HSNP
Hunger Safety Net Programme

Hunger Safety Net Programme

Evaluation of HSNP Phase 2: Inception Report

11 July 2015

National Drought Management Authority
NDMA
Committed to ending drought emergencies

UKaid
from the British people

Australian AID

Oxford Policy Management

EVALUATION OF HSNP PHASE 2

Executive summary

PROGRAMME BACKGROUND

This report sets out the Oxford Policy Management (OPM) proposal for evaluating Phase 2 of the Kenyan Hunger Safety Net Programme (HSNP). The HSNP is one of four cash transfer (CT) programmes in Kenya, which together form the National Safety Net Programme (NSNP). The NSNP is intended to harmonise these programmes within an integrated system of national social protection. While the other three CT programmes focus on children, older people and the disabled, the HSNP targets the extreme poor. It provides regular, unconditional CTs in four counties in northern Kenya: Marsabit, Mandera, Turkana and Wajir. These predictable transfers are supplemented by emergency payments to the rest of the population triggered by drought.

HSNP Phase 1 took place from 2009 to 2013. It supplied monthly CTs of KES 1,750 to around 69,000 households. Phase 2 began in July 2013 and will end in March 2017. Phase 2 transfers will increase to KES 2,450 (approx. $28) per month, payable every two months to up to 100,000 households. These households are selected through a combination of a proxy means test (PMT) and a community-based wealth ranking (CBWR).

Administration

The HSNP is operated by the National Drought Management Authority (NDMA) of the Government of Kenya (GoK), with support from a Programme Implementation and Learning Unit (PILU). This support covers assistance with programme delivery, a tool for rights and grievances, independent evaluation, and capacity building for programme management. The external implementing partners are HelpAge International (HAI), which is responsible for programme rights; Financial Sector Deepening Trust (FSD), which manages payments; and Equity Bank, which implements payments.

Funding

The HSNP is funded by the Government of Kenya, the UK Department for International Development (DFID) and the Australia Department for Foreign Affairs and Trade (DFAT). In addition, the World Bank's Programme for Results (P4R) provides funding to the NSNP, to which the HSNP belongs.

EVALUATION OVERVIEW

The independent evaluation will demonstrate the performance and impact of the HSNP for the benefit of implementers, funders, other stakeholders, and those interested in CTs more generally. The evaluation findings will feed into on-going programme operations and future programmatic and funding decisions.

Technical team

OPM's core evaluation team for this assignment combines international expertise and in-country knowledge with the selection of 12 experts, of whom five worked on the Phase 1 evaluation. The team is led by Roger Pearson, who has high-level experience with complex evaluations and social protection programming. This team leader (TL) is supported by a Kenyan national project coordinator (KNPC), who has strong experience of both M&E (monitoring and evaluation) and northern Kenya. Support to the core team is provided by a pool of

senior experts. The field research will be carried out by Research Guide Africa (RGA), with which OPM worked on the HSNP Phase 1 evaluation. Robust quality assurance (QA), both internal and external, is ensured.

Evaluation risks

Security is the greatest risk the evaluation faces. This is a particular concern in the two HSNP-covered counties of Mandera and Wajir, as the OPM team may need to travel into high-risk areas. To address this concern, OPM has drawn up a security plan and operating protocols in conjunction with the security adviser Spearfish. If the security situation demands, and only after discussion with DFID and other stakeholders as needed, OPM may need to temporarily suspend some activities of the evaluation.

OPM'S PROPOSED APPROACH

OPM conducted a rigorous evaluation of HSNP Phase 1, finding the programme to be an effective safety net. Our approach to the Phase 2 evaluation differs from that used for Phase 1, reflecting the changing priorities, contexts and evidence needed. The proposed evaluation will assess the impact of the HSNP on the larger community. It also seeks to build on the earlier work, enabling M&E functions to be embedded within the HSNP going forward. This will contribute to programme sustainability and capacity building within the programme team.

For Phase 2 OPM proposes the following core elements:

- impact evaluation (IE, including a local economy-wide impact evaluation [LEWIE]);
- operational evaluation;
- continuous independent programme monitoring; and
- policy analysis.

The more novel component supplementing these is a communications and learning workstream. This is for the purpose of disseminating the evaluation findings and building knowledge within the NDMA and PILU about how to carry forward the evaluation work. The communications and learning workstream will be delivered through tailored learning events and a comprehensive communication strategy.

The evaluation deliverables extend from the communications and learning strategy document in August 2015 to the anonymised evaluation datasets/metadata in September 2017.

Evaluation structure

OPM's approach answers all of the evaluation questions set out in the original terms of reference (ToR), including adjustments deemed necessary in subsequent discussions with DFID. The structure is made up of four workstreams. These address the Development Assistance Committee (DAC) criteria of impact, effectiveness, efficiency and sustainability as relevant. The overall structure is as follows:

- Workstream 1: IE
  - LEWIE model
  - quantitative impact analysis
  - qualitative analysis

  - special study on the Arid Lands Support Programme (ASP)
- Workstream 2: Operational M&E
  - process and institutional capacity assessment
  - process review of the two pilot emergency payments
  - costing study
  - bi-monthly operational monitoring and endline report
- Workstream 3: Policy analysis
  - targeting study
  - simulation analysis
  - registration instrument design
  - strategic review (including, at DFID's request, a review and redesign of the tool for registration)
- Workstream 4: Communications and learning

Each workstream is described in more detail below.

Workstream 1: IE

OPM's comprehensive approach to the Phase 2 IE calls for the use of mixed methods. The quantitative data will derive from a household survey, while the qualitative data will be collected in multiple rounds using participatory methods. There are four components:

- The LEWIE will draw on a survey of 6,468 households, selected through a multi-stage cluster sample design. It will simulate how the HSNP is affecting the wider local economy beyond the households receiving the CTs. This will draw out any multiplier effects of the programme, whether positive (e.g. increased income) or negative (e.g. higher prices). The LEWIE will address ToR question f (on multiplier effects). This research component will run from July 2015 to September 2016, at which point a report will be issued.

- The quantitative IE will also be founded around the household survey data. It will be built on a regression discontinuity (RD) design, used to estimate a programme's effect by comparing points above and below the programme eligibility cut-off. The quantitative IE will examine certain sub-samples, such as households that are smaller/larger and poorer/less poor. Impact will be measured across four ToR evaluation questions (a, b, d and e – covering welfare effects, sub-group differences, the influence of one-off transfers, and livelihood opportunities). The timeframe will be similar to that of the LEWIE.

- Qualitative research will analyse the context of the HSNP, the influences on the programme's outcomes, and non-numerical impacts such as social relations. As well as providing insight into the causal pathways of programme impact, it will help to triangulate the findings from the quantitative study. Using focus group discussions (FGDs), interviews, panel studies and observation, this research component will respond to ToR questions c, d, e, g, k and s. OPM has classified these into four categories: perceptions of wellbeing at individual, household and community levels; vulnerability and resilience; livelihoods and local markets; and informal institutions and social relations. The outputs of

the qualitative research will be three reports, one for each of three distinct rounds of research. The reports will be ready in October 2015, November 2016 and August 2017. The qualitative and quantitative methods will work in combination across three dimensions: the analytical concept of wellbeing, household-level analysis and community-level analysis.

- Finally, the special analysis of the ASP will determine how this programme interacts with the HSNP and local governance. Using both quantitative and qualitative data, this strand will respond to ToR questions e (the influence of one-off transfers) and m (the role of NGOs). This work will lead to a standalone report in March 2016.

Workstream 2: Operational M&E

Workstream 2 will evaluate the running of the HSNP, determining best practices, recommendations, efficiency and value for money. It is envisioned as comprising two components: standalone studies on particular operational matters and regular monitoring reports. Specifics are given below.

- The process and institutional capacity assessment will, through a combination of key informant interviews (KIIs) and quantitative data analysis, assess the effectiveness of resourcing to run the HSNP. This applies to ToR questions l (effectiveness of management) and q (future implementation capacity). The analysis report will be presented at a learning event.
- At DFID's request, an emergency payments process review will form part of Workstream 2. This will use KIIs to conduct a process tracing exercise laying out activities, decisions and outcomes related to four aspects of emergency payments: communication and sensitisation, targeting and verification, payments, and grievances. The report will be finalised during a tailored workshop.
- The costing study will analyse HSNP expenditure, using KIIs and quantitative analysis. The ToR questions incorporated into the costing study are n (cost-effectiveness of triggered vs. regular payments) and o (cost of Phase 2). These questions have been modified in discussion with DFID; for example, it has been agreed that the costing study will not cover the ASP. The study's report, together with a breakdown down of programme costs, will be presented at a learning event.
- Ongoing monitoring will be provided by a series of bi-monthly monitoring reports, from September 2015 to July 2017, and culminating in an endline synthesis report. Quantitative data and interviews will be used in tandem. This regular reporting will provide timely, actionable information on ToR questions i (targeting based on registration data), j (rights and grievances process) and k (benefits of financial services).

