is important to ensure that there is a shared understanding of the definitions of various terms used throughout this document.

Table 33 Key terms and their definitions

Activity – something the FSD does with its inputs, be it providing funding, technical assistance, data production, convening power etc.

Attribution – the degree to which an observed change was caused by programme activities, as opposed to external factors.

Contribution – the demonstration of a plausible link between an observed change and programme activities, but without fully isolating the effect of the programme and other external factors.

End-users/ end beneficiaries – users of financial services, both consumer and enterprises, who are ultimately expected to benefit from FSD programmes. In market development programmes, use of words such as users/ customers is preferred over beneficiaries, as there is no subsidy expected at the core market level (i.e. between demand and supply of financial services).

FSPs – organisations delivering financial services direct to end-users/ customers (consumers and businesses). A range of formal, semi-formal and informal financial service providers exist to which this guidance is applicable.

Impact evaluation – an adapted OECD definition is used: 'positive and negative, primary and secondary medium to long-term effects produced by a development intervention, directly or indirectly, intended or unintended'. Impact on a range of variables is considered, not just the FSD programme's end objective of improved livelihoods for the poor.

Impact-Oriented Measurement (IOM) – a wide-ranging guide for FSDs seeking to improve their measurement of impact. It consists of a set of principles for measurement, and steps for setting up, implementing, and bringing together measurement of an FSD's impact.

Impact pathways – impact pathways refer to the causal logic and assumptions that link inputs to outputs, outputs to outcomes and outcomes to impacts. Results chains tend to be a diagrammatic expression of impact pathways.

Intervention – a defined set of programme activities designed to effect a particular change in a market. This may be undertaken in the form of a specific FSD-supported project, or as activities without funding (for example, informal discussions with regulators).

Logframe (or logical framework) – a management tool used to improve the design of interventions, most often at the programme/project level. It involves identifying strategic elements (inputs, outputs, outcomes and impact) and their causal relationships, indicators, and the assumptions and risks that may influence success and failure. It thus facilitates planning, execution and evaluation of a development intervention. (OECD-DAC definition)

Macro level – government strategies, policies, laws and regulations, and sector norms that affect the enabling environment for the provision of financial services.

Market actor – organisations or individuals who are active in the market system; covers not only as suppliers and consumers, but also regulators, policy-makers and others who can influence the way markets operate.

Market facilitator – an actor outside of the market system who is temporarily 'intervening' within it (as opposed to a 'market system actor'). Although catalyst is also used interchangeably, the use of market facilitator is preferred because unlike a catalyst, which only triggers change but does not change itself, a market facilitator has to assess market response and re-adjust.

Market-level – refers to the part of the market in which the intervention or the outcomes of the intervention have occurred. There are three levels: macro, meso and micro.

Market system – the multi-player, multi-function constellation of system actors engaged in a core exchange (such as the supply and demand of financial services), plus the set of 'support functions' and 'rules' which shape the quantity and quality of the core exchange. See M4P/FSD 'donut' (Springfield Centre, 2014). Often viewed by FSDs as consisting of macro, meso and micro components.

Meso level – sector infrastructure and services.

Micro level – the transaction level at the core of the market where providers and consumers of financial services interact.

Outputs – at a project level outputs indicate the direct results of an FSD's activities. At a programme level, these too should indicate results to which an FSD contributes. However, there is some variation in relation to FSDs' logframes in this regard, where some intermediate outcomes, such as underlying market system changes, are represented as outputs despite only being partly a direct result of FSD activity.

Outcome – the central purpose of an FSD's programme or project, indicating what is expected to change and who should benefit.

'Final' outcome – at an FSD's programme level this tends to refer to increases in financial inclusion for households and enterprises

'Intermediate' outcome – in the IOM theory of change, this refers to the underlying 'systemic changes' that lead to the final outcome (on a sustainable, resilient and large-scale basis).

'Direct' outcome – outcome brought about through market actors/projects directly supported by an FSD programme.

'Indirect' outcomes – outcome brought about through systemic effects (wider changes in market systems/ actors).

Partners – organisations (including government, private sector, NGOs) to which FSD programmes have provided some support. This may be in the form of advice and other TA and not just direct funding.

Project – a set of FSD activities centred around a common workplan and aim, with associated budget.

Programme – a set of FSD interventions as a whole.

Results – the changes catalysed by the programme in financial market systems and the consequential benefits for poor men and women (or participating organisations, employees). An observed change cannot be classified as a result until plausible attribution/ contribution has been established.

Results chain – a visual tool to show what the programme is doing, and why. Results chains clarify the 'logic' of the programme, by showing how activities will lead to outputs, outcomes, and eventually development impact. (DCED)

Stakeholders – FSD stakeholders include funders, government, regulators, market players, and clients. Their interests are represented through governing body/ project investment committees, which oversee the functioning of the FSD programme.

Systemic change – a transformation in the structure or dynamics of a system. The systemic change in which FSDs are interested is change that leads to impacts on large numbers of poor people, either in their material conditions or in their behaviour.

Theory of change (ToC) – every programme is packed with beliefs, assumptions and hypotheses about how change happens – about the way humans, organisations, political systems, or eco-systems work. The ToC is about articulating many of the underlying assumptions about how change will happen in a programme. (OECD-DAC)

Annex B Developing a programme ToC

There is much existing guidance on how to develop a ToC, and therefore this is not specified in detail here. One of our favourite resources is the book *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models* by Sue Funnell and Patricia Rogers. The DCED Standard also draws on this book in its results chain guidance.

