Annex 6

Generic Programme Evaluation Terms of Reference

Prosperity fund Evaluation and Learning

March 2018

Table of Contents

1 Terms of Reference Notes 4
1.1 Process for the development of the ToR 4
1.2 Purpose and scope of the ToR 4
1.3 Sections requiring updates 4
2 Introduction 5
2.1 Background to the Prosperity Fund 5
2.2 The E&L Service 5
2.3 The scope of the programme evaluation 6
2.4 The target programme(s) 6
2.5 Evaluation architecture and users 6
2.6 Principles and approaches (ownership, learning, co-ordination) 10
2.7 Evaluation and Learning Cycle 12
2.8 Work done to date 12
3 Evaluation Framework 14
3.1 The Fund Theory of Change 14
3.2 Programme Families 14
3.3 Programme Theory of change 14
3.4 Core Programme Evaluation Question Matrix 14
3.5 Programme evaluation management 21
3.6 Quality Assurance 22
4 Evaluation Methodology [Revised depending on the programme and year] 23
4.1 The overarching methodology 23
4.2 Programme Level Evaluation Methods 23
4.3 Data sources and quality 24
4.4 Family-level analysis and synthesis 25
4.5 Thematic Syntheses 27
4.6 Evaluation Ethics 27
5 Evaluation Cycle and Products [Revised depending on the year] 27
5.1 Timing and Sequencing of Evaluation Activities Years 2-4 27
5.2 Year 1 Programme Level Evaluation Activities (revised each year) 31
5.3 Annual Programme Evaluation Products 34
6 Learning 35
7 Resourcing requirements 35
7.1 Expected Level of Effort 35
7.2 Evaluation Team 36
7.3 Logistics 36
8 Limitations, Conflict and Risk Mitigation 36
8.1 Limitations 36
8.2 Risk Mitigation 37
8.3 Dispute resolution (TBC) 37

Document Title	

	[Document ID]	Revision	0
Revision	**Written By**	**Checked By**	**Approved By**
Revision 0	P. Abbot	M. Schultz	
Revision 1	P. Abbot	M. Schultz	

This report was prepared during the inception phase of the Prosperity Fund Evaluation and Learning contract. It forms an annex to the main Inception Report.

It was produced and approved by the Prosperity Fund Management Office before the main Inception Report and Workplan were finalised and agreed.

If there is any inconsistency between this annex and the main Inception Report and Workplan, the main Inception Report and Workplan provides the agreed position.

Terms of Reference Notes

Process for the development of the ToR

The generic programme evaluation terms of reference (ToR) were prepared in consultation with the E&L Service's technical team. It incorporates aspects from other papers prepared by the E&L Service during its inception phase that are available as annexes to the E&L Inception Report. It is informed by the interactions of the E&L Service's engagement team with Programme Managers, and the regular interactions with the central PFMO-MREL team.

Purpose and scope of the ToR

The ToR are designed to give a programme evaluation team the background to; and the scope and requirements for, a programme level evaluation. It describes the necessary interactions with other E&L team members and with the primary users of the programme evaluations.

Evaluations in years 2 through to 4 are expected to follow an annual cycle. More detail is provided for the year 1 baseline and 1st formative learning evaluation. This section will need to be revised and updated each year.

Reference is made to the package of materials for programme level evaluations. As of submission of this version of the generic ToR, these materials are still in preparation. The materials will be in place prior to the initiation of the evaluations and made available to all Programme evaluation teams. They will be considered a work on progress, iteratively improved based on experience and feedback from the evaluations.

Sections requiring updates

The generic programme evaluation ToR contains some sections that will be common across all ToRs and other sections (e.g. information on the programme of focus, or year specific instructions) that need to be developed for each evaluation. They are shaded in Blue.

1. Introduction

1.1 Background to the Prosperity Fund

Many developing countries, including middle-income countries where around 70% of the world's poor live, face challenges such as rapid urbanisation, climate change and high inequality, including gender inequality. Together, these constrain the prospects for long-term and sustainable growth.

The Prosperity Fund (PF) is a cross-government Fund that aims to reduce poverty by supporting inclusive economic growth in middle-income countries. £1.2bn has been allocated from 2016/17 to 2021/22. The Fund is managed by the Prosperity Fund Management Office (PFMO) which is hosted by the Cabinet Office (CO). More information is available at: https://www.gov.uk/government/publications/cross-government-prosperity-fund-programme/cross-government-prosperity-fund-update.

The primary purpose of the Fund is to create opportunities for growth that will reduce poverty and improve the welfare of poor people in partner countries. Priority areas include improving the business climate, competitiveness and operation of markets; energy and financial sector reform; and increasing the ability of partner countries to tackle corruption. As well as contributing to a reduction in poverty in recipient countries, it is envisaged that a secondary benefit of the Fund will be the creation of opportunities for international business, including UK companies.

The Fund is not a unitary programme but a portfolio of programmes. The Fund's Ministerial Board has approved 21 concept notes. In its first year, 2016-17, it committed £55 million in Official Development Assistance (ODA) and £5 million in non-ODA funds to a series of smaller projects across 13 countries and in the South East Asian region. During that time, it began developing the structures and procedures required for programming a higher volume of funding over the remaining four years of the spending review period.

1.2 The E&L Service

The Evaluation and Learning (E&L) service to the Prosperity Fund (PF) is provided by three firms, led by WYG, together with Integrity and LTS. We work alongside the Monitoring and Reporting (M&R) service provider, PA Consulting, and their partner, The Economist Intelligence Unit.

The purpose of the E&L service is to generate lessons that can be used to enhance the effectiveness of the PF at project, programme, programme 'families' and Fund levels. Evaluations will respond to Programme Teams' learning and information needs: providing evaluation evidence they need to do their job better. We will also be evaluating what has been achieved, how and why.

E&L works closely alongside M&R. While the focus of the Fund's M&R systems is on accountability (i.e. identifying and explaining PF spending and results achievement); E&L processes focus principally on learning. The roles are nevertheless related, with data and findings from M&R expected to feed into programme and fund level evaluations. For this reason, the E&L and M&R service providers have and will continue to closely coordinate our delivery and service designs.

The Evaluation and Learning (E&L) team for the Prosperity Fund (PF) will conduct a range of evaluations across programmes and at the Fund level with the purposes of supporting

organisational learning and of contributing to the accountability mechanisms of the Fund. This dual purpose will be achieved through both the way in which it is conducted and the evidence it will generate. An annual evaluation and learning cycle will drive evidence-based review of progress (on what has been achieved, how and why) and facilitate the sharing of, reflection on, and use of the evaluation evidence to support decision making at project, programme and Fund levels. Learning about what works (or doesn't), why and how, is expected to strengthen the interpretation of the progress monitoring and to improve reporting. In this way the evaluation can support adaptation and course correction at programme level; and feed into adaptive management at the Fund level to improve performance of the Fund. In this regard, the evaluation will deliver evaluation findings on progress toward, and the achievement of:

- Primary Purpose (ODA requirements as articulated through the provisions of the UK International Development Acts 2006, 2014 (Gender) & 2015);
- Secondary benefit through creation of opportunities for international business including UK companies; and
- VfM in the way resources are planned, managed and used to achieve the results.

The E&L team will collate, analyse and present high-quality evidence and knowledge of what works, why, where, and for whom, in achieving the goals of the PF. If successful, there should be demonstrable impact of this knowledge at all levels of the PF, including centrally, as well as within individual programme management units and within Whitehall departments working with the PF.

The E&L service will be designed to be responsive to emerging learning and evidence from the range of different PF stakeholders (from Prosperity Fund Management Office (PFMO) to Programme Managers (PMs)) through structured consultative processes at the different levels.

1.3 The scope of the programme evaluation

The programme evaluation should be user focused; focused on the evaluation and learning needs of the Programme SRO and team. The Programme evaluations should support the Programme team understand evaluation, and how it can be used in improving performance and effectiveness. The evaluations will assess project and programme design, but they should not be used to do design work. The E&L team will meet the demand for support from the Programme to improve and then operationalise their ToCs and business cases a tool for setting – and monitoring - realistic programme objectives, advising the Programme Teams where additional support may need to be commissioned. In the first year, this process will provide an early opportunity to deliver usable recommendations for adaptation and course correction, and flows well from the core evaluation questions of the Fund.

1.4 The target programme(s)

Provide a short description of the target Programme(s), for the ToRs and a link to available documentation (business case and other relevant documentation).

1.5 Evaluation architecture and users

1.5.1 The four different types of evaluation

Annual evaluation cycles within the PF will be managed through four different types of evaluation (Figure 1). Engagement with evaluation users at programme level will inform the

planning for the next annual cycle, including the selection of themes, the sampling strategy and timing. This will ensure that the evaluation findings and learning meets the needs of the Programme managers.

To avoid evaluation fatigue and overlapping data collection exercises, most of the information for the evaluations will be collected during the programme evaluations and from the M&R system (Prospero) that collects data about programme performance. This information will then be assimilated in the thematic, "families" and Fund evaluations.

The evaluations are described below and will be integrated, so that evidence collection and analysis, and stakeholder interaction is coordinated to optimise efficiency and minimise duplication risks:

1. Programme level where the main user is the Programme Manager (PM) and Senior Responsible Officer (SRO), with evaluation evidence also being drawn up and used in broader cross-programme learning either at family level or Fund level (see Table 1 which describes the users for Programme level and family-level evaluations). These will be co-designed with the Programme Mangers and their teams and are intended to have a strong learning focus and deliver rapid results to aid decision-making and active portfolio management.