Workstream 3: Policy analysis

The policy analysis workstream will integrate sustainability and relevance into considerations of impact. The following three channels are envisaged:

- A targeting study will evaluate how well the registration process and selection criteria have targeted the poorest households as participants in the HSNP, corresponding to ToR question i. The data will come from the HSNP's management information system (MIS), supplemented by survey data. This

work is intended to be conducted in March 2016 and presented at a learning event about two months later.

- Micro-simulations will be conducted to simulate the costs and impacts that may stem from policy changes in the future (ToR question r). In the case of HSNP Phase 2, the micro-simulation analysis will inform the design of the core programme as well as the scope to scale up the programme. To this end, OPM will use IE survey data beginning in October 2016 to present two reports, one on the core programme and the other on the emergency payments component, at a learning event. This reporting will include various design possibilities, such as different transfer values, eligibility criteria and payment frequencies.
- OPM will design the registration instrument for Phase 3. This will ensure the next registration instrument is aligned with the evolving needs of the NSNP, optimised for future M&E purposes, and future eligibility for the programme is robustly determined.
- Lastly, the strategic review will take a broader view. Using a variety of research methods, it will address the fit among the HSNP's objectives, its implementation, and the national social protection strategy, along with the effects of policy dialogue (ToR question h), a combined approach (t), and relative cost-efficiency (u). The report will be made ready in April 2016 in order to contribute to the business case for Phase 3. This output will be supplemented by a presentation at a dissemination event involving key stakeholders, as well as a two-page brief.

Workstream 4: Communications and learning

The communications and learning workstream will work toward strong information sharing and dialogue about the HSNP with all stakeholders. This will have three components: programme learning, wider communications, and support to the HSNP monitoring framework. Following stakeholder analysis, OPM will work together with the PILU to deliver a strategy document on communications and learning that will be operational from June 2015 to September 2017.

Contents

List of tables, figures and boxes

List of abbreviations

ASAL	Arid and Semi-Arid Lands
ASP	Arid Lands Support Programme
CAPI	Computer-Assisted Personal Interviewing
CBA	Cost–Benefit Analysis
CBT	Community-Based Targeting
CBWR	Community-Based Wealth Ranking
CDC	County Drought Coordinator
CRA	Commission of Revenue Allocation
CT	Cash Transfer
CT-OVC	Orphans and Vulnerable Children Cash Transfer
CT-PWSD	Cash Transfer Programme for People with Severe Disability
DAC	Development Assistance Committee
DALY	Disability-Adjusted Life Year
DFAT	Australia Department for Foreign Affairs and Trade
DDFID	UK Department for International Development
DoC	Duty of Care
DR	Dependency Ratio
ERC	Ethical Review Committee
FAO	Food and Agriculture Organization
FCS	Food Consumption Score
FGD	Focus Group Discussion
FSD	Financial Sector Deepening Trust
GoK	Government of Kenya
HAI	HelpAge International
HSNP	Hunger Safety Net Programme
IE	Impact Evaluation
KES	Kenyan Shilling
KHIBS	Kenya Integrated Household Budget Survey

KII	Key Informant Interview
KNPC	Kenyan National Project Coordinator
LEAP	Livelihood Empowerment Against Poverty (Ghana)
LEWIE	Local Economy-Wide Impact Evaluation
M&E	Monitoring and Evaluation
MDP	Ministry of Devolution and Planning
MIS	Management Information System
MLSSS	Ministry of Labour, Social Security and Services
MMOU	Memorandum of Understanding
NDMA	National Drought Management Agency
NSNP	National Safety Net Programme
OPCT	Old Age Persons Cash Transfer
OPM	Oxford Policy Management
P4R	World Bank Programme for Results
PILU	Programme Implementation and Learning Unit
PMT	Proxy Means Test
PPS	Probability Proportional to Size
PSNP	Productive Safety Net Programme (Ethiopia)
PSU	Primary Sampling Unit
PSM	Propensity Score Matching
QA	Quality Assurance
RCT	Randomised Controlled Trial
RGA	Research Guide Africa
SAGE	Social Assistance Grants for Empowerment (Uganda)
SAM	Social Accounting Matrix
SPS	Social Protection Secretariat
SSC	Spearfish Social Consultant
TL	Team Leader
ToR	Terms of Reference
VCI	Vegetation Condition Index

WFP	World Food Programme

1. Introduction and background

1.1 THE KENYA HUNGER SAFETY NET PROGRAMME – PHASE 2

The GoK, with support from the UK's DFID Kenya and Australia's DFAT, started the second phase of the HSNP in July 2013. The HSNP is part of the GoK NSNP, which brings together four CT programmes in Kenya under one umbrella (see Box 1).

Box 1 Kenya NSNP

The four largest CT programmes in Kenya are implemented by two ministries: the Ministry of Labour, Social Security and Services (MLSSS) and the Ministry of Devolution and Planning (MDP). The three programmes housed in the MLSSS are separated among the Cash Transfer for Orphans and Vulnerable Children Programme (CT-OVC) in the Department of Children's Services; the Older Person Cash Transfer Programme (OPCT); and the Cash Transfer Programme for People with Severe Disability (CT-PWSD) in the Department of Social Development. The HSNP is being implemented under the NDMA, which reports to the MDP. An internationally procured PILU sits within the NDMA.

To increase the efficiency and effectiveness of the CT programmes, the government has established the NSNP. The aim is to create a framework around which the four main cash transfer programmes (CT-OVC, OPCT, CT-PWSD and HSNP) will be increasingly coordinated and harmonised. The NSNP has three objectives that aim to improve the efficiency and effectiveness of safety net support to poor and vulnerable populations in Kenya. The NSNP will: (i) create more robust and transparent systems for targeting, beneficiary registration, payments and monitoring, and strengthen the overall governance of these programmes; (ii) increasingly harmonise the four CT programmes to improve the coherence of the sector; and (iii) will expand the coverage of the four programmes in a coordinated manner that will progressively realise the right to safety net support.

Establishing the NSNP is the first critical step in a longer-term reform agenda that aims to establish a national safety net system as part of an integrated national social protection system. The Social Protection Secretariat (SPS), a body created by the National Social Protection Policy, provides sector-wide oversight and coordination.

The HSNP is an unconditional CT programme that focuses on people living in extreme poverty in four counties in northern Kenya: Marsabit, Mandera, Turkana and Wajir. The first phase, which ran from 2009 to 2013, provided around 69,000 households (approx. 496,800 people) with predictable electronic CTs worth KES 1,750 per month.[1] Under Phase 2, the transfer is worth KES 2,450 per month (approximately £18/$28). The transfer is made directly into recipients' bank accounts every two months.

HSNP Phase 2 will run till March 2017. It will continue to focus on the same four counties as HSNP Phase 1. DFID plans to provide £85.59 million and the GoK will also contribute funding as part of the NSNP – by 2017 it is envisaged that 49% of total programme costs and 54% of the caseload will be met by the GoK.

Under HSNP Phase 2, up to 372,806 households in the four counties have been registered for bank accounts and the HSNP will together provide regular CTs to up to 100,000 of these 'Group 1' households. The HSNP is targeted using a combination of a PMT and a CBWR. The allocation of beneficiary quotas to the four counties was made on a the basis of

[1] The HSNP originally provided KES 2,150 to each beneficiary household (or individual in the case of the Social Pension component) every two months. This was calculated as 75% of the value of the World Food Programme (WFP) food aid ration in 2006, when the value of the transfer was originally set. Over time, the value of the transfer has increased and at the end of the evaluation period stood at KES 3,500. The value of the HSNP transfer was initially increased from KES 2,150 to KES 3,000 with effect from payment cycle 16 (Sep/Oct 2011). It was subsequently increased to KES 3,500 with effect from cycle 19 (Mar/Apr 2012). A one-off doubling of the transfer occurred in Jul/Aug 2011 to support households coping with drought.

the PMT/CBWR distribution of the poorest households, adjusted by the NDMA using a modified version of the Commission of Revenue Allocation (CRA) formula, giving a greater weighting to poverty in allocating resources to the devolved county governments.[2]

As well as the predictable cash transfers, the HSNP also has provision for additional CTs to be triggered in the event of a drought shock (in this document referred to as 'emergency payments'). These are targeted at 'Group 2' households, which are not in receipt of the routine bi-monthly transfers.