As discussed, ToC is both a process and product. Funnell and Rogers (2011) suggest the following key questions to ask yourself when you are designing the process.

Funnell and Rogers (2011) also suggest three key activities to inform the development of a ToC. These are provided in Table 35, again with application to FSDs.

Table 34 Key questions for designing a ToC process

Question	Application to FSDs
Who should be involved in developing a programme theory, and in what capacity?	**Who will take a lead role?** FSD director, M&E adviser, external evaluator? **Who else will participate?** FSD staff, FSD board, funders, other partners, funded projects, external evaluator, financial sector experts? A**nd in what capacity?** e.g. source of information or participant in discussion?
What is an appropriate mix of approaches for developing a programme ToC, and how to go about it?	**Stakeholder mental models?** Which stakeholders to ask? In a group or individually? **Deduction?** e.g. market diagnosis and programme documentation, the wider research literature on financial inclusion; economic growth and poverty reduction? **Induction?** e.g. observe the programme in action; interview staff and other stakeholders.
How might workshops and interviews be used in developing the programme theory?[94]	**Workshop or interviews?** With whom? Or a combination of the two? **Which theory?** Should stakeholders be asked to describe how they think the FSD is supposed to work, or how it actually works?
When is it time to revisit programme theory?	**Strategy review?** More frequently – say every three years? At the time of M&E system development? When appointing a new FSD director? We recommend that, at the very least, the programme ToC be reviewed every year, and in-depth at the end of a strategy period.

Source: Adapted from Funnell and Rogers (2011)

Table 35 Key activities for developing a ToC

Activity	Further details	Application to FSDs
Situation analysis: understanding the problem, its causes, and its consequences	Identify the nature and extent of the problems or opportunities to be addressed; identify the known causes and consequences of the problem.	This is integral to the M4P approach and so FSDs will have already done this in their programme design. The complexity of market systems requires that intervention is guided by a good understanding of specific market systems, including a diagnosis of the symptoms and causes of underperformance. M4P interventions require a strong emphasis on information gathering and interpretation – not out of academic interest, but to shape design, ascertain prospects for sustainability and guide actions throughout the course of intervention.
Focus and scope	Set boundaries around the programme: which aspects of the problem – its causes and its consequences – are readily defined and acted upon; which are obscure but are within the scope of the intended programme; and which are outside an FSD programme's scope.	This could include: (a) what the role of the FSD should and should not be in terms of the impact it wants to achieve; (b) which market interventions the FSD programme is best suited to implement; (c) in regard to which market interventions is it most 'urgent' to achieve impact; and (d) which ones are likely to be the most important? Are these the same?
Outcomes chain	Show the hypothesised cause-and-effect relationships between immediate and intermediate outcomes and ultimate outcomes or impacts.	FSDs should also try, as far as it is possible (given the resources and levels of complexity), to identify as many as possible of the other major factors likely to affect outcomes and impacts, together with any foreseeable unintended consequences.
Focus on systemic change	Discussion should also focus on the types of systemic change that might occur. As Ruffer and Wach (2013) state, the ToC must clearly define the system to be changed.	Defining the financial system of interest and the assumptions regarding the ways in which this systemic change will be realised.

Source: Adapted from Funnell and Rogers (2011)

Annex C FSD indicators

The eight FSDs in Africa have similar ToCs, with most seeking market change, leading to financial inclusion and subsequent impact on poverty.[95] However, despite there being a fair amount of consistency in regard to objectives at different stages (i.e. impact and outcome statements) of their results matrices (i.e. the logframe), there are considerable differences in the indicators that FSDs actually use.[96]

Impact statements mostly focus on poverty and livelihoods, with a fair degree of commonality. In terms of impact indicators four FSDs[97] focus on national poverty levels – although there are doubts over how relevant an indicator this is for FSDs to measure, given that the difficulty in attribution jumps considerably from financial inclusion to country-level poverty reduction. There are also some significant differences in indicators used. For example, EFInA focuses on financial access at impact level (not poverty reduction), FSDZ measures the numbers of poor people experiencing an expansion in income opportunities or a reduction in *vulnerability* (i.e. not focusing on a change in poverty or income per se),[98] and FSDMoç includes an indicator for jobs created in the country. AFR includes an indicator for percentage savings to GDP, principally as a proxy for reduced vulnerability.[99]

Even at the **outcome** level, there are considerable differences. Four out of the eight FSDs use similar (albeit not directly comparable) indicators in terms of focusing on financial access. These include:

- the proportion of total adult population using services in formal (regulated) financial institutions (FSDK);
- the percentage of adult population using formal financial services (FMT and FSDT); and
- the number of poor people and microenterprises accessing a new financial service (FSDZ).

Three more FSDs have similar indicators but focus on direct FSD interventions, rather than country-level trends:

- the number of people in Rwanda using financial services as a result of AFR's interventions; and
- the number of adult individuals accessing financial services as a result of FSDU's interventions (disaggregated by type of product/service, gender and other categories).
- increased financial inclusion for poor Mozambicans and small businesses as a result of FSDMoç interventions (defined by first time usage of or using new financial products, disaggregated by type of product/ service, gender, urban/rural and other categories as required).

95. FSDA has been excluded from this analysis.
96. A full list of FSD indicators is provided at www.fsdafrica.org/knowledge-hub
97. FSDT had yet to define its impact indicators at the time of writing.
98. Although a reduction in vulnerability is a measure of reduced risk of entering into or increasing poverty.
99. Although the attribution to the FSD programme is a challenge.