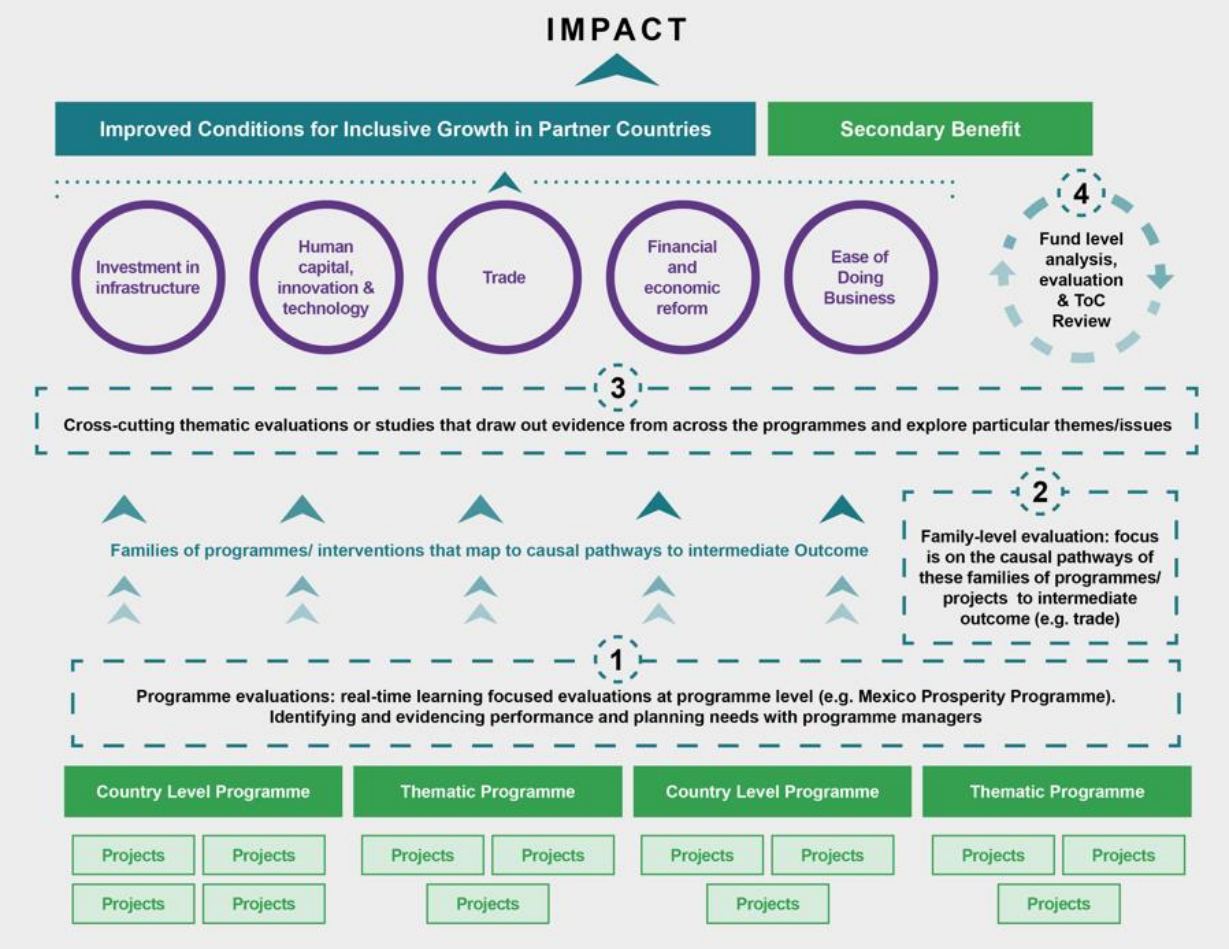

Figure 1: A Schematic of the four evaluation levels in the evaluation architecture

Table 1: Evaluation users and examples of the areas of interest in the evaluation

User	Examples of likely Areas of Interest
Programme Evaluation	
Commissioning Programme Managers	To validate or test plausibility of findings emerging from the monitoring data. To identify and capture specific learning issues related to projects or programmatic aspects across the programme or specific to one or more projects. How and when the evidence provided can be fed back to improve planning and implementation performance.
SROs	Plausibility of the results reported, validation of value for money, primary purpose and secondary benefits, potential links and synergies between the PF and other programmes they are responsible for, evidence to support the wider objectives of the SRO.
National (in-country) Stakeholders	To validate or test plausibility of findings emerging from the monitoring data, particularly the role and strength of PF in the changes observed. Lessons on what works that they may apply in future cooperation and partnerships.
Other Programme Managers	Lessons and best practice that they can introduce into their programmes.
PFMO	Plausibility of the results reported, validation of value for money, primary purpose and secondary benefits. Lessons what works where, why and best practice that can be shared through PF and beyond.
Other HMG Departments	Lessons of what works where, why and best practice that can be shared through PF and beyond to inform strategies for future work. How a cross-governmental mechanism can be used in ODA spending. Specific requirements related to their own mandate.
Other Development Partners	Lessons of what works where, why and best practice that can be shared through PF and beyond to inform strategies for future work.
Professional Communities	Best practice that may be developed tested and replicated elsewhere.
Family-level Evaluation	
Commissioning Programme Managers	To understand particular instances where there is strong learning potential that can be shared. Learning how projects within a family are working and the influence of country context, types of investments, sectors involved on how results are (or are not) being achieved; broader programmatic learning and opportunities for improvements
SROs	Plausibility of the results reported and comparison to progress and best practices in other similar projects or programmes in the PF
National (in-country) Stakeholders	To validate or test plausibility of findings emerging from the monitoring data
PFMO	Lessons and best practice: what works where, why, and what this means for devising strategies for ongoing or future management and design of the Fund Understanding how the observed changes are affected by country contexts, the institutional arrangements, and sectors, to achieve targeted intermediate outcomes. Understanding how these family of projects have been able to influence national and international processes to deliver value for money, primary purpose and secondary benefits
Other HMG Departments	Lessons and best practice: what works where, why, how the role of country context and investment type and sector in delivering results.
Other Development Partners	Lessons and best practice: what works where, why, how the role of country context and investment type and sector in delivering results. The potential for future collaboration between PF programmes and other development partners.
Professional Communities	Lessons and best practice: what works where, why, how the role of country context and investment type and sector in delivering results.

All programmes will be provided with support to undertake these evaluations focused on learning. A learning and knowledge management system is being developed by the E&L Service that will enable programmes to share lessons with each other. The learning generated through the programme evaluations will be brought together annually by E&L learning and knowledge management team and used to inform annual review and planning.

2. Family of projects where the main users are the PM, groups of PMs managing similar programmes or projects and; when synthesised and mapped to the Fund theory of change; the PFMO (see Table 1). The family-level evaluations focus on families of programmes or projects that aim to achieve the same or similar outcomes, work in similar sectors and contribute to the same Intermediate Outcome on the Fund ToC (see 2.2). They draw on the programme evidence, but there may be opportunities to undertake deeper analysis by way of case studies of particular projects with strong learning potential (e.g. performance lessons, best practice, management experiences) for the family. These will be captured as part of the programme evaluations and where possible integrated into the annual planning cycle (see section 1.7) to optimise synergies with any planned programme case studies. Cross-programme knowledge exchange will be promoted between projects within each family.

3. Cross-cutting thematic level where the main users are likely to be the PFMO and groups of PMs managing programmes of interest. Thematic evaluations will cover cross-cutting issues (e.g. gender, VfM, partnership building, capacity building) that are not specific to a particular casual pathway (such as trade, infrastructure) but which will contribute to cross-Fund learning and comparative assessment of performance, as well generate evidence on the Fund level management, governance and operational processes. As part of thematic evaluations, there may be opportunities to undertake deeper analysis through case studies of particular projects with strong learning potential for a thematic. These will usually be captured as part of the programme evaluations and where possible integrated into the annual planning cycle (see section 1.7) to optimise synergies with any planned programme case studies.

4. Fund level where the main users are likely to be the PFMO and the Fund's governance structures. These evaluations draw together findings from wider research and analysis, the family-level evaluations, thematic studies, programme evaluations and the case studies to draw conclusions and recommendations: i) in response to the evaluation and learning questions and ii) to test the assumptions regarding the pathways and envisaged mechanisms of change articulated in the Fund ToC.

The other potential users of programme and family-level evaluation findings (presented in Table 1 together with examples of the likely areas of interest) will need to be confirmed through a baseline stakeholder analysis at the start of each programme evaluation (see Section 0) and reviewed at each annual cycle.

1.5.2 The role of Programme Evaluation

The evaluations at programme level will play several functions, which means each evaluation will need careful coordination with the other evaluations during all stages of the annual cycle.

- The programme evaluations provide the PM and their SRO with a tool for generating learning about the performance of the programme; for testing the plausibility of the results and contribution claims reported by projects. The programme evaluations therefore need to be responsive to the needs of PMs' to understand their portfolio, and designed to provide information quickly to inform management decision-making, make course corrections and to support adaptation.

- The programme evaluations contribute to the accountability mechanisms of the Fund and there are a set of core programme evaluation questions (EQs) that will need to be examined across all programmes (see section 2.3). Whilst immediately useful to Programme Managers, these core programme EQs also allow for programmes to be compared and for information to be synthesised up for analysis at the family level.
- The programmes or selected projects of country programmes, can be grouped into 'families' of programmes/projects that map to output-intermediate outcome pathways of the Fund Theory of Change (ToC). This means the evidence and learning generated by the programme evaluations can be synthesised up to produce findings about that family of programmes/projects. This creates opportunities for cross-programme learning and knowledge exchange within peer-to-peer learning groups (see Section 5 below).
- As mentioned previously, projects with strong learning potential may be subject to more in-depth assessment undertaken as part of the family-level syntheses and/or the thematic studies. This includes examples where the PF influence on observed results is believed to be particularly strong and examples where signs indicate there are roadblocks to influence that offer learning opportunities.
- The evidence and findings from programme evaluations and case studies may also be used by the cross-cutting thematic evaluations.

1.6 Principles and approaches (ownership, learning, co-ordination)

1.6.1 Principles

The following E&L principles guide our overall approach to evaluation and learning. They also apply to the programme evaluations:

- Evaluation and learning activities will be guided by the evaluation framework and the underpinning Fund ToC. In line with the adaptive management of the Fund, the evaluation and learning approach will be reviewed and adapted annually based on evidence generated as part of the evaluation and learning cycle.
- As the central organisational structure guiding this evaluation, the ToC will be reviewed and updated annually based on evidence generated as part of the annual evaluation activities.
- Evaluation activities will have a learning focus in the sense that they will identify evidence to be used during planning and implementation (a formative approach) based on the information needs of stakeholders (user-focused).
- Interventions and activities will be based on 'do no harm' principles, promote women's active participation in the evaluation and incorporate ethical practices.
- The evaluation at the Fund and programme levels will incorporate gender and social inclusion-sensitive considerations.
- The evaluation will consider environmental and financial sustainability, drawing findings for programme delivery, and will seek to identify the PF value-add.
- Our learning approach will be based on a clear understanding of the different stakeholders, including beneficiaries and their differing monitoring, evaluation and learning needs. Findings will be fed back in inclusive ways.