The NDMA is responsible for leading the HSNP (see Box 1). It does this with the support of a PILU. The PILU helps to manage and monitor the delivery of the HSNP, provide oversight of a rights and grievances mechanism for the programme, and oversee the independent evaluation. It is also helping to build the capacity of the NDMA to manage the HSNP. The PILU reports to the NDMA and HSNP Steering Committee and comprises a mixture of internationally procured technical expertise (including both national and international staff), which will be imparted to the NDMA over the lifetime of the PILU.

The HSNP is delivered in partnership with implementing partners HAI, which manages the programme rights component, and FSD and Equity Bank, which manage and deliver the payments component respectively.

As part of the NSNP, the HSNP is a beneficiary of the World Bank's Programme for Results (P4R), which is supporting development of the NSNP. Some of the indicators that trigger payments to the GoK under the P4R rely on data from the HSNP programme and its evaluation.

The independent evaluation of the HSNP is required to provide evidence on programme performance and impact for use by all programme stakeholders, including the PILU, NDMA, DFID, NSNP and GoK, and other national and international stakeholders. The evaluation will inform future decision making and accountability for funding, as well as the wider community interested in CTs, both nationally and internationally.

1.2 OBJECTIVES OF THE EVALUATION

The previous evaluation of HSNP Phase 1, led by OPM, provided robust evidence that the HSNP works effectively as a safety net, particularly for the very poorest, directly supporting families to be more food secure, hold onto their assets during shocks, and spend more on health. The evaluation of the HSNP pilot phase was very rigorous, but it was also resource-intensive and placed large demands on the implementation of the programme in order to facilitate the community-randomised, staggered roll-out that underpinned our randomised controlled trial (RCT) design. This approach was appropriate for Phase 1, where the priority was establishing the impact of a highly innovative programme operating in an extremely complex context. After several years of implementation of HSNP, the policy and programme context have changed considerably, and so has the evidence needed to further inform the HSNP design and operation. Since there is already 'proof of concept', we are tailoring Phase 2 of the evaluation to respond to these needs, rather than to repeat the exercise of Phase 1.

In our proposal we set out how we will deliver on the ToR, implementing an **IE** (incorporating **a LEWIE**), an **operational evaluation**, **continuous independent programme monitoring**, and **policy analysis**. As agreed with DFID, the research to be conducted under these workstreams will not necessarily all be synthesised into a single assessment of the HSNP (though some of the results under different workstreams will speak to each other), but rather will achieve the broad objectives set out in the ToR via distinct sets of activities.

[2] NDMA opted to incorporate a modified CRA formula to allocate the final number or poverty ranked 100,000 households between the four counties. The CRA formula was modified by removing land area and fiscal responsibility and increasing the weight of the poverty count to 30%, resulting in the following weighting: 25% basic equal share, 30% poverty and 45% population. The poverty line calculating the poverty headcount rate component of the CRA formula was taken as the HSNP cut-off, i.e. household per adult equivalent monthly consumption of KES 442.6.

But we also want to go beyond simply delivering these standard evaluation objectives. Having undertaken the previous evaluation, we want to focus the Phase 2 evaluation in a way that builds on and complements, rather than replicates, our previous work. While periodic independent evaluation will always be required, as the HSNP matures it should be expected to incorporate much of the learning and accountability functions (impact and operational evaluation, ongoing monitoring and policy analysis) into its own routine M&E and policy development activities.

Our aim is therefore to design the core evaluation and monitoring instruments and technical approaches in such a way that they can be used as an appropriate basis for extending the scope of the programme's own routine M&E and policy development processes going forward. To support this incorporation of our instruments and approaches into the programme's own M&E and policy development processes, we will implement a **'communications and learning' workstream**. Through ongoing engagement with the NDMA and the PILU, specifically **tailored learning events** and a comprehensive **communication strategy**, our intention is to facilitate within the PILU and NDMA a firm understanding and ownership of the evaluation approach, such that core elements can be sustainably adopted by the programme itself going forward. For example, the evaluation will devise a methodology for routine monitoring of payments that could be taken up (and/or adapted) by the HSNP. Similarly, the evaluation will set up a structure for analysing programme costs that can be reported on in a routine way moving forward.

1.3 EVALUATION QUESTIONS

The original evaluation ToR put forward a series of key research questions. Our proposal addressed these questions by relating them to the DAC evaluation criteria and mapping them against the proposed workstreams and specific activities to be undertaken.

Table 2 presents a summary of the original evaluation questions, and whether there have been any changes or clarifications to them as a result of the discussions held during the inception phase.[3] Table 1 provides a summary matrix that links the evaluation questions to each workstream and component and points the reader to the relevant section of the document.

How these questions will be answered is set out in detail in the methodology sections for Workstreams 1–3. At the start of each of the relevant sub-sections, the specific evaluation questions to be addressed by that particular evaluation component are summarised in an evaluation matrix, together with the corresponding sub-questions, main indicators, data sources, data collection methods, data analysis methods, DAC criteria and evaluability issues. Table 1 summarises the questions covered by each matrix and its location in this document.

[3] During our inception mission, which took place between 19 January and 4 February 2015, the evaluation team met with the PILU and DFID, as well as other key HSNP stakeholders. The purpose was to discuss the proposal in detail, consider any changes in the needs and priorities of the programme and its stakeholders, and gain a more up-to-date and detailed understanding of the situation of the programme in order to work out what was technically, operationally and politically feasible in terms of each of the evaluation's various activities and outputs. Since that mission, further discussions have taken place with all key stakeholders, including during a second inception mission that took place between 3 and 8 June 2015. These discussions informed the evaluation design elaborated below, including the revised evaluation work plan.

TABLE 1 EVALUATION MATRICES IN THIS REPORT

WORKSTREAM	EVALUATION COMPONENT	EVALUATION QUESTIONS	EVALUATION MATRIX LOCATIONS: SECTION	EVALUATION MATRIX LOCATIONS: PAGE
Workstream 1: IE	LEWIE model	Impact: f	Section 2.2	12
	Quantitative impact analysis	Impact: a, b, d, e	Section 2.3	14
	Qualitative analysis	Impact: a, b, c, d, e, g, k, s	Section 2.5	29
	Special study on the ASP	Impact: e Effectiveness: m	Section 2.7	40
Workstream 2: Operational M&E	Process and institutional capacity assessment	Effectiveness: l, q	Section 3.2	45
	Costing study	Efficiency: n, o	Section 3.3	48
	Bi-monthly Operational monitoring and endline report	Effectiveness: i, j, k	Section 3.6	56
Workstream 3: Policy analysis	Targeting study	Effectiveness: i	Section 4.2	60
	Simulation analysis	Sustainability: r	Section 4.4	67
	Strategic review	Impact: h Sustainability: t, u	Section 4.5	69

Notes: Further to discussion with stakeholders during and following the inception mission, evaluation questions d, n, o and t have been modified, and question p has been dropped. See Table 2 for details.

The original ToR also requested that the evaluation include an evaluation of the DFID ASP. The ASP was originally envisaged as a programme that would complement the HSNP by focusing on longer-term resilience to drought and shocks – including supporting community adaptation strategies, establishing livestock insurance, encouraging livelihood diversification and using the new HSNP mechanism to provide emergency scale-up payments in times of drought. However, in reality the overlap between the HSNP and the ASP is very minimal due to the separation of the two programmes' design and implementation activities. This said, as the ASP is a significant programme in its own right, an assessment of it is incorporated into the evaluation of the HSNP in two ways. First, a special study will be conducted to assess the degree to which the ASP has been influential in county-level planning and budgeting processes. Second, the mixed-methods IE of the HSNP will include a component looking at the interaction of the HSNP with other interventions, including the ASP.

In addition to the above, and in response to requests made by DFID, two further elements have been incorporated into the activities to be conducted under Workstreams 2 and 3. These are:

1. a process review of the two pilot emergency payments that have been made so far, to complement the assessment of impact of the emergency payments that is built into the mixed-methods IE (see sections 2, 0 and 0 below); and

2. a review and redesign of the registration instrument to feed into DFID's ongoing design activities for HSNP Phase 3.