Table 36 Indicators for individuals/ households and MSMEs[100]
The figures in brackets show the number of FSDs using this or a similar measure.

Individuals and households[101]

Access (8)	*Vulnerability*	*Cost of access*
% of adults using formal financial services[102]	% adults with reliable access to lump sum equivalent to one month's expenses (1)	Total average annual cost of running an account as % GDP/ capita (1)
% adults using informal financial services (only)	Increased volume of deposits mobilised from poor people by (supported) providers (1)	
% adults excluded		***Volume of credit***
		Increased volume of credit to poor people (1)

Enterprises

Access	*Vulnerability*	*Volume of credit*
Increased no. (M)SMEs accessing financial services (formal only or formal + informal) (2)	Increased volume of deposits mobilised from (M)SMEs by (supported) providers (1)	Increased volume of credit to (M)SMEs (1)
Increased no. (M)SMEs[103] accessing new financial services (1)		Increased aggregate (M)SME loan portfolios of formal providers (1)

The table illustrates the extent to which FSDs have, in practice, developed common outcome indicators for financial access of individuals/ households, even though there are variances in the actual measures used. FinScope-type surveys typically provide these measures. These indicators raise several questions and issues (see Table 37). The table also shows less common ground when it comes to access for enterprises.

100. NB: this does not include the second outcome indicator for FMT, i.e. number of legal, policy or de facto barriers to cross-border capital flows removed.
101. Some FSDs refer to households in some of their access indicators.
102. Access figures are usually disaggregated by gender; they may also be disaggregated by wealth/ poverty levels, e.g. quartiles or quintiles.
103. FSDZ disaggregates this figure to show increased number of smallholders using a new agricultural financial service.

Table 37 Review of FSD indicators

Question/ Issue	Answer/ Recommendation
If financial access in some form is the only truly common outcome indicator across FSDs, is it sensible and feasible to seek others?	In its business cases for FSDs, DFID regards financial access as a common binding constraint on financial sector development and on economic growth. Common financial access indicators for MSMEs (subject to agreeing a common definition) as well as individuals does make sense and should be feasible, especially in the light of DFID's increased focus on its growth agenda.
Should vulnerability indicators be used at the impact or outcome levels, or are both acceptable?	Vulnerability is an indicator of the risk of poverty increasing (or decreasing). As such it seems to be sensible to include such measures at the impact level, and to ensure that all FSDs adopt vulnerability measures in line with their impact statements (i.e. they should not make a statement that refers to reduced vulnerability and then fail to include an indicator for this).
Volume of credit	*This falls into two categories:* (i) Credit to MSMEs (from FSD-supported partners) – given the increased focus on growth, it seems sensible to measure this explicitly and include it as an output indicator. Two FSDs do this in one form or another. (ii) Credit to individuals – this may be important as an indicator, particularly in terms of poor people's ability to secure credit for income smoothing and to reduce vulnerability to economic shocks. It is therefore recommended to measure this at the outcome level. One FSD does this at present, but it should not pose much of a problem for others.
Cost of access	This reflects both the affordability of financial services and, albeit indirectly, market efficiency. It is thus a useful measure that can also reflect systemic change. However, the sources of this indicator are pricing surveys, central bank data or proxy indicators, such as FSP overheads. These might not always be available to, or available at an acceptable cost to, all FSDs. It is therefore not recommended that this be a common indicator for all FSDs. However, those that wish to and can feasibly adopt it should do so.[104] FSDK has recently changed this indicator to 'Cost of a KSh 500 electronic transfer across most widely used retail payment platform' and will be able to confirm in a few months whether this indictor works.
Should indicators be linked specifically to FSD interventions or measure overall changes, only some of which may be related to FSD interventions?	– At the outcome level, Mozambique, Rwanda, Tanzania and Uganda indicators specify links with the respective FSD interventions, whereas Kenya, South Africa and Zambia do not.[105] – The advantage of defining indicators with specific links is that this strengthens the case for the contribution that FSDs make. The disadvantage is that it may miss other factors that influence outcomes, but still relate to those cases where FSDs have intervened, especially where there are unintended outcomes that might be more clearly seen over a wider sample than just FSD interventions. – It is up to each FSD to work with their funders to decide whether the indicators should be limited to FSD intervention. However, we would recommend (either captured as part of the logframe or separately) that FSDs use indicators that are not limited to FSD interventions. However, in drawing up samples for measurement, FSDs include some cases where they have intervened, as well as ones where they have not.

104. At the overall sector level the 'tracking financial sector development' paper provides indicators of efficiency which may provide some indications of cost of access. See, www.fsdafrica.org/knowledge-hub

105. An update of Nigeria's outcome indicators had not been provided at the time of writing.

The final point to make here is that indicators should always be defined so that FSDs can measure change. Simply stating numbers at any point in time may fall short in this respect. At the outcome level in particular, it is also important to measure the effect of that change. Thus, an indicator that measures a given number of individuals or enterprises that have received support from an FSD does not capture the effect that support may have had (intended or unintended) on those individuals or enterprises.