- Evaluations will be conducted at appropriate times to feed into key decision timelines of the Fund and its programmes.
- Learning-focused programme evaluations will be tailored to the needs of the individual programmes and PMs but their results will also inform the overall understanding of PF performance. To ensure that these evaluations fit with both their primary and secondary purpose, the frameworks for each individual programme evaluation will be developed with the overarching framework in mind and specific tools for matching and 'translating' data and findings across will be developed.
- Methods will address Organisation for Economic Cooperation and Development (OECD) evaluation best practices and respond appropriately to the EQs, data sources, beneficiary feedback and PF Fund and programmes to minimise limitations of evaluation findings
- A robust multi-layered QA process will be applied to ensure the generation of high quality evidence. This will ensure the evaluation is governed by explicit principles covering: inputs (methodologies, people, resources); processes (oversight and governance, management arrangements, stakeholder engagement, communications); and deliverables (outputs).

1.6.2 Evaluation Ownership

In addition to these overall principles, it is important that the evaluation team understands that the programme evaluations are there to meet the demands, information and learning needs of programmes in the first instance. The family-level synthesis evaluations should meet the information and learning needs of the programmes represented in those 'families'.

The evaluation and learning cycle (Section 1.7) is designed to embed that ownership and direction from the programme teams. An important part of this ownership is the identification of programme learning or evidence needs and the development of questions that respond to these in each annual cycle.

1.6.3 The E&L team Coordination

Planning and conducing the programme evaluation will involve close collaboration with all other E&L activities, as these each relate to evaluating the change brought about by the PF in one way or another. There will be several programme evaluations happening simultaneously; with information from these feeding into several family evaluations; whilst cross-cutting thematic evaluations may require additional evaluative assessments on particular themes. All these evidence sources also contribute to the overall assessment of the Fund's performance and its contribution to primary purpose and secondary benefit.

Coordination of timing and mechanisms for engagement with the programmes and with family groups will be managed through dedicated E&L staff to support synergies and joint learning. The engagement with PMs should be as seamless as possible: the E&L service is a single service and that should also be the view from programmes.

1.6.4 Coordination with the MR Service

There is a sister service to the E&L Service, known as the MR (Monitoring and Reporting) Service, run by PA consulting. The MR Service is responsible for establishing an online reporting platform (Prospero). The PMs will be expected to enter data into it on a regular basis: They will select indicators and then report progress against these indicators on the platform.

The platform will display statistics and dashboard information to which the E&L Team will also have access. As the evaluation proceeds, the indicator results will provide the E&L service with the 'what' in terms of reported results. The evaluators will test the plausibility of the results, by asking the 'why', the 'how' and the 'how much' questions.

The E&L team has a focal point for interaction with the MR Service. The programme evaluation team should liaise with this person to manage interaction with the MR Service as whole.

1.7 Evaluation and Learning Cycle

The evaluation is applying a user and learning focused approach. This emphasises the value of stakeholder participation in design, analysis and follow-up to making evaluations useful. The E&L Team have put in place an annual evaluation and learning cycle that supports all its evaluations (Figure 2). This cycle will drive evidence-based review of progress (on what has been achieved, how and why) and facilitate the sharing of, reflection on, and use of the evaluation evidence to support decision making at project, programme and Fund levels. This ensures that evaluation activities will be designed to facilitate a learning process which responds to PMs' learning needs. The first annual cycle will differ slightly from subsequent years as programmes are still coming on stream and various baseline and early formative activities are planned (Section 4.2).

The E&L cycle comprises five steps and within each of those steps, there are learning touch points that give opportunity for the PMs and their workstream/component/project leads to learn through participation within evaluation processes. The programme evaluations follow this E&L cycle.

User participation in the evaluation will also build stakeholder relationships, awareness and ownership of evaluation objectives and support the use of evaluation results by the different users. These learning touch points should also be used across Family-level evaluations, creating opportunities for peer-to-peer learning points. The programme evaluation's learning approach is covered in Section 5 of this ToR.

1.8 Work done to date

Describe the E&L work done to date that provides relevant context and background reading. Reference to the package of materials for programme evaluation can be provided.

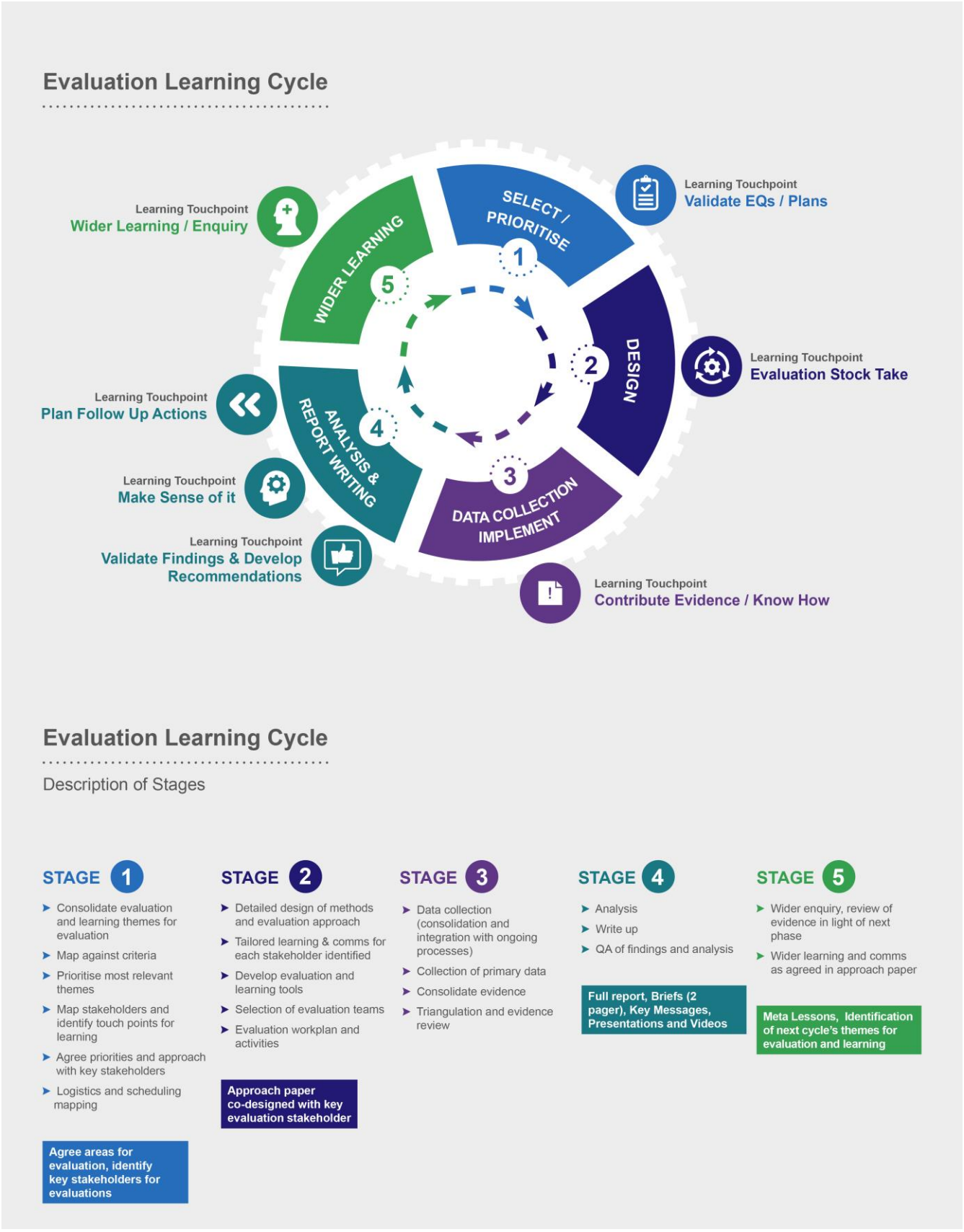

Figure 2: Annual evaluation and learning cycle

2. Evaluation Framework

2.1 The Fund Theory of Change

The Fund level ToC is the overall framework for the evaluation (Figure 3), within which programmes are nested (*i.e.* programme outcomes and impacts must align to the intermediate outcomes, outcomes and impacts of the Fund).

The E&L Service is taking a theory-based approach to the evaluation of the Fund and programmes. This means the evaluation will draw on and triangulate evidence from multiples sources, including the results monitoring, evidence emerging from the evaluation and learning activities and wider evidence bases to test and refine assumptions underlying the causal pathways from inputs to intermediate outcomes, outcomes and contributions towards impact. The ToC at Fund and programme levels are viewed in a dynamic way and causal pathways and assumptions may be modified and refined in the light of evidence from the evaluation.

However, more work is needed to flesh out the ToC at programme level and at the level of contracted projects that sit underneath and feed into different causal pathways at the Fund ToC level. This will be undertaken as part of the Year 1 baseline work.

See Fund ToC and the E&L Evaluation Framework in the package of *materials for programme level evaluations*.

2.2 Programme Families

An initial portfolio analysis has been made of the alignment of projects and programmes to the 'lower level' of the Fund ToC (inputs to intermediate outcome). These groups of projects/programmes have been termed "families" and form the sub-frame for the Family-level evaluations. Whilst these family groupings will require validation during the programme baseline work, they provide a basis for bringing together programmes and projects in similar sectors and towards similar PF outcomes for learning; and for synthesising assessment of causal contribution to the Fund's ToC. The families will be reviewed in the first year, as part of the programme evaluability assessments (see Section 4.2).

2.3 Programme Theory of change

Presentation and brief explanation of Programme ToC for respective programme(s)

2.4 Core Programme Evaluation Question Matrix

There are a set of core programme evaluation questions (EQs) that will need to be examined by the programme evaluation (see Table 2). These are expected to form a sub-set of the evaluation questions for each programme evaluation that are developed by the programme evaluation team; together with the PM; as part of the select/prioritise stage of each annual cycle. The programme evaluation team will prepare the evaluation matrix (see 3.2) based on the evaluation questions developed.