1.4 RISKS TO THE EVALUATION

The evaluation approach described below comprises a series of technical studies, some of which are largely desk based and some of which require primary data collection activities to be conducted in Kenya. Each of these studies contains its own set of particular risks and challenges around availability and quality of data, but the biggest risk to the evaluation is the general security situation in Kenya, and particularly in the two north-eastern counties of Mandera and Wajir that are both covered by the HSNP.

The evaluation team structure comprises a mix of leading international expertise and Kenyan knowledge and experience. This is embodied in the pairing of an international TL and full-time KNPC, both based in Nairobi for the duration of the project. All of the field research will be conducted by an experienced Kenyan field research company, RGA, with continuous and extensive involvement from relevant international team members under each of the four proposed workstreams.

While OPM does not have a DoC for the field research team (this is the responsibility of RGA, as specified in its sub-contract with OPM), the programme will employ a core evaluation team of 12 experts, supported by a pool of senior experts, all of whom will come under OPM's DoC. In order to deliver the project successfully, in addition to the TL and KNPC based in Nairobi, other members of the core evaluation team and senior experts will be required to work and travel to Kenya and potentially into a number of high and extreme risk areas throughout the country. While it is likely that travel will initially be contained within Nairobi and some of the larger towns/cities, travel into more remote and potentially riskier areas may be required under some of the evaluation components as the project progresses.

Security in each of these locations must be managed to a standard that meets the requirements of the OPM DoC statement and DFID's DoC requirements as set out in the HSNP ToR. To this end, OPM, in partnership with its security adviser Spearfish, has prepared a comprehensive security plan and set of standard security operating protocols (see 0).

OPM reserves the right to suspend operations in the case of security concerns. Such a decision would take place in discussion with DFID and other relevant stakeholders.

1.5 STRUCTURE OF THE REPORT

The remainder of this document is structured as follows: Section 2 describes our approach to the IE. Section 3 presents the design for the operational M&E. Section 4 describes the various policy analysis studies to be undertaken. Section 5 details how the evaluation will ensure that the key programme stakeholders are able to engage with and make use of the various outputs and data produced by the evaluation. Section 6 presents the evaluation work plan and timeline. A short glossary provides brief explanations for some of the technical terms used in this report. 0 provides supporting technical analysis for the IE. 0 provides information regarding the evaluation security plan. 0 includes an extract from the evaluation ToR (purpose and objectives, as well as scope of work).

TABLE 2 FINALISED EVALUATION QUESTIONS MAPPED TO WORKSTREAMS AND EVALUATION COMPONENTS

DAC CRITERIA	ORIGINAL EVALUATION QUESTION	CHANGES / COMMENTS	WORKSTREAM	EVALUATION COMPONENT
Impact	a. What are the overall effects of the CTs in terms of: consumption, asset retention, wellbeing (PMT as measure), food diversity, health, education and self-esteem, comparing beneficiaries and non-beneficiaries?	Focus on consumption, poverty, asset retention/ accumulation, nutrition (dietary diversity), financial inclusion (saving, borrowing and credit), subjective wellbeing, social networks, conflict/social tension.	WS1: IE	Quantitative impact analysis Qualitative analysis
	b. For which sub-groups are effects most pronounced (taking account of poverty status, household size, family composition, geographic location, livelihood base, gender and disability)?	Heterogeneity of impact will be analysed across a number of dimensions: poverty status; household size; programme exposure ('smooth' vs 'lumpy' payments) – this latter dimension will depend upon the distribution in the data.		Quantitative impact analysis Qualitative analysis
	c. (How) Do CTs impact on women's control of cash within their (often polygamous) households and their wider empowerment?			Qualitative analysis
	d. How do the effects of predictable transfers compare with those of short-term transfers triggered in response to acute shocks? How do the larger one-off transfers some households will receive due to the later than anticipated start of the programme impact on those households?	We will only be able to address the first element of this question qualitatively, as we have no way of quantitatively comparing the impacts of the regular HSNP transfers with those of the emergency transfers, since these are targeted at distinct population groups. The second element we modify to: How do the larger one-off transfers some households will receive due to the later than anticipated start of the		Quantitative impact analysis Qualitative analysis

		programme impact on those households?		
	e. Does the combination of CTs and wider livelihoods activities open up new livelihoods opportunities/income-generating activities for poor households? How?	We will assess this question qualitatively, supplemented by descriptive analysis of the quantitative data where feasible (depending on the degree to which the household sample contains households also benefitting from other livelihood interventions).		Qualitative analysis Quantitative analysis Special study on the ASP
	f. What kinds of multiplier effects are found in local economies?			LEWIE
	g. Is there evidence of the programme having an impact on community relations – both within and between communities?			Qualitative analysis Quantitative impact analysis
	h. To what extent does NDMA policy dialogue contribute to changes in the budget allocations of the GoK in respect of the NSNP? (this is also a sustainability issue)		WS3: Policy analysis	HSNP strategic review
Effectiveness	i. How well does the system based on use of HSNP registration data work as a basis for targeting the transfers and selecting households for a triggered drought response payment?		WS2: Operational M&E	Quarterly operational monitoring and endline report
			WS3: Policy analysis	Targeting study
	j. How well does the rights and grievances process work?		WS2: Operational M&E	Quarterly operational monitoring and endline report
	k. Do the new payment platform and expansion of financial services provide benefits for beneficiaries and non-beneficiaries?		WS1: IE	Qualitative analysis
			WS2: Operational M&E	Quarterly operational monitoring and endline report

	l. Do the HSNP management structures and processes work well and provide value for money?	We will answer this question in relation to the capacity of the implementing partners (NDMA/PILU, HAI, FSD, Equity Bank) and the objectives and expectations of the funding agencies (DFID, DFAT, GoK).	WS2: Operational M&E	Process and institutional capacity assessment
	m. How effective is the work with NGOs to understand how communities respond to shocks under ASP? Does this get reflected in county plans?		WS1: IE	Special study on the ASP
	n. Are triggered payments more/less cost effective than longer-term predictable CT payments?	Modified: this question aspires to a direct comparison of the impact of emergency versus non-emergency payments, but such a comparison is not possible because the payments are targeted to different types of households, for different lengths of time and under different circumstances. However, we are able to measure cost-efficiency distinguishing between routine and emergency payments.	WS2: Operational M&E	Costing study
Efficiency	o. What are the costs and benefits of HSNP 2? Can we undertake cost–benefit analysis (CBA) using disability-adjusted life years (DALYs), rate of return on education, and changes in welfare or assets?	Modified: How much has it cost the HSNP to deliver Phase 2 of its programme, for regular beneficiaries and for emergency payments? It has been agreed that the original question is not appropriate because it requires assigning a monetary value to non-monetary programme benefits such as the sense of dignity or security, which would not be part of a financial costing study. The 'benefit' component of a CBA is dependent on the IE, which will not be collecting data on e.g. the education benefits and changes in life	WS2: Operational M&E	Costing study

			expectancy referred to in the proposition.		
	p.	Do the gains associated with a package of complementary activities to support livelihoods and CTs represent value for money compared with just providing the CTs?*	Dropped: this question implies a quantitative estimate of the impact of the ASP, which we have agreed will not form part of this evaluation or the costing study.	N/A	N/A
	q.	Is the NDMA developing the capacity to deliver HSNP in the future?	We will look at the capacity of the NDMA as it pertains to delivery of the HSNP Phase 2. A separate consultancy has been commissioned to assess capacity requirements for the NDMA going forward.	WS2: Operational M&E	Process and institutional capacity assessment
Sustainability	r.	Can we learn anything about the transfer value for future transfers? Could a lower amount have a similar impact for some households? Could more have sufficient impact to justify the extra cost?		WS3: Policy analysis	Simulation analysis
	s.	Do the reliable CTs build people's resilience to climate variability?	Moved: now under impact questions.	WS1: IE	Qualitative analysis
Relevance	t.	In areas where the HSNP is linked to other activities, is the combination of a targeted CT, mechanism for additional triggered CTs and additional work on livelihoods an appropriate response to the situation in the four counties?*	Modified: Is the combination of a targeted CT with a mechanism for additional triggered support an appropriate response in the four programme counties? It will not include a quantitative estimate of the contribution of additional complementary livelihoods interventions as this question would require a quantitative evaluation of the ASP, which has been omitted by agreement. The latter part of the original question may be	WS3: Policy analysis	Strategic review

		answered in a partial way depending on how possible an assessment of the interaction between the HSNP and ASP is under the qualitative component of the IE (see evaluation question e above).		
u.	Where comparison is possible, how does the cost/ impact of HSNP/ASP compare with other interventions (e.g. food aid) used to support households in the four counties?	We will compare cost-efficiency (cost of dollar transferred and cost per transfer) of HSNP vs. World Food Programme (WFP) food aid if those data are available from the WFP.	WS3: Policy analysis	Strategic review

2. Workstream 1: IE

2.1 OBJECTIVES OF THE WORKSTREAM

The evaluation of HSNP Phase 1 clearly demonstrated the impact of the CTs at the beneficiary level with the use of a robust experimental design. However, the methodology did not allow for a comprehensive analysis of the broader effects of the programme on the local economy. As HSNP scales up under Phase 2 it may generate a wide spectrum of effects at different levels and for different groups within society and households. Our proposed evaluation approach for HSNP Phase 2[4] thus places a stronger focus on this aspect. This takes the form of a multi-method IE that aims to disentangle this variety of effects and understand through which channels and with what observable results the HSNP is changing the lives of people in north-eastern Kenya.