Output indicators

Consistent with the common analytical framework that all the FSDs use in approaching market failures in the financial system, the broad themes of the outputs across the countries relate to:

Macro – Improvements in policies and strategies, along with legal and regulatory frameworks relating to the financial sector

There are three main types of interventions on which FSDs focus at the macro level, with corresponding programme indicators:[106]

1. Focus on changes in rules:
   – change in number of financial sector policies, strategies and activities that are aligned with appropriate international codes or standards;
   – FSPs, surveys or FSAP report improvements in regards to identified constraints in enabling environment (e.g. access to credit indicators/ microfinance business environment); and
   – number of policies/ regulations/ administrative procedures improved (supported by an FSD).
2. Focus on regulator/policy-makers:
   – change in and number of research products being used by policy-makers; and
   – improved capacity of policy-making bodies to formulate and implement effective financial sector policies and regulations.
3. Focus on groups interacting with regulators/ policy-makers:
   – working/ advocacy groups strengthened (as reported by both groups and central banks/ policy-makers); and
   – number of organisations demonstrating improved effectiveness in advocacy.

Meso – This can be divided into three main components:

1. improved capacity of financial institutions, including trade associations, platforms, and policy-makers to deliver appropriate products and services (e.g. EFInA and FSDU);
2. enhanced knowledge and information as a public good, e.g. FinScope and similar surveys of both individuals and small businesses; and
3. improved financial capability delivered through financial education (e.g. EFInA, and FSDMoç).

Micro – Greater financial access provided to poorer individuals and small enterprises. In several cases these two segments are treated as separate groups.

As expected at the output level, most indicators focus on 'how many' of a particular thing were achieved; for instance, number of people reached, or number of knowledge products disseminated. Given that most of these indicators are based on programme-level interventions, the source for the information is largely based on M&E system aggregation.

One key point is the link between outputs and outcomes, which is also discussed in the section on ToCs. While outputs like the number of workshops held or the number of times FinScope and similar surveys and databases are accessed and reportedly used are useful, they do not take the essential next step of assessing how FSPs and policy-makers use the information to make better decisions. While this is perhaps more of an outcome level measure, it is critical to assess what impact(s) such additional knowledge actually has in practice. A similar point could be made about the provision of financial education.

106. As adapted from FSD logframes.

Annex D Indicator sheets – systemic change

The tables in this annex show types of changes and related indicators that are general and not specific to a country context. They have not been presented as 'SMART' (specific, measurable, achievable, relevant and time-bound) indicators but are provided here for illustrative purposes to summarise a number of group sessions during the workshop. They are also not provided as a template that can be slavishly followed. Which type of indicator should be used, and in which column, is more of an art (i.e. requires the exercise of judgement) than a science.

Before finalising indicators, FSD teams may consider how these can be made SMART. As summarised below, each letter refers to a different criterion for judging objectives. Different sources use the letters to refer to different things. Typically accepted interpretations for SMART indicators are as follows:

Table 38 Developing SMART indicators

Letter	Common interpretation	Alternative interpretation
S	Specific	Significant, stretching, simple, sustainable
M	Measurable	Motivational, manageable, meaningful
A	Achievable	Agreed, assignable, attainable, actionable, action-oriented, adjustable, ambitious, aligned with corporate goals, aspirational
R	Relevant	Realistic, results-oriented, resourced, reasonable
T	Time-bound	Time-oriented, time-based, time-specific, time-sensitive, time-frame, testable

Source: Adapted from https://en.wikipedia.org/wiki/SMART_criteria

As noted in section 5.1.3, the IOM will need to track three types of changes: i) partner changes supported by an initial project; ii) partner changes beyond an initial project; and iii) broader market changes, i.e. signs that the broader sector is adapting and changing with emerging evidence of scale and breadth of change. This is where what has been a fairly FSD-centric monitoring perspective (bottom-up) connects with a broader sector tracking perspective. Evidence around these last two areas (e.g. partners' adopting, adapting, scaling up, replication, demonstration, crowding in effects, changing incentives of market actors resulting from structural change and changing resilience and responsiveness) can help articulate the narrative around systemic change.

For micro interventions see Table 22 in Section 5.2.3. The tables 39 to 41 below give examples of meso and macro level changes. Changes on the part of the partner as well as the wider market have been broken into during and after FSD support. However, the systemic change indicators cannot be seen as a linear progression and will be influenced by factors such as country context, diversity of projects, feedback loops across projects and dependence on many actors for timing and pace of change.

D.1 Supporting function/ meso: services

Table 39 Scenario: your FSD is supporting one product design service partner that aims to develop products appropriate for low-income consumers

Categories of change	Change in attitude, skills, knowledge and behaviour of partner	Market change occurred (often presented as an output in a FSD logframe)	Changes on the part of the partner beyond the initial project	Broader market changes
Detailed measurement question	How has the product design service (partner) demonstrated skills to meet the needs of low-income consumers?	Is a new support service established in the market for low-income consumers?	Will the product design service (partner) continue to provide design services to target low-income consumers after FSD support ends?	How many new product design services companies have incorporated design services for low-income customers in their long-term business plans? How has the FSD partner and those influenced by the partner contributed to this change?
Progress indicator/ change of interest	– FSD partner hires staff with skills in low-income market – Partner invests in training of staff and associate consultants – Partner carries out research on low-income market	Partner contracts with one or more FSPs to provide product design services focused on low-income consumers	– Partner develops good working relationships with FSPs (scale) – Credibility of product design partner in the market (adopt/scale) – Product design partner being hired by FSPs (scale)	– Other capacity service providers enter market (replication) – Increased number of new products on offer in the market (scale across market) – Increased competition among capacity service providers (scale across market) – FSPs develop internal units to deliver product design (respond)
Data sources	– Market research – Partner information on staff and skills – Periodic reports to FSD	Partner contract(s) interviews with FSPs which contracted partner	– Interviews with market players – Contract with capacity partner and FSPs	– Survey of design capacity service providers and FSPs