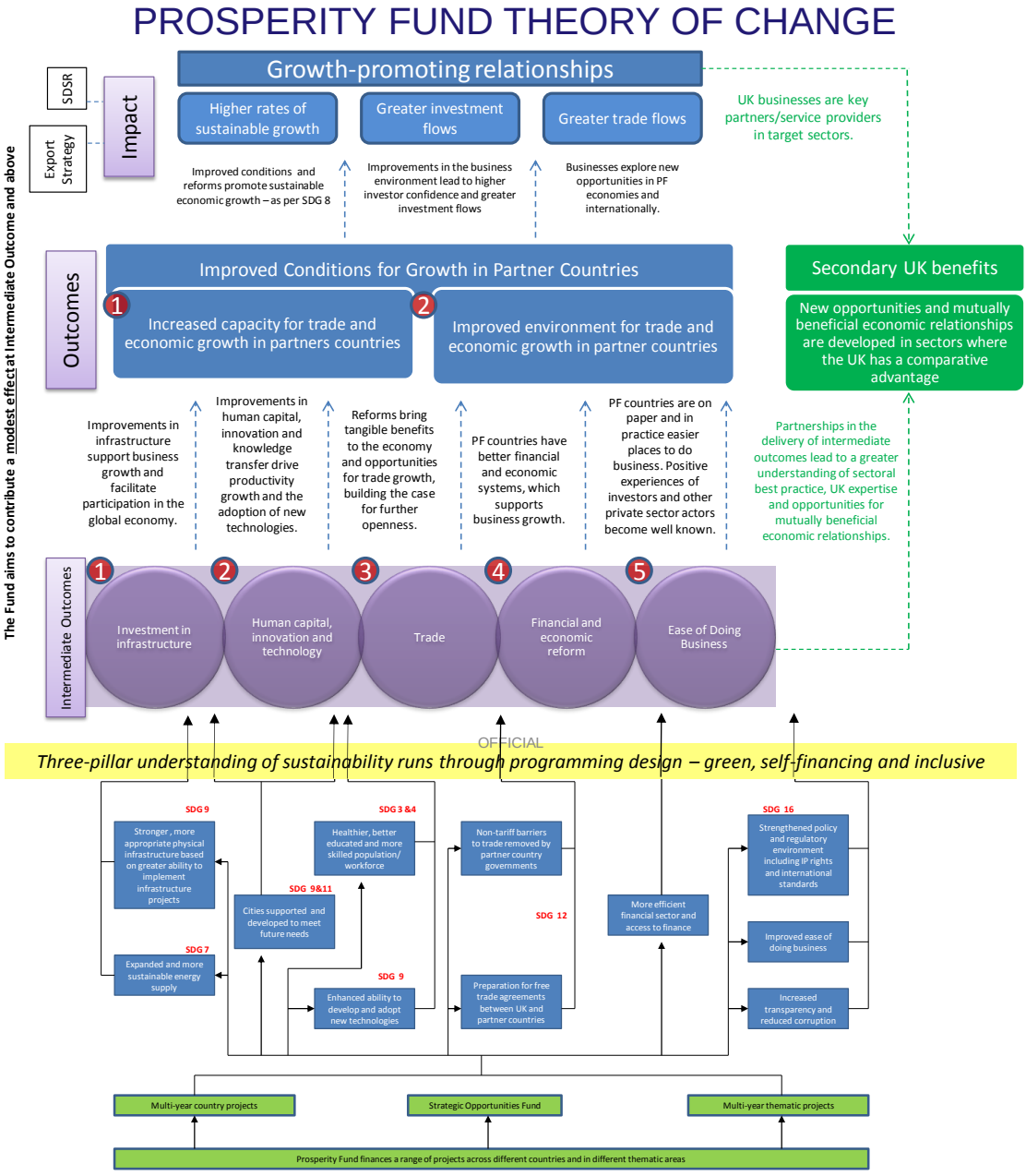

Figure 3: Prosperity Fund Theory of Change

Table 2: Core Programme Evaluation Questions

Column One	When	Types of information needed	OECD-DAC Criterion
What has been or is likely to be achieved as a result of the Prosperity Fund? (intermediate outcomes and outputs)			
EQ1 What evidence is there that the Programme or its projects are likely to contribute to the intended outputs and intermediate outcomes in the ToC, as well as any unintended or unexpected effects?	Year 1: Programme/Project ToC analysis for evaluability and context, define outcome and contribution claim. Combine with EQ2,3 Annually review the evidence from the evaluation to draw interim conclusions. Synthesise to tell contribution story at family level. Becomes more summative in years 3 and 4.	MR programme level indicators Records of the initial contribution claims and ToC assumptions Context analysis Data on outputs and outcomes observed, disaggregated by gender where appropriate Secondary data that can trace traces the contribution of the intervention to the recorded outcomes Data that establishes any influence of the Programme beyond the aggregation of specific interventions (secondary backed up by primary data for triangulation or validation).	Relevance Efficiency
EQ2 Which types of interventions and which sectors or country settings have been more and less successful in contributing to the achievement of primary benefits?	Year 1: Programme/Project ToC analysis for evaluability and context, define outcome and contribution claim. Combine with EQ1,3 Annual Family-level synthesis to draw comparative assessments between types/sectors and country settings. Final summative contribution analysis in year 4 at family level	MR programme level indicators Programme level primary outcome results evidence triangulated through primary data collection. Context analysis Data on outputs and outcomes observed, disaggregated by gender where appropriate Secondary data that can trace the contribution of the intervention to the recorded outcomes In depth case study data through primary data collection and/or detailed secondary data collection, processing and analysis; Information on external contextual factors influencing programme implementation (e.g. programme/institutional, contextual issues such as legislation, issues related to beneficiaries themselves).	Relevance Effectiveness
EQ3 Which types of interventions and which sectors or country settings have been more and less successful in contributing to the achievement of secondary benefits?	Year 1: Programme/Project ToC analysis for evaluability and context, define outcome and contribution claim. Combine with EQ1,2 Annual Family-level synthesis to draw comparative assessments between types/sectors and country settings. Final summative contribution analysis in year 4 at family level	MR programme level indicators Programme level secondary benefit outcome results evidence triangulated through primary data collection. Context analysis Data on outputs and outcomes observed, disaggregated by gender where appropriate Secondary data that can trace the contribution of the intervention to the recorded outcomes	Relevance Effectiveness Impact

Column One	When	Types of information needed	OECD-DAC Criterion
		In depth case study data through primary data collection and/or detailed secondary data collection, processing and analysis; Information on external contextual factors influencing programme implementation (e.g. programme/institutional, contextual issues such as legislation, issues related to beneficiaries themselves).	
EQ4 What evidence is there that the changes supported by the Programme's interventions will be sustainable and ensure environmental sustainability, will be self-financing and lead to inclusive growth that reduces inequality?	Year 1: establish baseline data against which inclusive growth can be measured and articulate causal pathways for reducing inequality / which beneficiaries are targeted, and how. Combine with EQ9 (equity and inclusive growth). Programme inputs into thematic study of sustainability. In Y1, this will cover aspects of design and process at Programme level (how programmes have designed sustainability into their projects, how programmes integrating issues of sustainability and inclusive growth into their risk management and review processes). Annual family-level analysis to provide learning on how sustainability and inclusive growth is being managed across the Fund. Final sustainability assessment using rubric and environmental and social impact assessment.	MR programme level indicators Evidence of likely environmental impact, processes to manage and mitigate environmental and social impacts Assessment of contextual (constraining/enabling) factors for a sustainability analysis. Evidence of mechanisms for self-financing of results post-PF	Sustainability Impact
What factors have contributed to or hindered these achievements?			
EQ5 What factors have contributed to the achievement of primary and secondary benefits?	Year 1: establish baseline contextual analysis for Programme ToCs to identify key influencing factors and assumptions related to these. Interim conclusions produced on basis of Annual programme evaluation findings. Annual family-level synthesis and barrier analysis to provide cross programme comparison at family level. Summative assessment in final year.	MR programme level indicators Contextual factors (e.g. programme/institutional, contextual issues such as legislation, issues related to beneficiaries themselves) - case studies will includes enabling and preventative factors as identified by ToC analysis Contextual Factors correlating with causal pathways identified as credible / evidence-based against a pre-defined rating system	Relevance Effectiveness Impact

Column One	When	Types of information needed	OECD-DAC Criterion
EQ6 How has the balance and relationship between primary and secondary outcomes across the Programme's portfolio influenced the achievement of results?	Year 1: Programme/project level ToC analysis to assess assumed interaction and causality between the two outcomes. Mid-point programme assessment fed up into Thematic evaluation. Final year programme assessment and fed up into thematic evaluation, allowing large pool for comparison and learning.	MR programme level indicators. Records of the initial contribution claims and ToC assumptions Context analysis. Data on outputs and outcomes observed, disaggregated by gender where appropriate. Data on how projects / activities / etc. rated (by EL) have had lesser or greater focus on primary or secondary benefits Studies on design and management decisions establishing how these decisions shaped the achievement of effects on primary and secondary benefits, drawing on analyses of programme contributions to these effects	Relevance Impact
EQ7 Which assumptions and causal pathways outlined in the ToC remain valid, which have been adapted and what refinements need to be made?	Year 1 baseline assessment of programme/project ToC assumptions as part of evaluability assessment. Annual review of ToC assumptions based on evidence from programme evaluation; informs annual programme level ToC review.	MR programme level indicators Baseline context and assumptions evidence Regularly updated and structured context and assumptions evidence Evidence of causal and affective factors including evidence of what interventions' contribution to primary and secondary benefits, disaggregated by gender	Relevance Impact
EQ8 To what extent is the institutional governance set-up of the Programme and its projects more or less effective in achieving i) primary benefits; ii) secondary benefits; iii) other results?	Year 1 baseline assessment of existing and planned governance arrangements for programme and projects. Feed into baseline thematic assessment of governance arrangements across the portfolio, by type, country context, and sector. Combine with EQ9. Institutional set up used as a factor in analysis of EQ5 evaluations. Finalised statements in final programme, fed up to fund evaluation.	MR programme level indicators Data / process mapping on governance and delivery mechanisms and processes at programme and project level, including stakeholder mapping, critical path analysis, analysis of roles and responsibilities including soft power, etc. and comparisons with other relevant funds Data on causal pathways and interventions' contribution to change - intended and actual - against primary, secondary benefits and disaggregated by gender and social inclusion (EQs 1-3, 5-7) Explanations for causes of change (especially those given by beneficiaries), where these refer to governance and delivery mechanisms and processes.	Relevance Effectiveness Efficiency Impact
EQ9 What types of approaches, governance and management arrangements have been more and less effective for achieving results and demonstrate good approaches	Year 1 baseline assessment of existing and planned governance arrangements for programme and projects. Review existing and planned management arrangements for managing VfM. Combine with EQ4, 8.	MR programme level indicators Data on programme outcomes and outputs disaggregated by gender and social inclusion Programme and project expenditure data.	Effectiveness Efficiency Sustainability