The IE workstream will address the evaluation questions reported in the evaluation matrix presented above each workstream activity. The aim is a comprehensive evaluation of the effects produced by the scaled-up HSNP CTs on targeted households and individuals, as well as on the communities and local markets in which they live and work. A range of analytical methods will constitute the basis for a mixed-methods evaluation of the impact. While the qualitative approach will rely on multiple rounds of qualitative research based on participatory methods, we will undertake a single round of quantitative data collection based on a household survey that will underpin both the LEWIE and the quantitative IE. The quantitative IE will thus be based on a single round of post-treatment data collection, in the form of a large household survey. There is no scope for a 'before' and 'after' comparison (i.e. pre- vs post-treatment) since the current phase of HSNP began in July 2013, well before the start of this evaluation.

The different methods are specifically associated with the relevant evaluation objective that they are expected to achieve.

- The first objective of the workstream is to investigate the wider effects of the HSNP CTs on the local economy through the use of a **LEWIE**. This approach will enable us to understand the local supply response to the anticipated increase in demand due to the injection of cash into the local markets brought about by the HSNP transfer. On the one hand, there could be positive economic spill-overs in the surrounding economy, giving rise to a local income multiplier and amplifying the HSNP's overall impact. On the other hand, if the local supply is not responsive to the increase in demand, the programme's benefits could be undermined by higher prices. **The LEWIE analysis will therefore shed light on the potential multiplier effects of the HSNP**, by simulating HSNP impacts on the entire local economy and on groups of households and production activities (see Section 2.2.4 for more information on the LEWIE method).
- The second objective of the workstream is to assess the effects of the CTs at the beneficiary level using a **quantitative IE approach based on a RD design.** (RD is explained in more detail in Section 0 below and a summary description of RD is provided in the glossary.) Quantitative data will be collected for this impact analysis through the same survey underpinning the LEWIE analysis. The considerable size of the sample used for this IE will provide the opportunity to also carry out a sub-sample analysis of the HSNP impact across a number of household categories of interest (see Sections 2.3.1 and 2.4 for more information on this).
- **Multiple rounds of qualitative research deploying participatory methods** will complement the quantitative approach by: providing an **understanding of the context** within the programme is operating, and how this affects and is affected by the CT; **capturing experiences and processes that produce outcomes of interest**; enabling an **assessment of impacts that are difficult to cover quantitatively** (such as social cohesion and inter- and intra-household relations); and providing complementary data on some of the topics covered by the quantitative survey, and in doing so **triangulating, validating and providing depth** to the quantitative findings.

[4] We will refer to HSNP Phase 2 as HSNP throughout.

- A **special study on the ASP** to look at the interaction of the HSNP with a package of complementary activities to support livelihoods as well as provide an insight into the contribution the ASP makes to county planning and budgeting processes.

The remainder of this section outlines, for each proposed activity, the timeframe, the research questions it proposes to answer, its intended audience and use, and the research method.

2.2 LEWIE

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DATA ANALYSIS METHODS	DAC CRITERIA	EVALUABILITY ISSUES
f. What kinds of multiplier effects are found in local economies?	Is there an increase in aggregate demand in the local economy due to the HSNP CTs? Is the local supply responsive to the increase in demand? Who benefits from any multiplier effects? How?	Total amount of income created by HSNP cash transferred to eligible households: types and number of households that receive the benefits of project spill-overs; effects of other interventions that increase the local supply of food and other goods and services	Data from surveys of HSNP beneficiaries as well as of households and businesses in the surrounding economy	Quantitative survey (same survey used for the quantitative IE) comprising household and business instruments	LEWIE simulation modelling	Impact	High. A robust LEWIE model will be able to simulate the HSNP's impacts on the entire local economy, as well as on individual groups of households and production activities.

2.2.1 Rationale

HSNP injects a considerable amount of cash into households and through them, into local economies. By transferring cash to beneficiary households, the programme also affects the local economy to which those beneficiary households belong. Viewed from a local economy-wide perspective, the beneficiaries are a conduit through which cash is channelled into the local economy. As the beneficiary households spend their cash, the programme's impacts spread from the cash recipients to the surrounding economy, in the form of an increased demand for goods and services.

If venders and producers expand their supply to meet this new demand, we would expect the HSNP to contribute to local economic expansion. Purchases in local stores and markets can create positive economic spill-overs within the surrounding economy, as suppliers' profits and demand for labour and other inputs increase. As profits and wages rise, so does the demand for goods and services by non-beneficiary households. The result may be a large local income multiplier.[5] In other words, the HSNP's total impact on income in the local economy may exceed the amount transferred to the beneficiary households. If this is the case, the HSNP would not only fulfil the function of protecting vulnerable households but also stimulate local economic growth.

On the other hand, if the local supply is not responsive to increase in demand (for instance because local markets are very remote), there could be a situation of more money chasing the same quantity of goods, leading to inflation instead of real growth in the local economy. If this happens, the programme's real benefits will be muted by higher prices.

[5] For example, a regional LEWIE for the CT-OVC programme produced income multipliers of KES 1.34 and 1.81 in the west and east, respectively (FAO, 2013).

Higher prices would erode the programme's real benefits for the beneficiary households. They would also dampen the real benefits for non-beneficiary households, many of which contain both producers (who could benefit from higher prices for the goods they sell) and consumers (who are harmed by higher consumption costs). Because the local supply response is important in shaping CT impacts (especially in the case of HSNP given the constraints to market integration in northern Kenya), it is critical to understand the supply response, identify any constraints on expanding local supplies in response to increases in demand, and take measures to alleviate these constraints. For example, complementary interventions enabling local pastoralists to increase their production might increase the programme's benefits for both beneficiaries and non-beneficiaries.

Finally, some impacts of CTs leak out of the project area, potentially unleashing income multipliers in other parts of Kenya. For example, vendors who sell to beneficiaries or to non-beneficiaries may acquire their inventories from other parts of the country. Market linkages potentially may spread the effects of the HSNP to other regions in Kenya.

The LEWIE study provides an assessment and measure of each of these dimensions of the possible impact of the HSNP on the local economy.

2.2.2 Timeframe

The LEWIE study will be undertaken between July 2015 and September 2016. The full study design will be developed and finalised between July and January 2016. Fieldwork will be undertaken between February and April 2016. Data processing and analysis will take place between April 2015 and July 2016, with the report submitted at the end of September 2016.

2.2.3 Research questions

The key research question to be answered by the LEWIE is: What is the multiplier effect of the HSNP CT on the local economy in programme areas? The indicators to be produced by this component of the evaluation include the following:

1. The total amount of income created by each KES transferred to eligible households – that is, the real income multiplier. This multiplier will exceed 1.0 if, in addition to each KES transferred, the HSNP creates positive spill-overs within the local economy and does not lead to significant price inflation. For example, a multiplier of 1.5 would indicate that, for every KES transferred to an eligible household, total income in the project area rises by the KES 1 transferred plus an additional KES 0.5.

2. A disaggregation of this multiplier to ascertain how the benefits created by the HSNP get distributed among households in the local economy. All of the direct benefits accrue to the beneficiary households, but which households receive the benefits of project spill-overs, if present? For a multiplier of 1.5, KES 1 goes to the beneficiary households; but of the KES 0.5 of spill-overs how much goes to HSNP beneficiaries and how much to non-beneficiaries? Both are important. To the extent that part of the spill-over goes to beneficiary households, the programme's impacts on the beneficiary households include some of the spill-over. To the extent that spill-over goes to non-beneficiary households, many or even most of the programme's economic impacts may be missed by a conventional IE focusing on eligible households. Many or most of the ineligible households are poor–just not poor enough to qualify for the HSNP. It is important to document how they, as well as the beneficiary households, might benefit from this programme.