D.2 Supporting function/ meso: infrastructure

Table 40 Scenario: Your FSD is supporting the development of a collateral registry

Categories of change	Change in attitude, skills, knowledge and behaviour of partner	Market change occurred (often presented as an output in a FSD logframe)	Changes on the part of the partner beyond the initial project	Broader market changes
Detailed measurement question	Has the project helped change the attitude of the national stakeholders involved in the establishment of the collateral registry?	Is the collateral registry operational/ has it been strengthened?	Has the project helped the collateral registry to become more operationally efficient and accessible to FSPs?	Has the collateral registry increased the security of collateralised debt? Has this led to increased supply of credit by FSPs, especially to poorer segments of society?
Progress indicator/ change of interest	– Stakeholders agree on clear action plan for establishing collateral registry – Private and public sector agree/ sign up to action plan – Expert(s) hired to support process – Government and private sector invest in its establishment – Government allocates budget to the registry – Contractor appointed for the registry – Policy roadmap developed	– Collateral registry established – Number of members registered – Number of items registered – Number of queries submitted to the registry – Any regulation or policy change needed are identified and implemented – Registry creating awareness with FSPs	– Financing of registry improved (adopt) – Increased outreach to smaller FSPs (adapt/ scale) – Increased usage by FSPs (scale) – Usage costs falling as economies of scale develop (adopt) – Registry adapting to new market opportunities (adapt) – New channels for submitting information to registry (adapt)	– Volume of lending to new segments (incentive structural change /scale across market) – Increase in average SMEs lending (incentive structural change /scale across market) – New credit products developed to target new segments (incentive structural change /scale across market) – Perception of risk of lower-income segments by providers (incentive structural change /scale across market)
Data sources	– Policy paper – Minutes of meetings – Memorandums of understanding – Press releases – Contracts – Interviews with public and private players involved	– Registry MIS – National Financial Inclusion reports – FSDs own observations/ interviews – Press releases	– Registry MIS – Small business surveys – Annual report of collateral registry – National financial inclusion reports – FSD programme staff's own observations/ interviews – Press releases	Small business surveys Interviews with FSPs and MSMEs Annual report of collateral registry Regulator announcements Press releases

107. NB: in this example the FSD partner is the central bank and/ or MNO regulator.

D.3 Macro/rules and norms

Table 41 Scenario: Your FSD is advocating for a new mobile financial services regulation

Categories of change	Change in attitude, skills, knowledge and behaviour of partner	Market change occurred (often presented as an output in a FSD logframe)	Changes on the part of the partner[107] beyond initial project	Broader market changes
Detailed measurement question	Supervisors' and regulators' understanding of the business case for the provision of services in low-income and low-density areas?	Is the regulation in place and being effectively implemented?	Have the banking and MNO regulators learned from the experience, e.g. about the value of consulting with stakeholders, and the role of digital financial services in increasing financial access?	Does the mobile financial services regulatory framework incentivise actors – or remove constraints/ reduce risks and costs – to extend services to low-income and low-density areas? Has there been an increase in usage of mobile financial services by low-income people in low-density areas since the regulation was implemented?
Progress indicator/ change of interest	– Different regulators (banks, MNOs) collaborating – Regulator consulting with providers and other industry stakeholders – Regulators agree that change is needed – Regulations analysing/ using information about risks – Regulators commissioning analysis of risks and costs for providers – Regulators conducting diagnosis of access, usage and uptake – A new MFS unit set up by the government – AFI/ national commitment/ included in national financial inclusion	– New regulation and guidelines enacted – Effective implementation by regulator(s) – coordinating where necessary – Initial challenges faced by customers, agents and providers identified	– Regulators adapt in how they tackle constraints on increasing financial access and how to deal with innovation and new technologies (adapt) – Improved capacity to enforce MFS regulation (adopt) – Improved internal processes and decision-making relating to new regulation (adopt)	– Increase in access points in low-income areas (incentive structural change) – Lower entry barriers for new actors or new actors enter market (depends on market) (incentive structural change /scale across market) – Increase in number of entrants into the market (incentive structural change /scale across market) – Average number of transactions carried out by agents, especially in rural areas (incentive structural change /scale across market) – Private sector actors making additional investments, including in agent networks (respond) – Uptake of mobile financial services (number of active users and active agents increases) (incentive structural change /scale across market) – Changes in volume of transactions (incentive structural change /scale across market) Quality of service – Redress mechanism for clients operational (respond) – Satisfaction of clients with (a) service and (b) redress mechanisms Resilience – MFS product innovation [resilience/respond] – Regulator uses risk based approaches to supervision [resilience]

Table 41 (continued)

Categories of change	Change in attitude, skills, knowledge and behaviour of partner	Market change occurred (often presented as an output in a FSD logframe)	Changes on the part of the partner[106] beyond initial project	Broader market changes
Data sources	– Minutes of meetings between regulators and industry stakeholders – Press releases or news articles – Interviews with regulators and other stakeholders – Analysis/ reports – Minutes of FSD meetings with regulators, policy makers – Industry reports	– Announcements by regulators – FSP and MNO interviews – Regulator interviews	– Interviews with regulators – Interviews with FSP providers (FSD partners and others)	– Mapping of access points (financial inclusion reports or FSP maps) – Regulators' data + announcements/ websites – Interviews with stakeholders – Press releases/ government gazettes – IMF FAS – Information from providers – Industry tracking of price data – Reports on the state of the market

Annex E Adopt, adapt, expand, respond model

The adopt, adapt, expand, respond (AAER) model is a framework that can help to measure if a project is being implemented in a way that will contribute to systemic change: i.e. are systemic change mechanisms present? There is also a lot of overlap with the indicative indicators provided in Table 11 and Table 12 in the main report (and in the Annex D above), although the AAER provides a slightly different framework. In theory it can be used to measure all interventions, but it is much more suited to interventions at the micro level.[108] FSDZ currently uses this framework.