Column One	When	Types of information needed	OECD-DAC Criterion
to supporting inclusive growth and VfM?	Feed into baseline thematic assessment of capacity to manage VfM across the portfolio, by type, country context, and sector. Annual Family-level synthesis, to allow comparison between projects within intermediate outcome. Final assessment at programme level against results assessment; synthesised across family to allow comparative assessment within the intermediate outcome. Thematic synthesis to allow comparative by type and country context.	Description of programme management cycles and fund governance as described in programme / project and fund documentation (for a sample of progs / projects and the fund). Indicators of inclusive growth (TBC) and VfM (TBC), as measured as standard per programme (e.g. as part of the disaggregated Core Indicators or in other programme monitoring), disaggregated by programme type (sector, approach) and by governance/management arrangement type (potentially including cost per VFM relevant output/outcome indicators from programmes/fund log frame). These should have been identified in response part of EQ4. Description of programme approaches, governance and management arrangements (VFM1 (process). Existence, relevance & robustness of VfM Indicators; VFM2. (economy) Procurement & other strategies in place to manage costs; VFM3. (efficiency) Efficiency of PF Interventions ("Measure extent to which outputs are cost-efficient, timely"); VFM6. (effectiveness) Strategies & measures adopted to enhance delivery & mitigate risk) Systems developed and communicated for measuring and utilising VFM data Evidence of good practice approaches to inclusive growth and VfM from other programmes (benchmarking)	
EQ10 To what extent has the Programme and its interventions contributed to results that support gender equality, women's economic empowerment and social inclusion in line with the UK's Gender Equality Act and the Prosperity Fund Policy and Guidance and the Prosperity Fund Gender and Inclusion Framework?	Year 1 Initial baseline assessment of gender accountability processes and how gender has been considered in programme/project design. Feeds into Fund level assessment. ToC baseline assessment of causal assumptions for gender equality, empowerment and social inclusion. Using case studies, annual programme evaluations capture contribution claims related to gendered impacts and benefits. Final programme evaluation contribution analysis tests causal assumptions for gender equality, empowerment and social inclusion.	MR programme level indicators Analysis of gender differences in the achievement of gender and inclusivity outcomes across different interventions, sectors and activities and mechanisms of change, contextualised by the baseline and any additional local data on gender and inclusivity (see below), with interventions' contributions unpacked and contextualised within other success factors' and barriers to project implementation and project achievement. Information on gender inclusivity might include: Existence of gender-responsive infrastructure contributed to by PF interventions; Women's access to services (e.g. health and infrastructure); Gender/inclusion-sensitive business processes/policies (e.g. mat.leave); Women-owned SMEs with higher productivity/income/links to bigger businesses etc.; Women's participation in labour market/income generation; Women's labour market participation in different sectors/roles (reduction of gender-based occupational segregation); Women's roles in businesses/	Relevance Impact

Column One	When	Types of information needed	OECD-DAC Criterion
		in the workplace (where in the hierarchy); Women's access to income and productive assets	
What can we learn from the Prosperity Fund experience to date to improve on-going and future programming?			
EQ11 What kinds of learning processes are evident within and across programmes	Year 1 as part of 1st formative evaluation of the programme, assess processes in place to capture and use evidence for organisational learning and active portfolio management. Incorporation into annual learning and reflection points.	Units of 'evidence' resulting from evaluation activities (e.g.TOC workshops, KIIs, facilitated learning events) designed as part of the various PF evaluations during implementation. Description of programme management cycles and fund governance as described in programme / project and fund documentation (for a sample of progs / projects and the fund). Sources of information on learning processes, engagement with learning and action upon it (e.g. results of TOC workshops, KIIs, facilitated learning events).	
EQ12 What Programme level learning is sufficiently robust for wider circulation between programmes?	Incorporation into annual learning and reflection points.	Lessons learned as described in Fund/programme / project documentation Lessons and underlying evidence identified through PF evaluation activities Criteria of external validity described in PF evaluation framework	

The core programme EQs are derived from the Fund level EQs. The Fund level EQs will be reviewed annually based on feedback from a range of PF stakeholders to reflect the information needs of the Fund at different levels, and revised if necessary. Changes in Fund-level EQs will have to be reflected in updates to core programme EQs.

By having a set of core programme EQs, this allows PMs to share common questions, which is valuable for comparative learning purposes. It also enables evaluation information to be synthesised up for analysis at the family level.

2.5 Programme evaluation management

2.5.1 E&L team roles

Each programme evaluation will be led and managed by a *Programme Evaluation Lead* (PEL) who will be responsible for the team delivering a specific evaluation.

The PEL will interact with a range of other members of the E&L service, so that the work is organised, efficient and unitary at the point of delivery. This is described in Table 3 below.

Table 3: Key personnel and their roles in the E&L Service

E&L Personnel	Role and likely interaction points for the Programme Evaluation Lead (PEL)
Portfolio Leads	They are responsible for organising and managing family synthesis along a particular causal pathway family. They will ensure consistency in data collection and analysis, ensure evaluators properly briefed / trained. The PEL will report to a designated Portfolio Lead. The PEL will work with the Portfolio Lead(s) to: identify and plan integration of any relevant family evaluation aspects into the annual programme evaluation cycle; ensure the evaluation team are properly briefed on necessary instruments; and participate in the synthesis of programme evaluation findings for use at the family level.
Thematic Lead	They oversee thematic evaluations, but also ensure other evaluations cover important thematic issues relating to accountability – e.g. the '3 pillars', Gender etc. The PEL will work with the Thematic Lead(s) to: identify and plan integration of any relevant thematic aspects into the annual evaluation cycle; ensure the evaluation team are properly briefed on necessary instruments; synthesis programme evaluation findings for use in thematic analysis.
E&L Technical Lead (Programme & Portfolio)	The E&L Technical Lead (Programme & Portfolio) has oversight of all Programme and Family-level evaluations being undertaken. The PEL will engage with the E&L Technical Lead (Programme & Portfolio) to discuss and agree the evaluation and learning approaches and methods being used in the assignment.
E&L Team Leader	The E&L Team Leader has overall management responsibility for the E&L client deliverables. The Programme Evaluation Lead will engage with the E&L Team Leader through the joint planning and learning meetings of the E&L team. The E&L Team Leader will ensure coordination between the evaluation and learning/knowledge teams in the E&L service. The PEL will interact with the E&L Team Leader for the internal QA of the programme evaluation products.
E&L Learning & Knowledge Management Lead	The E&L Learning & Knowledge Management Lead oversees the knowledge, learning and communications work of the E&L Service. The PEL will interact with the E&L L&KM Lead for training/guidance in learning approaches, in the preparation of evidence for use in learning processes and products and for supporting facilitating programme learning from evaluations.
Programme Evaluation Team	The PEL will lead the Programme Evaluation Team (PET) comprising several evaluators / sector experts, depending on the size of the respective programme and its technical or geographic focus. Country programmes are likely to include in-country experts on the team to aid ongoing interaction and improved understanding of contextual, data and stakeholder issues. Each programme evaluation will have a designated contact point from the E&L core team as part of the PET who will support the PEL in planning the evaluation, and incorporating any necessary case studies or thematic work into programme evaluation cycles in consultation with the PM.

2.5.2 Reporting Arrangements

E&L

Within E&L, the Programme Evaluation Lead (PEL) for Programme [add name] will report to the Portfolio Lead for [add name].

The Programme Evaluation Team (PET) for the [name] evaluation report to the PEL [name].

The Programme

The evaluation is undertaken for the programme and reports to the SRO for the programme. The evaluation products are prepared for the SRO, the Programme Manager and the Programme Team. For this evaluation that is [add titles, names].

The PET will need to interact with programme teams at various points throughout the evaluation cycle – as part of the planning and management of the evaluation; as well as part of the evaluation, for example identifying evaluation needs, interacting over information sources, contacts and key informants, discussion of findings. Designated contact points for each [project or equivalent] under the Programme will be identified from the Programme and from the PET to facilitate this process.

2.5.3 Evaluation Governance

The PET is a 'critical friend' to the programme: It will be expected to operate as the Programme Manager's evaluators, and the PET will need to work closely with the Programme Manager and any delegated staff throughout the evaluation cycle. The programme evaluation should set up a governance structure to oversee the evaluation process and assure its independence. This should be agreed during the initial interactions with the PEL and Programme Manager and their teams. The details of this are likely to be specific to each evaluation, and it should not be an over-burdensome instrument. However the following aspects should considered:

- Chairpersonship of Governance structure (likely to be the SRO or their delegated representative).
- Membership of governance structure.
- Terms of Reference: Frequency of meetings, quorate, role in strategic oversight of evaluation process and evaluation products.

The Approach Paper should set out the expectations for the interaction between the PET and the Programme Team, describing the expectations of involvement of each party.

2.6 Quality Assurance

The programme evaluations and Family-level evaluations will be subject to internal QA processes.

The E&L service has an approach to Quality Assurance of evaluation products and evaluation process that is set out in a separate Quality Assurance Approach Paper. This is available in the package of *materials for programme level evaluations*.

3. Evaluation Methodology [Revised depending on the programme and year]

3.1 The overarching methodology

The E&L is taking a theory based approach to the evaluation of the Fund and its Programmes. Within this approach, the E&L service will use contribution analysis to assess contribution claims of the Programmes and the Fund. Contribution analysis approaches are good at answering 'what', 'how' and 'why'.

Such methods will help cover a clear identification of the theory underlying the Programme's role in anticipated and later observed change, including hypotheses about how change occurs and the Programme's role in it.

It is not expected that econometric methods, primarily experimental and quasi-experimental design, will be suitable for questions about policy influence and complex change under uncertain situations and long-term time horizons.

The Technical Lead (Programme & Portfolio) will convene regular E&L working meetings to which PELs will be invited, so that emerging plans on methods can be shared, discussed and critiqued.

3.2 Programme Level Evaluation Methods

During the design phase of each annual evaluation cycle, the programme evaluations will be jointly developed with the PM, based on the results monitoring and learning needs identified across the [Name] Programme's portfolio of projects. The aim is for all programmes to be supported with these learning focused evaluations. The learning findings will be synthesised annually and shared through the knowledge management system to be developed.

In line with the user and learning focus, the identification and selection of programme evaluation methods that these evaluations apply should consider the needs of the evaluation prime users and the comparative strengths and weaknesses of different methods in terms of how they can be used[1]. Once the information needs of programmes have been drawn out and evaluation questions devised to answer the necessary evidence or learning, the PET should set out each method's:

- Ability to answer a series of specific evaluation questions
- General ability to fulfil a series of tasks or reach specific goals
- Demands on the capacity of the Programme Manager and the evaluation team to accommodate its specific requirements.

This process will enable evaluation stakeholders and the E&L service to jointly refine, clarify and articulate the reasoning behind the choice of methods in a logical and structured way and in so doing build their understanding of evaluation.