3. Estimates of the potential impacts of complementary interventions on programme impacts, particularly interventions that increase the local supply of food and other goods and services. For example, as part of the simulation analysis under Workstream 3 we will estimate the change in the local income multiplier that might result from the HSNP emergency transfers. Complementary interventions need not be limited to HSNP-eligible households. The LEWIE model can thus be used to explore the market multiplier implications of changes in

targeting, scale, or other potential future modifications to the HSNP, which will be done as part of the Workstream 3 simulation analysis (see Section 4 below).

2.2.4 Method

The LEWIE methodology was designed to understand the full impacts of projects on local economies, including on the income and welfare of those who are affected indirectly. In this case the LEWIE of the HSNP will embed models of beneficiary and non-beneficiary households within a model of the local economy, highlighting the economic linkages between the two. Many households are both producers and consumers of agricultural and livestock goods. Agricultural household modelling provides a way to represent this dual nature of many rural households. Constructing agricultural household models for beneficiaries and non-beneficiaries entails estimating production functions, on one side, and household expenditure functions, on the other. Once the household models have been constructed, the supply and demand for goods and services can be summed up across households to obtain total supply and demand in local markets. For goods traded with outside (regional or national) markets, prices are given, and the difference between local supply and demand determines net sales or 'exports' to outside markets. For non-traded goods and most services, the interaction of local supply and demand determines prices as well as quantities supplied and demanded. The novelty of LEWIE that it embeds models of heterogeneous households into a model of the whole local economy in this way, making it possible to evaluate outcomes at both the household and local economy levels.

The LEWIE model will be estimated using data from surveys of beneficiaries as well as of households and businesses in the surrounding economy. Once the LEWIE model is constructed, it will be used to simulate the HSNP's impacts on the entire local economy, as well as on individual groups of households and production activities.[6]

2.2.5 Output

The output from the LEWIE will be a report detailing the methodology and findings from the study and representing the impact estimate of the HSNP in terms of the multiplier effect of the programme. The report will incorporate findings from the first round of qualitative research as relevant (see Section 0 below).

The LEWIE model will also be used to model the local multiplier impacts of policy changes as part of the simulation modelling to be undertaken under Workstream 3 (Section 4).

2.3 QUANTITATIVE IMPACT ANALYSIS

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DATA ANALYSIS METHODS	DAC CRITERIA	EVALUABILITY ISSUES
a. What are the overall effects of the CTs on household and individual welfare, comparing beneficiaries and non-beneficiaries?	What are the effects in terms of consumption, poverty, asset retention/ accumulation, nutrition (dietary diversity), and financial inclusion (saving, borrowing and credit)?	Average per adult equivalent consumption; proportion of households under the poverty line; average dietary diversity score; proportion of households owning assets (by type); proportion of households	HSNP beneficiaries and non-beneficiaries, with data collected at household and individual levels.	Multipurpose quantitative household survey, with thematic modules on demographic characteristics, consumption, assets, agriculture, and use of financial services	Econometric analysis (quasi-experimental RD approach, with propensity score matching (PSM) as a back-up)	Impact	Medium to high. There are methodological challenges related to the feasibility of the RD approach (i.e. existence of discontinuities at baseline) that may affect the robustness of the analysis.

[6] Monte Carlo methods, developed by Taylor and Filipski (2014), will be used to construct confidence bounds around these simulated impacts. See the glossary for a description of Monte Carlo methods.

		owning livestock (by type); proportion of households with access to formal credit; proportion of households with formal savings					
b. For which sub-groups are differences in wellbeing most pronounced?	Are there observable differences in impact over poverty status, household size, family composition, geographic location, livelihood base, gender and disability?	As above	As above	As above	Extended RD design analysis: estimation of heterogeneous local average treatment effect	Impact	Medium to high. Given the considerable size of the sample used for the IE, we will be able to carry out a heterogeneity analysis between sub-samples that represent half the size of the main sample. Unbalanced and smaller sub-samples will be analysed descriptively.
d. How do the larger one-off transfers some households will receive due to the later than anticipated start of the programme impact on those households?	Are there observable differences in impact over transfer size and frequency?	As above	As above	As above	As above	Impact	Unless these sub-groups are sufficiently similar in size (approximately half of the main sample), a robust heterogeneity analysis cannot be guaranteed and their analysis will only be descriptive. A preliminary analysis of the MIS shows that 67% of household received a payment in Mar 2014, while 31% had still not received any payments in Jan 2015.
e. Does the combination of CTs and wider livelihoods activities open up new livelihoods opportunities/	What are the effects in terms of improved livelihood opportunities and income-	Proportion of individuals engaged in paid employment; proportion of individuals engaged in	As above	As above	Econometric analysis (quasi-experimental RD design approach, with PSM as a back-up)	Impact	Medium to high. There are methodological challenges related to the feasibility of the RD design approach that

income-generating activities for poor households? How?	generating activities?	business activities	may affect the robustness of the analysis.

2.3.1 Rationale

The HSNP Phase 1 evaluation provided a comprehensive assessment and quantitative estimates of impact across a wide range of domains. In Phase 2 the programme has undergone some changes, including the types of household to be targeted and the value of the transfer (which has increased over time). It thus remains important to monitor impact on a number of core impact areas as the HSNP has evolved over time. These include consumption and poverty, food security, asset retention and building, and financial inclusion.

The ToR were very clear that obtaining robust quantitative measures of programme impact was a priority for this evaluation, something that was confirmed during discussions with programme staff, DFID and other stakeholders during the inception phase. The quantitative impact analysis is therefore designed to provide robust impact estimates across these core domains, based on a quasi-experimental RD approach. In addition, a heterogeneity analysis will be undertaken to detect the HSNP impact on sub-samples of households belonging to categories of interest, such as smaller and larger households or poorer and less poor households.

2.3.2 Timeframe

The quantitative component of the IE will be undertaken between July 2015 and September 2016. The study design will be developed and finalised between July 2015 and January 2016. Fieldwork will be undertaken between February and April 2016. Data processing and analysis will take place between April and August 2016, with the report submitted at the end of September 2016.

The timing to the IE study balances the need for the results to feed into the design of Phase 3 (which will take place in 2017), with the desire to schedule the IE as late in the evaluation period as possible, so as to maximise the time for impact to occur.

2.3.3 Research questions

The broad research question that will be addressed by the quantitative component of the IE is focused on a post-treatment comparison between treatment and control groups, in terms of general welfare as captured at the household and individual levels. In particular, the proposed quasi-experimental RD design allows us to measure observable differences between HSNP beneficiaries and non-beneficiaries over a range of welfare indicators and for a number of relevant sub-samples. This enables us to answer three crucial research questions:

1. What are the overall effects of the CT on household and individual welfare, comparing beneficiaries and non-beneficiaries?
2. What are the specific effects in terms of consumption, poverty, nutrition (dietary diversity) and financial inclusion (saving, borrowing and credit)?
3. What are differences in the effects produced by the CTs on household sub-groups, including smaller and larger households or poorer and less poor households?

The analytical levels of these research questions are therefore households and the individuals within these households, and confidently answering these questions provides an understanding of the quantitative impact of the HSNP transfers on the intended beneficiaries. What difference does it make for beneficiary households and individuals to receive the CT on their overall wellbeing? What would have happened to their degree of wellbeing if they had not received the transfer

(control group)? Are the effects more or less pronounced at the household level or the individual level? What domains of household and individual welfare are found to disproportionately benefit from the extra inflow of cash resources? What types of households benefit more from the CTs? What types of households benefit disproportionately less?

Given the considerable size of the sample used for the collection of quantitative data, our impact estimation models will have the power to detect differences in outcome indicators between treatment and control groups belonging to separate categories. However, as mentioned above and explained in detail in Section 0 on the sampling strategy, the analysis will provide a robust disaggregation of the HSNP impact only for household groups that represent half the size of the main household sample. For other potential sub-samples of households that are unequal in size (for example households that have received different types and levels of exposure to the programme, i.e. 'smooth' vs. 'lumpy' payments), the ability to detect impact will depend on the distribution of the data.[7] If impact heterogeneity variations cannot be estimated robustly using the RD approach, the data collected through the quantitative survey will form the basis for a descriptive analysis of differences in mean values for a range of outcome variables. These descriptive statistics will represent a summary analysis of average characteristics of different individual and household groups and will facilitate the sub-group analysis of qualitative information.