It can be used in three ways:

- As a useful sense check for monitoring how interventions are effectively contributing to systemic change: e.g. a quarter/bi-annual check to see in which quadrant you can identify evidence, and if interventions need to adapt.
- To build the evidence base for evaluation and annual reports.[109] For example, this year x number of projects exhibited adapt mechanisms, as shown by these results.
- Rather than use it as a single consistent framework, FSDs can also borrow those indicators they view as relevant to be used alongside a project's result chain.

Table 42 AAER indicators

Quadrant	Description	Measurement/ indicator[110]	FSD examples
Adopt **If you left now, would partners return to their previous ways of working?**	– New or improved product/ service offer, business model, or the uptake of a new role/responsibility	*Can be measured during the life of intervention (signs of adoption) as well as after intervention.* – Partner contribution to the pilot – Long-term viability (financial) /benefit of practice (inspires) – Partner satisfaction and intent to continue (e.g. ownership, future work plans, costing strategy etc.) – Partner ability to continue (e.g. financial, HR etc.) – Target group's satisfaction with, and benefit from (or at least signs of / or theory that it will benefit from) new behaviour from partners	– Financial institution increases sales (financial) – Financial institution targets new market segment (inspires)

108. Its design largely evolved from enterprise programmes that tended to carry out the majority of their projects at the micro level. While it is possible to apply the boxes at other levels (i.e. meso and macro), the frames of perspective change (for example, you do not want policy to be replicated, although you may want responses to that policy change to be replicated) and it becomes less relevant.
109. Interventions can ex ante specify what they expect to achieve in relation to the AAER model to help design the results chain, BUT it is a measurement rather than design tool.
110. See http://www.springfieldcentre.com/wp-content/uploads/2014/06/2014-03-Adopt-Adapt-Expand-Respond-Briefing-Paper1.pdf for a list of more in-depth indicators.

Table 42 AAER indicators

Quadrant	Description	Measurement/ indicator[109]	FSD examples
Adapt **If you left now, would partners build upon the changes they have adopted without us?**	– The partner acts independently of project support to continue, in some manner, with the role change and/or innovation that they originally piloted – The player is keen to improve the performance of the innovation further and works to tailor it, making investments that support its continuous and perhaps improved operation	– Independent investment improving changes beyond FSD project objectives (e.g. has new money been devoted to it)? – Has project changed (been made more efflcient; has more experimentation occurred; has it been extended to new areas)? – Has there been any behavioural change in relation to innovation in the partner? (e.g. new resources devoted to innovation / branding of innovation etc.) – Target group benefits sustained (i.e. change of direction still provides benefit to the poor)	– FSD partner organisation monitoring project beyond project agreement to better align with their objectives – Innovation within one part of FSD partner (dept.) has spread to another – FSD partner organisation moves into new markets
Expand **If you left now, would (significant) pro-poor outcomes depend on too few people, firms, or organisations?**	– A number of other market players have adopted the innovation, or clear variants thereof – BUT can also examine nature of system as well (to accommodate such expansion) – can be linked to any work FSD undertaking in this area	– Competitors or similar types of organisations are copying innovation[111] – Partner is scaling up, with innovation becoming mainstream and/or new business practices pushing innovation to scale[112] – Ability to accommodate competition or collaboration (depends on the nature of the system) *Nature of system* – Ease of entry for new players – Respect for rules/regulations/standards (e.g. adhering to voluntary/ industry codes of conduct and compacts, etc.).	– Market organisations copy new business practices of an FSD partner organisation – FSD brings two organisations together – FSD produces data to show demonstration effects that are taken up *Nature of system* – Are industry bodies playing a pro-active role in facilitating organisations to broaden outreach?
Respond **If you left now, would the system be supportive of the changes introduced (allowing them to be upheld, grow and evolve)?**	– The innovation triggers a secondary response from players in the wider system such as changed or new supporting functions and rules – BUT largely based on analysing how FSD projects at micro (and to some extent meso) level can lead to wider changes (i.e. not interventions directly related to changing rules and supporting functions)	– Supporting systems respond to help organisations engaging in FSD project area (e.g. new service providers emerge; rules have been adjusted; add-on products emerge) – Long-term resilience and sustainability of partners (and others copying them) to cope with shocks and 'move with the times' (e.g. change in economic circumstances; change in rules etc.)	– Regulatory bodies have changed or are considering changes to adapt to new products – New players are developing products (e.g. mobile payment add-ons) – How responsive is the market compared to comparative countries?

111. When working with less commercially-oriented partners, rather than looking at copying it may be more appropriate to consider how the influence of the project has 'spread' and if practices are being used elsewhere.
112. e.g. working with 'apex' players that have influence over a large number of relevant players.

Annex F Narrative analysis for the period (quarter/ half-year)

A number of methods can be used by FSDs to pick up signals from the market that are not easily captured in pre-defined indicators. Ideally these are not just shared amongst the team (to both triangulate evidence and improve programming) but also a record is kept to aid measurement. Table 43 notes a few of these, with some of their pros and cons. The template below the table suggests the types of issues that can be explored.