By taking into consideration the type of EQ, the PET should be able to identify which methodologies are best suited to answering the EQ. Questions such as 'how much', 'what',

[1] Following Befani, B. (2016) Choosing Appropriate Evaluation Methods: A Tool for Assessment & Selection. Bond, London; and Stern *et al* (2012) Broadening the range of designs and methods for impact evaluations. DFID Working paper 38.

'how' and 'why' require different methodologies, as statistical approaches are good at answering 'how much' questions but not 'how' and 'why'. The features of the projects and interventions should guide choices on methodology at the programme level. As for some interventions, the attributes of the intervention may make it difficult or impossible to use some methodologies: The PEL should designate one to two members of the PET to conduct an evaluability assessment to guide methodological choices (see 4.2 below).

The PET should explore with the Programme Manager whether there are any other areas of interest; beyond just looking at the EQs; that the evaluation should cover.

The PET should prepare an evaluation matrix, which sets out the preferred methods to be used. This can be discussed with the Programme Manager and should be included in the evaluation Approach Paper.

3.3 Data sources and quality

3.3.1 Baseline secondary data assessment

Desk research using secondary data sets will be a key information source for the programme evaluation. The baseline assessment included an assessment of available secondary data for answering the core programme evaluation questions. This should be referred to and built annually as the additional evidence is added to our evidence base and the utility of secondary data sets is tested. It is available in the package of *materials for programme level evaluations*.

3.3.2 Programme and Project Indicator data

The indicators developed for programme and project enable the projects to report against the indicators and milestones set out in project logical frameworks. The indicator reporting will provide a valuable tool; at family level they may be used for identifying trends and patterns, and for identifying areas of strong learning potential that will inform evaluation planning at the select and prioritise stages.

3.3.3 Primary Data Sources

The exact primary data requirements will depend on the EQs and methods chosen. However the evaluation is expected to involve new data collection through methods such as interviews, surveys to validate or triangulate the secondary data or to fill information gaps. The programme evaluation team will provide analysis of both potential methods and focus areas for primary data collection, *i.e.*, which themes, countries and/or projects merit deeper analysis, as part of the design stage of their work.

3.3.4 Data quality control

All pieces of evidence that are collected including secondary data and from stakeholder interviews should be assessed for the "strength of evidence" following the categories listed in the rubric outlined in Table 4 below. This will generate a score for each piece of evidence. This will enable an assessment of the strength of evidence of each finding to be clearly presented based on the aggregated results of the strength of evidence supporting that finding.

The evaluation should seek to triangulate evidence across multiple plausible evidence sources to have greater confidence in the findings. The potential limitations in analysis and findings, as a result of weak evidence should be clearly explained in the evaluation product.

Table 4: Strength of Evidence Rubric

"Strength of Evidence" Assessment for each evidence collected		
2	Verifiable evidence	Refers to data that are both plausible and possible to verify. Such evidence generally describes quantifiable measures that can be physically counted. For example, the MW rating of installed energy capacity or the number of jobs in a company at a given time.
1	Plausible evidence	This includes evidence which may make a plausible claim but may draw heavily on assumptions from secondary literature, for example those used to calculate greenhouse gas emissions avoided. Alternatively, it may refer to evidence which is the plausible conclusion drawn by an expert stakeholder or observer. There may be evidence presented to justify this view but no methodology against which the validity of the conclusion can be verified.
0	Minimal evidence	Some documents may simply claim an outcome but there may be no information about the data or methodology used to evidence this claim.

3.4 Family-level analysis and synthesis

In year 1 of the E&L Service the projects and programmes[2] will be allocated to 1 (or more) families. This provides an opportunity for answering evaluation questions around input to outcome causality within a family and for facilitated learning across programmes and projects with similar outcomes. The E&L Service has appointed Portfolio Leads to lead each family level evaluation.

The Portfolio Leads will engage with all PELs as part of the 'prioritise and select'; 'design'; 'data collection and analysis'; and 'report writing' stages of the family-level evaluation (Figure 2). It will be undertaken jointly with the Thematic Leads, to ensure the work is properly co-ordinated and the links between the different parts of the evaluation are joined up. The synthesis strategy is available as an Annex to the E&L Inception Report.

3.4.1 Prioritise and select engagement

The respective Portfolio Lead will identify the information requirements from programme evaluations for the purposes of family-level synthesis. This includes the selection and prioritisation of Evaluation Questions to be addressed in the family-level evaluation. This should draw on inputs from the Secondary Benefits, VfM and Gender Thematic Leads, given the cross-cutting nature of those fields. Where practical, the PEL should build these requirements into the programme evaluation so that the information is generated in one evaluation process managed by the PET.

There is a need to ensure consistency across the various evaluation teams assembled for specific assignments. The Portfolio Lead will convene an induction meeting for the PEL's within their management to ensure consistency in the use of evaluation methods and instruments for those EQs that will be synthesised under the family evaluation.

The sampling of projects for case studies under the family-level evaluation will be purposive. The criteria used to select samples is likely to vary between families, depending on the key design attributes and contextual factors of the programmes and projects in that family. PEL's

[2] Thematic programmes (e.g. trade) and projects of country programmes (e.g. a trade project of a country programme) may be grouped under a 'trade' family for family-level analysis as they contribute to the 'Trade' intermediate outcome of the Fund ToC.

will identify potential sample projects together with the PMs. The final selection of samples will necessarily be a negotiation between the PEL's, PMs and the Portfolio Lead; leading to an optimum selection of samples. The types of criteria likely to be considered are:

- Learning Potential: in line with our learning focus, the selected cases should have potential for providing useful learning to the programme teams participating in the family. Those that have a strong demonstration effect (such as early start or particularly successful projects) can provide useful lessons for other projects in the family.
- Policy Environment: PF may operate in countries with either a strong or weak regulatory and policy environments supporting the target sector. Private investment will be more challenging in countries with weak or limited policies and regulatory contexts. Cases that reflect different policy context settings will allow a broader assessment of the contextual factors affecting the causal pathway.
- Coverage: the case studies should be representative of the family and together capture the breadth of contexts and mechanisms observed within the causal pathway for that family.
- Quality of evidence: an obvious criterion: those projects that will not be able to provide data that are plausible and verifiable will be less useful to the contribution analysis and less likely to be selected as case studies.

Based on the resources available to the E&L Service, it is expected that 4 to 6 project case studies will be purposively selected per family grouping. The selection of case studies and the detailed data collection and case study synthesis methodology will be prepared in the approach paper for each family-level evaluation.

3.4.2 Design engagement

The PEL will be responsible for the design of the programme evaluation, in consultation with the PM, considering the needs of the family level evaluations. Some of the family evaluation requirements may speak to the immediate information needs of the PMs whilst others will be tailored to answer the EQs at family level, and therefore must be consistently designed across the family. Additional resource requirements and organisation of the team to enable programme evaluation to work in a coordinated manner will be incorporated into the design.

The Portfolio Leads will be responsible for designing the family synthesis evaluations that responds to the EQs, as articulated in the context of an individual family and its technical specificities. The Portfolio Leads will be in regular contact with the respective PEL during this process.

3.4.3 Data collection engagement

During the data collection, the Programme Evaluation Lead will have light-touch engagement with the Portfolio Lead; the latter playing an oversight and trouble-shooting role in ensuring the consistent use of data collection instruments, coding procedures etc.

3.4.4 Analysis and Report writing engagement

The data from the programme evaluations will be assessed for relevance to the family synthesis and where relevant will be integrated into the synthesis work. During the data analysis and report writing stage, two levels of sense making will be necessary for the family-level synthesis:

1 Sense-making between evaluation teams, where findings are compared and discussed, probably through a sense making workshop facilitated by the Portfolio lead.

2 The synthesised findings sense-checked with stakeholders to ensure it is sufficiently representative and that relevant findings haven't been overlooked. This should be done through the family-level peer-to-peer learning groups setup to support learning through the family-level evaluations.

3.4.5 Learning and Reflection engagement

The learning stage will coordinate with the portfolio learning, pulling in any additional data to circulate to the family peer-to-peer learning group.

3.5 Thematic Syntheses

The Thematic evaluations will be led by a Thematic Lead (Table 3). The Thematic Lead will engage with all PELs as part of the 'prioritise and select'; 'design'; 'data collection and analysis'; and 'report writing' stages (Figure 2).

The interactions that PET's have with thematic evaluations will mirror the process as set out in 3.4 for family-level evaluations in most cases. As discussed above, there may be situations where specific evaluation field work, separate to that of the programme evaluation, is necessary for specific thematic studies commissioned by the PF. It is expected that these will still be built into the 'select/prioritise' and 'design' stages of the evaluation and learning cycle, so that their inputs are planned for and can be organised in the context of the ongoing evaluation at programme level.

Training and guidance will be provided to the evaluation team from designated Thematic Leads to ensure consistency in data collection, interpretation and analysis.

3.6 Evaluation Ethics

The programme evaluation team are expected to adhere to the highest standards of integrity. The team should uphold the *Research Code of Conduct and Ethics Policy* of the E&L Service, the *DFID Principles on ethical practice in research and evaluation* and the guidance in the *Prosperity Fund Gender and Inclusion Framework*.

These principles and guidance are available in the package of *materials for programme level evaluations*.

4. Evaluation Cycle and Products [Revised depending on the year]

4.1 Timing and Sequencing of Evaluation Activities Years 2-4

In keeping with our user and learning focus, the timing and sequencing of the E&L evaluation and learning cycle is aligned to the PF planning and review cycles – so that evaluations inform annual review processes and plans for subsequent year's activities. The Fund level cycle will be June to June, taking into consideration the June Annual Review process.

Figure 4: Indicative timings and sequencing of annual evaluation and learning cycle

Figure 4 schematically presents the expected timings and sequencing for the annual cycle. Within the constraints of PF fixed reporting dates, and the resourcing requirements of evaluations, the cycle will be flexible and adaptive; itself reviewed annually. Equally there may be specific interim products produced, to support ongoing programme of Fund evidence or learning requirements that will need to be incorporated into the cycle. Year 1 will be slightly different because programmes and projects are expected to come on stream throughout the year and is described in 4.2.

4.1.1 Prioritise and Select (June/July) [revise for specific programme]

This timing coincides with the PF annual review submissions: it should follow reflection on the previous year's evaluation findings, information in the review of the Programme ToCs and informing upcoming programme management decisions.

The PEL will consolidate evaluation and learning themes for evaluation. The coordination with relevant Portfolio Leads will be necessary (3.4 above).