2.3.4 Method

2.3.4.1 Why an experimental (RCT) approach is not feasible

The main challenge to be addressed when designing a rigorous quantitative IE is the identification of a statistically valid counterfactual. This is needed in order to answer the question as to what would have happened to the beneficiaries if they had not received the intervention. Answering this question requires the construction of a 'control' group formed by a group of households that are as similar as possible to the HSNP beneficiary households, with the only difference being that they do not receive the HSNP.[8]

The so-called 'gold standard' approach for constructing a control group in IEs is to use an experimental or RCT design, under which beneficiary status is randomly allocated across all eligible households. This was the approach used for the IE of the first phase of HSNP. This was made possible by the fact that the programme was in its initial phase and was therefore not operating across all areas within the four HSNP counties. There were therefore communities containing eligible households that were not being covered by the programme, from which the control group could be drawn.

However, under Phase 2 the programme was taken to scale, with the intention of covering the entirety of the four HSNP counties. In fact, the HSNP secretariat registered almost 375,000 households, with the target of enumerating all qualifying households within the four counties. Therefore for this evaluation it is not possible to randomly allocate a set of eligible households to not benefit from the programme, which would provide us with a control group.

In summary, an experimental (i.e. RCT) IE approach is not feasible in this context. Therefore, in order to produce robust quantitative measures of programme impact we have to rely on alternative, so-called 'quasi-experimental' approaches.

---

[7] A preliminary analysis of the MIS data supplied to the evaluation team in January 2015 showed that approx. 67% of (Group 1) beneficiary households had received their first transfer by March 2014, while approx. 31% of (Group 1) households had still not received any transfers in January 2015.

[8] Hence, treatment and control groups must have the same characteristics on average and would have reacted to the programme intervention in the same way if they had both been targeted. In addition, for the two groups to be comparable it is also important that they should not be differentially exposed to other interventions during the period of the evaluation. A control group that satisfies the above conditions will be robust to selection bias. This is a central condition for achieving robust impact estimates, as systematic differences between treatment and control groups (such as unobserved differences in ability or political connections) would mean that the evaluator could not be sure whether observed differences in key impact indicators are a causal impact of the HSNP, or the result of pre-existing differences between treatment and control groups.

The key design features and considerations relating to our proposed approach (RD design) are set out below in the subsequent sub-sections of Section 0.

2.3.4.2 RD design

The previous sub-section sets out the reasons that an experimental (i.e. RCT) IE approach is not feasible in this context. As a result, during the inception phase we comprehensively scoped out the feasibility, relative advantages and risks of alternative quasi-experimental approaches. We have determined that an RD approach is the best option for providing robust measures of programme impact. Specifically, we are proposing a 'fuzzy' RD design with the inclusion of instrumental variables. The key features of this approach and the feasibility assessment analysis that provided the basis for this decision are set out in this sub-section and the remainder of Section 0.

It is the judgement of the evaluation team that the RD approach is the best available option for providing robust quantitative measures of programme impact. However, it should be noted that the RD approach is not without risks (set out in Section 0). Therefore we have built into our design the potential for using an alternative quasi-experimental approach (PSM) as a fall-back option for the quantitative IE study methodology. The full details of this back-up PSM approach are set out in Annex A.2.

The RD method exploits one of the key design features of the HSNP Phase 2 beneficiary selection process, namely the use of a PMT score as one of the eligibility criteria. The PMT score is a measure of a household's relative socioeconomic status. A PMT score is constructed by conducting statistical analysis to estimate the relationship between a household's measured consumption expenditure and a restricted set of readily observable household characteristics, in this case based on the latest Kenya Integrated Household Budget Survey (KIHBS) data. This analysis provides a set of weights that can be used to construct the PMT score –which technically speaking is a predicted measure of household consumption expenditure– based on the same restricted set of household characteristics, which were captured as part of the HSNP Phase 2 household registration process specifically for this purpose. As mentioned above, the registration process was intended to cover all households in the four HSNP counties. The programme's MIS data contains the registration data and calculated PMT score for all 375,493 registered households, and provides the basis for the RD approach.

Household were selected onto the programme based on a combination of the PMT score and a CBWR.[9] The way the PMT score and CBWR ranking are combined to determine household eligibility is as follows:

1. The number of households within a village with a PMT score below the county-specific PMT threshold provides the absolute number of households that are selected within that village.[10]
2. The PMT score is then used to create four PMT-based income groups within each village, with the range of PMT scores equally divided into four and used to assign households into four corresponding groups. The households with the lowest PMT scores in the village are assigned to Group 1 ('very poor'), the households in the next range of PMT scores to Group 2 ('poor'), those in the next range to Group 3 ('middle'), and finally those with the highest PMT scores to Group 4 ('better off').
3. In parallel to this, community members in each village are asked themselves to rank each household into one of four CBWR groups on the basis of their relative wealth, with 1 being poorest, and group 4 those better off.[11]

[9] We use the term CBWR, as opposed to the more standard term community-based targeting (CBT), in line with the programme's terminology.

[10] Each of the four counties was assigned its own specific cut-off point on the basis of the distribution of poor households across the four counties, adjusted using a modified version of the CRA formula that the GoK uses to allocate resources to devolved county governments (see Section 0).

[11] For households in Marsabit, a preceding step was required before the wealth group merging could take place, as the CBT undertaken by CARE did not produce four wealth groups but an individual household rank.

4. The PMT-based income groups and the CBWR wealth groups are then combined to produce a final PMT-CBWR 'ranking' group for each household.[12]

5. Households are then ordered by their 'ranking' group (lowest to highest), and then within each 'ranking' group by their PMT score (lowest to highest). Finally, households are selected into the programme based on this final ranking up to the village's beneficiary quota (defined under Step 1).

Although the final targeting mechanism is based on a combination of both the PMT score and a CBWR, our analysis shows that beneficiary selection is associated closely enough with a household's PMT score alone to enable it to be used as the basis for a RD approach (see Section 0 below).[13]

The RD design works by exploiting the county-specific eligibility threshold for the PMT-based part of the selection process to identify statistically valid treatment and control groups. In broad terms, we compare households just below the PMT eligibility threshold (treatment households) with households just above the PMT eligibility threshold (control households). However, since programme eligibility is based on both a household's PMT score and its CBWR group, being above or below the PMT threshold does not correspond exactly to beneficiary status. In other words, due to the way the targeting system works, some beneficiary households do not meet the PMT eligibility criteria (i.e. they have a PMT score above the PMT threshold) but still qualify due to their CBWR ranking. Conversely, there are some households with a PMT score below the PMT threshold (i.e. are eligible according to the PMT score) that do not qualify due to their CBWR ranking.

Since the PMT score (referred to technically as the 'forcing variable') does not perfectly predict the beneficiary status of the households, this requires the adoption of a 'fuzzy' RD design. Under a 'fuzzy' RD design, some members of the group that is considered the treatment group in the analysis (i.e. those households below the PMT threshold) are not receiving the transfer ('no-shows'), and some members of the analytical control group (those households above the PMT threshold) are receiving the transfer ('crossovers').

In theory a 'sharp' RD design –whereby being above or below the PMT threshold corresponds perfectly to beneficiary status– should have been feasible. This would have entailed using retrospectively constructed implicit village-level PMT thresholds, applied to a restricted category of households (those in the same combined 'ranking' group as the beneficiary household with the highest PMT score in that village). However, despite extensive efforts to clean the data, it was impossible to accurately identify each household's village in the programme's MIS dataset.[14] Without being able to group households in the dataset by their village, the 'sharp' RD approach based on constructed village-level PMT thresholds was not feasible.

2.3.4.3 Impact indicators

As mentioned above, the HSNP Phase 1 evaluation provided a comprehensive assessment and quantitative estimates of impact across a wide range of impact areas. In order to enable the use of a single household questionnaire that would accommodate the needs of both the quantitative impact analysis and the LEWIE study, it was agreed during the inception phase that the Phase 2 evaluation will focus a restricted number of impact areas that relate to the core aims

[12] In Turkana, Wajir and Mandera. If household size>2 or no children <18 then final wealth group=the average of the PMT and CBT wealth group number ((PMT_WG+ CBT_WG)/2). If household size≤2 and no children <18 then final wealth group=CBT wealth group.