Table 43 Beyond monitoring methods

Method	Summary	Pros	Cons
Market scanning by FSD staff (written)	Project staff record observations in template	– Leverages staff's understanding of market – Written record	– Difficult to incentivise staff to do additional reporting
Market scanning by FSD staff (video)	Project staff record observations in videos (i.e. M&E officer interviews them each quarter)	– Leverages staff's understanding of market – Potentially more buy-in	– Labour-intensive to analyse video recordings
FGDs with market actors	Bring together a group of senior market actors (periodically) to discuss trends in the market	– Non-FSD perspectives – Relatively non-labour-intensive (if FSDs have contacts)	– Difficult to arrange (unless there are existing fora to leverage)
Media analysis	Collect and analyse financial sector media reports	– Leverage data already collected – Broad sector perspective	– Labour-intensive to analyse – May not be directly relevant

Narrative analysis for the period (quarter/ half-year – illustrative template)

This template is an illustrative of a example of a tool FSD can use to monitor changes that are not adequately captured by the set of traditional quantitative indicators. This checklist is an extension of the key indicators FSDs will typically use to measure progress in their projects and other interventions. An important source of information for this narrative analysis will be FSD staff, who can provide a narrative each quarter on what changes they are seeing in different characteristics of the system and then discuss and triangulate this with other FSD staff.

Area of focus	Description	Examples	Sources of data
Qualitative indicators	Measuring outcomes that are in the ToC (programme or theme) but are not easily captured by quantitative indicators or project result chains	Market players see value in continuing to offer new service and/or have plans to upgrade or roll it out to new market segments There is a 'change driver' – an institution or set of entrepreneurs driving the market change process. For example, a new player enters the market (with or without FSD support) with an improved business model to reach increasing numbers of poor people, and this in turn increases competition There are changing relationships – including competitive dynamics – amongst different market players (e.g. financial institutions and MNOs) Improved flows in learning and transfer of information (e.g. new fora or institutions that facilitate information-sharing) There have been changes in the overall business environment (e.g. financial inclusion regulation and policy) which enable more pro-poor businesses	Field observations Narratives from FSD staff Surveys Media monitoring FGDs Key informant interviews
Learning questions	Capturing key lessons learned and insights that have been prioritised as learning areas that can contribute to the sector or to the programme strategy	What are the main constraints in the market? What are the drivers of change in the market?	
Reviewing the ToC	Testing the assumptions, including the causal links, within the ToC. (Note collecting these data does not constitute an evaluation, but will help provide valuable information and insights for the evaluations)	Have the assumptions held true? (Review the specific assumptions of interest that were developed with the ToC) Have the causal links held true? (Review the specific causal links of interest within the ToC) Were there any unanticipated results or factors? Or anything that surprised you?	

Annex G Implications for FSDs of different approaches to impact evaluation and type of evaluation support

Section 7.5 notes that in addition to internal capacity, FSDs have three possible options to inject greater expertise and independence into their IOM approach. These scenarios, which are not mutually exclusive and have potential overlaps, are analysed in more detail below.

Table 44 Trade-offs in regard to independent evaluation

Category of analysis	Internal IOM systems (Option a)	Collaborative relationship with external learning partner (Option b)	Independent one-off evaluations (Option c)
Description	Rely on the IOM system to produce evidence, with the process of implementing the IOM tested by an independent evaluator	A broad/concurrent partnership between the FSD and an independent organisation(s) to assess the impacts of the programme, and make changes to programme design and monitoring systems to make it more evaluable, using both monitoring data and specifically commissioned impact research	Periodic external impact evaluations to assess if the programme is delivering expected results, e.g. at mid-point and end-point of the strategy, but using the evidence collected through the IOM
Ability to support real-time learning	High: monitoring data likely to be collected regularly and understood by FSD teams and partners. Relatively short time gap between data collection, analysis and consideration of possible changes in programme focus or features	Medium: Can help to strengthen monitoring system incrementally, but in the meantime external partner can help to produce periodic assessments	Low as only a couple of assessments likely (e.g. mid-term and end-term)
How 'impact' focused it is (compared to monitoring)	Low–medium: Based largely on monitoring system to provide learning and information for adaptation. However, if the IOM guidance is followed by the FSD team and there is some external validation, then impact information can be provided. This can also be augmented with additional evaluations	Medium: Can work with monitoring data to capture real-time changes but also develop robust evaluative techniques	High: Undertaken by independent evaluators with clear ToR that focuses on rigorous testing of pathways although dependent on quality of available data and willingness of FSD partners to share information (e.g. sensitivities around commercial data from the private sector)
HR required for FSD	Significant: Requires dedicated M&E capacity and buy-in from FSD management, theme leaders and funders	Medium: Will require additional M&E capacity with at least one senior expert within FSD to provide a focal point for a learning partnership	Limited: Up-front engagement around changes to the monitoring system will be required by FSD but then largely an oversight role (although will need some technical expertise to play this role). Some FSD support also needed in facilitating access to FSD partners