This will require the PEL to validate the EQs and prioritise the learning needs with the PM. This is best done as part of a 'prioritise and select' meeting in which the evaluation stakeholders (including main users, information sources and influencers) are mapped and the approach to be taken with key stakeholders is agreed, as well as the touch points for learning through the evaluation schedule.

Key evaluation team members and the focal points within the programme team should be selected.

4.1.2 Design (August) [revise for specific programme]

The PEL will be responsible for preparing the annual programme evaluation approach paper, using PET members and drawing in capacity from the E&L team (including the respective Portfolio and Thematic Leads, and the Learning and KM Lead). The approach paper should set out the objectives and scope for the evaluation, the methods and data collection techniques to be followed. The approach paper (including evaluation team, workplan and activities) is agreed with the PM. The criteria used for the selection of projects that will be evaluated should be set out in the approach paper.

The Programme evaluation is designed to be strongly user-focused: rapid turn-around of evaluation findings may be necessary to provide key learning evidence or to unblock constraints to progress. Apart from detailing evaluation design, the approach paper should include:

- A plan for how the evaluation products are to be communicated, the production of interim products or touch points on emerging results and findings, to keep the evaluation relevant and responsive.
- Agreed timings and expectations in terms of inputs from the Programme teams and relevant stakeholders.
- Nominated contact points from each team for each aspect of the evaluation.
- Other evaluation inputs (e.g. if there are Thematic evaluations the programme could contribute to) than need to be incorporated into the evaluation plan for the programme.
- Links to other evaluations [likely to be relevant to MDB and possible DFID programmes]

The annual programme approach paper will be subject to QA, managed by the Technical QA Lead within the E&L Service. The PEL should discuss feedback with the PM and the approach paper finalised.

4.1.3 Data collection (September to November) [revise for specific programme]

The evaluation field work will be undertaken in close coordination with the PMs and their respective project/ workstream/ component leads. Depending on their own constraints, they may be involved in aspects of data collection: They can contribute evidence and know-how to build the evidence base to answer the EQs. The evidence will be collated and triangulated.

Interaction with the respective Portfolio and Thematic Leads will be necessary to ensure data collection is consistent with requirements of those synthesis evaluations.

4.1.4 Analysis & Reporting (December to February) [revise for specific programme]

It is likely that it will be possible to start the analysis prior to all field work being completed, but by December, analyses should have started across all programme evaluations. The E&L cycle should support the co-production of findings together with the evaluation users. Touch points are identified that programme evaluations should follow:

1. Validate the collated evidence and analysis and develop recommendations to ensure preliminary evaluation findings are robust and that evaluation recommendations are relevant to the users;

2. Sense-making exercise to put evaluation findings and recommendations back into the current programme context and use them as a critical perspective to review assumptions and causal pathways with consideration for active portfolio management. It should include all key evaluation users (PMs and their project leads, implementing parties or suppliers).

The evaluation write-up should be organised and managed by the PEL with clear roles and delegated responsibilities given to their respective team members. Coordination and communication will be necessary with other E&L team members for purposes of consistency and cross-programme experience sharing:

- With other PELs managed through the Technical Lead (Programme and Portfolio)
- With the Portfolio and Thematic Leads to inform the production of the synthesis evaluations. The Portfolio and Thematic Leads may hold a synthesis workshop that brings together the PELs to jointly assess the synthesised findings for the family and thematic reporting.
- With the Learning and Knowledge Management Lead for identifying cross programme and wider best practice and learning opportunities that can be picked up by the Learning and Knowledge team.

The draft evaluation product will be subject to QA, commissioned by the Team Leader. QA feedback will be discussed with the PM and the programme evaluation paper finalised.

4.1.5 Wider Learning & Reflection (March to June) [revise for specific programme]

Once the evaluation products have been prepared, there is an opportunity for a facilitated process of reflection and incorporation of evaluation findings into future planning and action, including subsequent evaluation priorities. There are several important touch points that the PET can have with the PM and their teams in this stage, to action the evaluation findings and learning outcomes:

1. Plan follow-up actions that takes the sense-making exercise to the next level by learning from the evaluation, drawing out operational implications, then facilitating a process for allocating responsibility for taking action on the learning gains as part of active / adaptive management.

2. Annual ToC assumptions-testing, evidence-building and review of the Programme/Project ToCs ("*how far have we come*", "*where are we now*"). This considers not only the findings of the programme evaluation and the programme monitoring and reporting, but the findings of the synthesis evaluations, and the opportunity for comparative assessment of and learning from sister-programmes or projects with in the PF that this presents. This evidence is collated, summarised and analysed using the baseline reference point, the ToC and the programme or projects' results framework(s).

3. Using the reflection and progress review process, the ToC is revised accordingly and planning priorities established for the coming period by the PM and their team. Information gaps and priorities that can feed the next year's select/prioritise evaluation stage are identified through a facilitate process with the PEL.

The programme evaluation team will be interviewed by the Learning and Knowledge Management Lead (or a member of the learning team) to validate findings and identify

opportunities for wider learning and communications that may be disseminated through the learning and communications team. This leads onto the fourth touch point in this stage:

4. To ensure that E&L plans for sharing of externally valid lessons from the evaluation or for commissioning further evaluations to address gaps in knowledge are appropriate to the programme's own learning needs and opportunities.

4.2 Year 1 Programme Level Evaluation Activities (revised each year)

In the first year of E&L implementation (starting April 2018) the programme evaluation team will undertake the evaluation baseline and the first formative learning evaluation.

4.2.1 Initial Engagement

The initial interactions are critically important in establishing a constructive working relationship between the PM and their team and the E&L's team that will work with them. The initial 'hand-over' from the Inception Phase Engagement Team to the designated PET will be completed at the beginning of the implementation phase. The PEL and [designated team members] will introduce the E&L approach, principles and overall plan and role for the evaluation in year one. Timings for and the process for the first engagement with the evaluation team will be agreed.

The PEL and [designated team members] will:

- Discuss with the PM, the programme's early learning and information needs to support the prioritisation and design of the evaluation approach for year 1.
- Cover three early evaluative assessments that will be part of the 'prioritise' and 'design' stages of year 1: *i.e.* the stakeholder mapping, context mapping and secondary data assessments (detailed below). These assessments are necessary to inform the selection of evaluation methods and approach for year 1.

4.2.2 Baseline Evaluation

The early baseline activities will require light levels of interaction with the Programme teams, and draw on a range of programme specific documentation and publicly available data sets and information. The findings will be reviewed collaboratively as part of the 'stock-take' learning point and inform the design and approach paper. The follow-on baseline work (described below in 4.2.2.5-6) is likely to require engagement with the project implementers in many cases, and may have to wait until those are in place to be completed.

Early Baseline

Stakeholder Mapping

The stakeholder mapping is necessary to establish:

- The wide range of users of the evaluation evidence and learning for each programme
- Who the key influencers are with respect to the programme/project ToC
- Who represent potential sources of information or evidence that will inform the evaluation
- How stakeholders may be engaged as part of the evaluation process.

The stakeholder mapping and context mapping exercises may be done concurrently, because of the importance of understanding the political economy context for stakeholder analysis.

The method for stakeholder mapping should follow the approach used for the Fund level stakeholder mapping exercise.

Context Mapping

The success or otherwise of the programme will be highly dependent on its ability to understand and navigate the contexts in which in works. Most of the programmes seek to effect regulatory and institutional changes in target countries aimed at improving the environment for pro-poor sustainable economic development: this is subject to many influences that are outside of the control of any particular intervention and are similarly uncontrollable and unpredictable. It is clear therefore that no matter how robust the rest of the Theory of Change, unless each programme is able to manage the contextual factors in which it finds itself, it will struggle to achieve anticipated outcomes and contribute to the achievement of Fund level aims.

Therefore a thoughtful assessment of context will be needed that covers the following dimensions:

- Programme and project-level context
- Country political economy context (for country programmes only)

A consistent method for context mapping will be used across the programmes, drawing on the approach used for the Fund level context analysis. It is available in the *materials for programme level evaluations*.

Secondary data review

The type, quality and availability of secondary data sources for answering each of the EQs needs to be assessed to inform the selection and prioritisation of the EQs and the design of the evaluation methodology. The survey should include free and paid-for secondary data sources.

A consistent methodology for mapping and rating data will be used across the Fund and should draw on the approach used for the Fund level review of secondary data.

Ongoing coordination with the MR service will be required in the review of secondary data so that opportunities for improving potential indicators and data sources can be communicated and further developed iteratively as programmes come on stream.

The initial overview of secondary data sources produced during the inception phase is available in the *materials for programme level evaluations*.

Primary data requirements

Primary data may be required to triangulate claims, fill missing information gaps or strengthen weak evidence. Where there is scope to ‘piggy-back’ on existing surveys through payment for example, this may prove a cost-effective option and should be explored.

Follow-on Baseline

ToC Evaluability

The programme evaluation team are expected to review existing; and gather baseline information on; the evidence base for programme or project Theory of Change, specifically the assumptions and casual pathways. This information can be combined with the early

baseline data collection on context, stakeholders and data to effect an evaluability assessment of the programme ToC (or its constituent ToCs).

Clear, specific outcomes and plausible contribution claims for the programme/projects need to be established against which progress may be reported annually and tested by the evaluation. It is likely this will require consultation with programme teams together with the implementers/suppliers.

A consistent approach to this across the programmes is recommended. Guidance on the attributes that make a programme ToC evaluable are available in the methods toolbox.

Assess Programme Results Baselines

Once the ToC has been reviewed and specific outcomes have been defined by the programme teams, the programme results need to be baselined for evaluation purposes. This allows the later evaluation of the project's results and the change claimed by the programme to be measured against that baseline, as well as to evaluate the relative strength and direction of that change. Ongoing coordination with the MR service will be required in the establishment of results baselines to ensure they set a baseline appropriate for evaluation.

Baseline of Programme Families

The ToC analysis and the results framework for each programme provides a 'menu' of change areas that will allow a review of the "family" groupings initially developed in the inception based on business case data.

Once the groupings have been confirmed in discussion with the PMs, joint learning groups for peer-to-peer learning across the families of programmes or projects will be set up by the Portfolio Lead together with the Learning and Knowledge Management Lead. This will create a structure for future family-level evaluations and a structure for synthesising evaluation evidence under Intermediate-Outcome causal pathways.

4.2.3 First formative learning evaluation

The baseline work should provide immediate information to the PMs on the strength of the programme's intervention designs. The first formative learning evaluation should help the PM and their team build an understanding of how to use the evaluation strategically in their management cycle and support readiness for implementation. It will contain an early evaluation of some core programme EQs and EQs set by the programme teams that reflect their priority learning needs at this early implementation stage.

Identify the early learning needs of Programme Managers

The programme evaluation team should work with the PM and their teams to identify their early learning needs to:

- Inform the definition of the EQs for the programme beyond the core questions;
- Establish the scope for cross-programme learning and evaluations (using the 'family' groupings) where similar learning needs are identified;
- Set out a process for evaluation information to contribute to programme management cycles; and
- Design the first formative learning evaluation approach.

Assessment of capacity for delivery at scale

The evaluation will provide evidence of the programme readiness for implementation and inform a wider cross-programme assessment. It will cover *inter alia:*

- How the programme and planned programme and project governance structures will support programme accountability;
- The programme's capacity for adaptive or active portfolio management;
- How gender, equity, VfM, environment and sustainability have been incorporated into project design and how these factors will be monitored and reported against.

4.2.4 Year 1 Scheduling

Across the Fund, the baseline is likely to be staggered because not all programmes will be ready to receive the evaluation team from the beginning of April. The initial scheduling will be agreed between the Programme Evaluation Leads and PM. However, it's likely to follow a timetable similar to that set out in Table 5.

Table 5: Outline scheduling for the Year 1 baseline and 1st formative learning evaluation

Timeline	Main evaluation activities
April	Initial hand-over from the Inception Engagement Team to the designated PET team members. Introductions and timings for Y1 evaluation and the first engagement with the evaluation team agreed.
Engagement Starts	
Weeks 1-2	Programme Evaluation Lead and the designated team members discuss with the PM the programme's early learning and information needs and early evaluative assessments for year 1. Context mapping, stakeholder mapping and secondary data assessments initiated.
Weeks 3-6	Context mapping, stakeholder mapping and secondary data assessments completed Programme EQs prioritised, agreed and evaluation methods designed
Weeks 7-8	Approach paper completed and consulted
Weeks 11-15	Data collection and analysis for 1st formative assessment and baseline
Weeks 15-17	Data validation and sense making exercises - Interactive ToC evaluability assessment - Programme results baseline assessment - Inter-programme family groupings for learning and synthesis agreed
Weeks 17-20	Completion of Year 1 programme evaluation report and lesson learning exercise

4.3 Annual Programme Evaluation Products

The Programme Evaluation teams will be responsible for producing several evaluation products and contributing to others. Each programme evaluation team will be responsible for producing:

- Annual Programme Approach Paper
- Annual Programme Evaluation Report

Beyond the main evaluation products (*i.e*. approach papers and annual evaluation products) the programme evaluations would be expected to deliver interim products and learning outputs that support the learning and information needs of the PMs and the synthesis evaluations. These interim products would be detailed in the annual programme approach paper.

Each programme evaluation team will be responsible for contributing to:

- Annual Family-level Evaluation Synthesis Report as defined in advance by the Portfolio Leads.
- Annual Thematic Evaluation Report as defined in advance by the Thematic Leads.
- Learning materials that draw on strong learning and good practice findings from their evaluation.

5. Learning

The evaluation and learning cycle described in Section 1.7 above is designed to maximise opportunities for the evaluation to facilitate a learning process which responds to PMs' learning needs. The process, with embedded learning touch points, should enable the PM and their team to learn with the evaluation.

There is however great opportunity for wider learning to happen alongside the evaluation. This wider learning will be supported by the PF through:

- Knowledge Products
- Synthesis of lessons from the breadth of the evaluation and research
- Peer to Peer learning (e.g. through the families of project/programmes that create opportunity for wider communities of practice)
- Annual Reflection

The E&L Learning Plan details the learning and knowledge management initiatives of the E&L service and is available in the *materials for programme level evaluations*.

The E&L learning component is led by the Learning and Utilisation Lead who will provide a briefing to all Programme Evaluation Leads on the learning component and the role that the programme evaluation team is expected to play in supporting this.

6. Resourcing requirements

6.1 Expected Level of Effort

The E&L Service has a finite resource for conducting its evaluations at Fund, programme and family levels. The E&L service is not able to evaluate every project or sub-component of every programme. Projects will be sampled based on the information and accountability needs of the PM and wider family or thematic requirements and will be clearly explained in the approach paper. Opportunities for joint evaluations between projects – on common learning themes for example – could reduce costs whilst enhancing learning across programmes. Such opportunities will be explored during the design stages of the programme evaluations.

- The programme baseline and 1st formative evaluations are expected to require at least 250 person days each.
- The annual programme evaluations are estimated to be 250 person days.
- The resourcing requirements for the annual evaluation of MDB projects is expected to be lower at 100 person days per year.
- The Family-level syntheses will involve an additional 180 person days per family, incorporating both the additional data collection and analysis work that may be necessary.

6.2 Evaluation Team

The specific requirements of each evaluation team will depend on the scope of each programme and the evaluation focus in any particular year. In terms of Year 1 baseline and 1st formative learning evaluation, the following team members are expected:

- Programme Evaluation Lead
- National Expert(s) (stakeholder analysis, context mapping, secondary data analysis)
- 3x Evaluators incorporating expertise in: Secondary data analysis, political economy analysis, evaluation methods, gender, VfM, economics
- Specific sectoral expertise with short term analytical and technical inputs.

Each programme evaluation team will be supported by the wider E&L team comprising Thematic Leads, the Portfolio Lead, Learning and Knowledge Management Lead, Technical Lead, and the respective Programme Evaluation Leads; each of whom will have time dedicated to supporting the evaluation. The available level of effort will be made clear to the Programme Evaluation Lead who will be expected to manage the inputs available to the team to deliver the evaluation.

6.3 Logistics

The PET will be supported logistically in terms of travel and accommodation requirements by the E&L Service's central operations function. This function will also conduct travel risk assessments on behalf of the E&L service prior to PET country visits.

Programme Teams (through the Programme manager) will be asked to assist in making introductions to stakeholders (e.g. Key Informants), assisting in making information available to the evaluation teams, providing accompanying to team members when necessary.

7. Limitations, Conflict and Risk Mitigation

7.1 Limitations

7.1.1 Scale and Scope

The programme evaluations are limited in the scale and scope of their potential coverage. With the resources available only some projects can be sampled and only the priority learning and information needs of a PM and their team can be tackled in any one year.

The programmes are large and many are diverse either geographically or in the themes that they cover. In many cases they represent a microcosm of the Fund itself – *i.e*. they are not

unitary programmes but portfolios of projects within a programme. Therefore it is important that expectations on evaluation are matched to evaluability of the programmes within the resource constraints of the E&L service. This should be clearly set out in the annual approach paper.

[Other Factors identified by respective Programme]

7.1.2 Design Limitations

The most appropriate design for the programme evaluations will be contingent on the evaluation questions, intended uses and attributes of the intervention, and evaluation process. There may be no single 'best' method but rather an optimum combination of methods that the evaluation can use. With this in mind, it is important that the evaluation clearly sets out the strengths and weaknesses of the methods selected for the programme evaluation in the annual approach paper.

[Other Factors identified by respective Programme]

7.2 Risk Mitigation

The annual approach paper will outline possible risks, dependencies, and contingencies that could affect the process or outcomes of the evaluation. A risk, for example, could (hypothetically) include fundamental differences in stakeholder perspectives on what change has been achieved; activities taking substantially longer than envisioned, affecting the timing of final evaluation results and products; and the chance that findings could have unanticipated implications on the programme's design or approach.

7.3 Dispute resolution (TBC)

This covers internal disputes within the Programme evaluation team, disputes between the programme evaluation team and the MR Service, and dispute between the programme evaluation team and the PM.

7.3.1 Internal

Dispute Management

The E&L Service has internal management structures that should deal with potential conflict or dispute resolution. If there is a conflict with the programme evaluation team, the issue should be arbitrated by the E&L Technical Lead (Programme and Portfolio). If this cannot be resolved by the Technical Lead or the parties do not accept the decision of the Technical Lead, this may be escalated to the E&L Service Team Leader. The Team Leader will hold a hearing on the matter where the issue is raised and discussed. The Team Leader will have final say on the matter.

Mitigation

The risk of internal dispute will be mitigated through:

- The establishment of clear roles and responsibilities for each evaluation team member, the expectations for them and how they should engage with other team members and with evaluation stakeholders.
- The roles of the Portfolio Lead and Thematic Leads and their relationship to the programme evaluation will be defined in the approach paper for the evaluation.

- The Programme Evaluation Lead will hold regular weekly stock take meetings with their evaluation team so that issues may be aired and discussed and risks managed before they become conflictual.

7.3.2 External

Between the evaluation team and the MR Service

An MoU governs the management of relations and the management of disputes between the E&L and MR Service. Regular communication and coordination meetings between the parties are forums for communication, updates and raising any issues of concern affecting each Party's respective service contracts including the areas of Joint Activity. The MoU sets out procedures for escalation of issues and settlement of disputes.

Between the evaluation team and the Programme Manager or their team members

Management actions

To manage potential conflict situations during a programme evaluation:

- If either party has any disagreement or issues that may give rise to a conflict, the first step is to seek to resolve the issue by a process of consultation at the level of PM and Programme Evaluation Lead. The Portfolio Lead with oversight of the evaluation should be informed to understand and see if they can facilitate the consultation.
- If the issue cannot be resolved within a reasonable period of time, the matter shall be escalated to the programme SRO, and a meeting held between the SRO, the relevant E&L Lead and the PM, to consider the case and shall decide on the appropriate course of action to take. The decision of the SRO in consultation with the PFMO shall be final. The issue, process and decision made will be reported back through the contract review meeting mechanism that governs the contract overall.

Mitigation actions

To mitigate risks of conflict:

- Good and open communication between the PM and the Programme Evaluation Lead, through regular evaluation update meetings (either phone or video conference, face-to-face where necessary), that provide a forum for updates, monitoring risks and raising any issues of concern affecting the evaluation process.
- The annual approach paper should clearly set out the roles and expectations of the different parties in the evaluation, including roles and expectation from each party in terms of involvement in the completion of the evaluation's expected outputs.
- Establish and jointly agree with the PM the 'ground rules' under which the evaluation team will need to conduct the evaluation, in terms of: being able to work freely and without interference, there is transparency on conflicts of interest and independence of data sources and that differences of opinion will be fully acknowledged in the evaluation.

Finally, the contract between the E&L Service lead (WYG) and Cabinet Office provides a forum for dispute resolution of a serious nature insofar as such disputes could impact on the E&L head contract. At the central level, the E&L Service holds monthly contract review meetings

that includes representation from Cabinet Office as well as PFMO. Updates, risks and issues of concern are raised and discussed at this meeting.