Category of analysis	Internal IOM systems (Option a)	Collaborative relationship with external learning partner (Option b)	Independent one-off evaluations (Option c)
Cost	Low–medium: Likely to impose some extra overheads (given staff costs) but generally monitoring data is relatively inexpensive Low cost for FSD partners	Medium–high: Will require some increase in FSD overheads plus cost of impact research Medium cost for FSD partners – both FSD and external agency seek information	Medium-high: High cost for each evaluation but as undertaken infrequently, the cost overall is spread out. High cost for FSD partners – external agency needs significant data and context in a short period (in addition to reporting to FSD team)
Additional data collection/ analysis required	Limited: Largely based on monitoring data although some external testing or additional studies may be required to develop annual assessments. Help may also be required to set up monitoring system	Medium: External partner can work with FSD to undertake or commission additional data collection/studies and analysis, but will also assist in developing programme-level monitoring systems	High: Significant conceptual understanding and data collection (e.g. quantitative and qualitative) required to test pathways
Independence of impact evaluation	Limited: In part, real-time IA will be carried out in-house with annual reports provided. However, these (and the data/ assumptions contained within them) could be verified by external consultants periodically through an audit	Mixed: While external agency will have a collaborative relationship with FSD, the advantage of appointing a learning partner (compared to full-time consultants) is that their independence is largely maintained. Independence can be enhanced through contracting arrangements, e.g. evaluators recruited by and reporting to a committee of FSD governing body rather than FSD management	High: Clear externally verified information with few incentives to present 'success' stories
Main advantages	– Highly operationally focused – Impact evaluation undertaken on FSD terms – Use FSD expertise regarding context and programme – Flexible (emerging experience in this area can augment this approach in the future) – Faster feedback loops and closer links between M&E and programme management	– Evaluators can build up familiarity with the programme – Able to substitute for a lack of FSD capacity – Flexible (as no best practice in this area) – Able to assist FSDs in wider knowledge agenda, including research products that go beyond IA	– High degree of independence – Access to outside expertise and perspectives – Less reliant on FSD capacity – Reduces bias of the promoters and implementers – Strengthens credibility of the findings
Main disadvantages	– Risk of bias (inflating success; reinforcing existing mental models) – Will not be viewed as sufficiently robust in terms of rigour, as not independent – Requires comprehensive monitoring systems and data – Risky as dependent on the skills, motivation and capacity of the FSD M&E staff and engagement with FSD technical staff	– Very reliant on picking a 'good' partner and clear scope of work/ working arrangements – Potentially may fail to build FSD's capacity, as dependent on others – Can be expensive – May lack ownership (and adaptive qualities) if inputs from evaluators are too thinly spread out over a long period	– Limited FSD ownership over findings (absence of learning and adaptation) – Evaluators may not fully understand the context – Reliant on availability of data, and if not adequately planned often reliant on largely secondary data – Less able to track long-term change given their one-off nature

About this guidance document

This assignment was commissioned by FSD Africa to facilitate peer learning among the nine FSDs in Africa, help them adopt more robust approaches, and develop a crisper message across the FSDs in regard to both measuring and reporting their results. This assignment has been facilitated by an OPM core team (Sukhwinder Arora, Sarah Keen, Ian Robinson, Robert Stone and Richard Williams). The OPM team was supported by a panel of experts including Thorsten Beck, Susan Johnson, Celina Lee and Alan Roe. The OPM team has also greatly benefited from frequent consultations with and guidance from FSDs, FSDA and CGAP teams. Contributions, especially from Mark Napier, Joe Huxley, Mayada El-Zoghbi, Karina Nielsen and Krisana Pieper are greatly acknowledged. Once this core assignment is completed by OPM in January 2016, FSD Africa seeks to work with DFID and the FSD Network in Africa to support its implementation and periodically review and update the guidance.

About FSD Africa

Financial Sector Deepening Africa (FSD Africa) is a non-profit company, funded by the UK's Department for International Development, which promotes financial sector development across sub-Saharan Africa. FSD Africa operates as a catalyst for change, working with partners to build financial markets that are robust, efficient and, above all, inclusive. It uses funding, research and technical expertise to identify market failures and strengthen the capacity of its partners to improve access to financial services and drive economic growth.

FSD Africa is also a regional platform. It fosters collaboration, best practice transfer, economies of scale and coherence between development agencies, donors, financial institutions, practitioners and government entities with a role in financial market development in sub-Saharan Africa. In particular, FSD Africa provides strategic and operational support to the FSD Network. FSD Africa believes that strong and responsive financial markets will be central to Africa's emerging growth story and the prosperity of its people.

About the FSD Network

Today, the FSD Network:

- Comprises two regional FSDs – FSD Africa based in Kenya (est. 2013) and FinMark Trust based in South Africa (est. 2002) – as well as seven national FSDs, in Kenya (est. 2005), Moçambique (est. 2014), Nigeria (est. 2007), Rwanda (est. 2011), Tanzania (est. 2005), Uganda (est. 2014) and Zambia (est. 2013);
- Is a world-leading proponent of the 'making markets work for the poor' approach;
- Specialises in inclusive financial sector development, through interventions such as SME finance, agriculture finance, housing finance, savings groups and digital financial services. A number of FSDs are starting to explore financial sector development for growth, through capital market development interventions such as secondary stock exchange development, capacity building and skills development;
- Represents a collective investment of $450+ million by DFID, the Bill & Melinda Gates Foundation, SIDA, DANIDA, Foreign Affairs, Trade and Development Canada, Royal Netherlands Embassy and the World Bank;
- Spends $55+ million per year, predominantly through grant instruments; and
- Employs over 100 full-time staff across sub-Saharan Africa and uses a wide range of specialist consultants.

FSD Africa, Nairobi, Kenya
info@fsdafrica.org
@fsdafrica

fsdafrica.org

Department for International Development
enquiry@dfid.gov.uk
@DFID_UK

gov.uk

Oxford Policy Management
admin@opml.co.uk
@OPMglobal

opml.co.uk