[13] Using the combined PMT/CBWR ranking as the basis for the RDD was considered but was not possible for two reasons: (1) The ranking variable is not continuous (it is an ordinal measure taking discrete integer values), which creates complications for the RDD analysis; (2) The ranking variable cut-offs would have to be defined at village level and this was not possible due to village name errors in the MIS dataset which meant that households' villages could not be accurately assigned.

[14] Around 60% of the villages named in the MIS dataset presented some spelling and data entry mistakes, which even after extensive data cleaning led to the incorrect grouping of households by village.

and objectives of the programme. These are: consumption and poverty; food security; asset building and retention; and financial inclusion.

The indicators we are proposing to measure are listed in Table 3. As far as possible we will use the same indicators as the Phase 1 evaluation to maximise comparability. The precise specification of these indicators may change slightly during the design phase as we seek also to ensure optimum comparability and complementarity with other evaluation data collection efforts currently being conducted under the NSNP. In addition, a selection of intermediate and contextual data (not listed below) will be gathered and analysed descriptively.

TABLE 3 INDICATORS PROPOSED FOR QUANTITATIVE IMPACT ANALYSIS

IMPACT AREA	INDICATOR
Consumption and poverty	Mean total monthly household consumption expenditure per adult equivalent
	Proportion of households in bottom national decile
	Proportion of households below absolute poverty line
	Poverty gap
	Severity of poverty
	Mean monthly per-capita health expenditure per household
	Mean monthly household education expenditure per child
Food security	Mean monthly food consumption expenditure per adult equivalent
	Mean food share of consumption expenditure
	Mean food consumption score (FCS) of dietary diversity
	Mean Food and Nutrition Technical Assistance Project household hunger score
	Proportion of households that were food insecure in worst recent food shortage period
Asset building and retention	Proportion of households owning livestock assets (goats/sheep, camels, cattle)
	Proportion of households owning non-livestock productive assets (animal cart, water drum, plough, wheelbarrow, sickle, pick axe, axe, hoe, spade, machete)
	Proportion of households purchasing livestock (goats/sheep, camels, cattle) in last 12 months
	Proportion of households selling livestock (goats/sheep, camels, cattle) in last 12 months
	Mean value of livestock purchased in last 12 months
	Mean value of livestock sold in last 12 months
	Proportion of households purchasing non-livestock productive assets in last 12 months
	Proportion of households selling non-livestock productive assets in last 12 months

	Mean value of non-livestock productive assets purchased in last 12 months
	Mean value of non-livestock productive assets sold in last 12 months
	Households currently owning agricultural land
Financial inclusion	Proportion of households that currently have cash savings
	Mean value of cash savings
	Proportion of households saving via formal bank account
	Mean value of formal bank account savings
	Proportion of households saving via HSNP bank account
	Mean value of HSNP bank account savings
	Proportion of households that have borrowed money in last 12 months
	Mean value of outstanding debt
	Proportion of households that have taken out formal loan in last 12 months
	Mean total value of all formal loans taken out in last 12 months
	Proportion of households that bought something on credit in last three months
	Mean value of outstanding credit debt

2.3.4.4 Impact heterogeneity

To understand how impact varies across different types of households, a heterogeneity analysis will be undertaken. Specifically this analysis will attempt to assess variations in programme impact by:

1. household size (i.e. smaller vs larger households);
2. socioeconomic status (i.e. poorer vs less poor households);
3. Phase 1 HSNP status (i.e. Phase 1 HSNP beneficiaries vs. Phase 1 non-beneficiaries);
4. transfer 'lumpiness' (i.e. those who received regular bi-monthly payments vs. those who experienced delays and therefore received a large initial one-off transfer due to accumulated back payments); and
5. sex of household head.

These impact variations were identified as priority areas for the analysis based on discussions with programme staff and DFID during the inception phase.

To maximise the power of this analysis, where possible it will be based on defining the sub-groups as two groups of equal size. This is feasible for household size and socioeconomic status, but is not possible for the analysis of variations by Phase 1 HSNP status, payment 'lumpiness', or sex of household head, since these distinctions do not correspond to even splits in our treatment and control groups. Specifically:

- Only around 20% of households registered for Phase 2 were beneficiaries under Phase 1.

- Our analysis of the programme's January 2015 MIS dataset indicates that 67% of Phase 2 beneficiaries had received regular payments as intended, with no accumulated 'lumpy' back payments.
- The Phase 1 evaluation survey indicates that only around a third of households are headed by a female.

In the case of categories for which the main sample of households may not be divided into two groups of equal size, the feasibility of a robust impact estimation will be assessed depending on the relative size of the relevant groups. At the very least, if the proportion and absolute number of households in the two groups do not allow for an IE due to sample size issues, a descriptive analysis of the differences in impact indicators between these sub-groups of households will be performed instead.

It should be noted that this analysis will be affected if the sub-groups are not equally balanced across the treatment and control groups, for example if the poorer households are more likely to be found among the treatment group compared to the control group. The RD methodology is still valid with an unbalanced sample (i.e. treatment and control groups of different sizes), though this will be less efficient (in a statistical sense) and potentially slightly more prone to bias due to model specification issues. However, where viable in such cases, the appropriate econometric adjustments will be applied to mitigate any potential issues.

2.3.4.5 Testing the viability of the RD approach

Internal validity tests were performed to determine whether the key assumptions underpinning the RD design are satisfied. 0 presents the results of these validity tests in detail, which taken overall suggest that our proposed RD approach is viable and formed the basis for our decision to go with this as our preferred methodology for the quantitative impact analysis.

However, it is important to highlight two key concerns emerging from the tests. Firstly, the estimates of the HSNP programme impact will be diluted by the fact that not all treatment households receive the transfer, while some of the households with scores above the PMT county cut-off point do receive the transfer due to the fact that beneficiary selection is also based on a combination of the PMT score and CBWR grouping. Secondly, the significant pre-intervention discontinuities found across a range of covariates highlight the risk that any difference in outcome indicators between treatment and control households may be due to pre-existing baseline differences rather than the effects of the HSNP CT.

Both of these concerns will be mitigated through the use of an instrumental variable (IV)[15] approach to the impact estimation. A fuzzy RD design is required when treatment status is not determined purely by the assignment score, and the probability of treatment becomes conditional on one or more covariates. In fact, a basic fuzzy RD design is econometrically equivalent to an IV approach, where the instrument used for the treatment status is the point of discontinuity in the probability of treatment. However, because a fuzzy RD design produces impact estimates that are econometrically equivalent to those produced by an IV approach, we will use a number of covariates in the IV specification to control for the key variables for which we are observing the largest and most significant pre-intervention discontinuities. In other words, our proposed IV regression specification will include the variables showing significant discontinuities to control for these discontinuities. Depending on whether these variables are endogenously or exogenously correlated to the treatment status, they will be included as instruments or covariates in the first stage of the IV specification. The resulting IV specification constructed for our analysis will thereby allow us to obtain robust impact estimates.[16]

[15] For a summary explanation of IV see the glossary.

[16] Although the IV approach does help to remove potential bias in the estimates of the programme impact, it is important to highlight that the selection of 'good instruments' could be challenging. IVs are required to have two key properties: they must be related to the explanatory variable(s), but must not be correlated with any unobservable factors (i.e. errors in the estimation) influencing the outcome variable. An informed and valid set of instruments will be selected once the sample for the analysis is drawn.

2.3.4.6 Risks around the RD approach

It is the opinion of the evaluation team that the RD approach set out in the preceding sub-sections offers the best option for delivering robust quantitative measures of programme impact in the context of complete programme roll-out and no possibility of any pre-intervention baseline data collection. However, we acknowledge there are some significant risks to the approach, which are set out below in this sub-section.

During the inception phase the option of entirely dropping the quantitative impact analysis in light of these risks was discussed. The decision essentially comes down to judgement as to whether the additional resources required for the quantitative impact analysis and extra survey sample size over and above that required for the LEWIE study are too high given these risks. The marginal cost of the quantitative impact study is provided in Section 0. Given the emphasis that DFID and programme staff have placed on the importance of the evaluation providing quantitative measures of programme impact for the design of national social protection policy going forward, it is the view of the evaluation team that the risks associated with the RD approach do not outweigh the cost, and therefore the quantitative impact study should remain a component of the evaluation.

It should also be noted that our design is informed by our team's first-hand involvement in the Uganda Social Assistance Grants for Empowerment (SAGE) evaluation (which OPM implemented) and the Tigray evaluation in Ethiopia (which was overseen from the client side by our TL). In both cases an RD approach was used for the quantitative impact analysis, and in both cases this was associated with problems.

In the case of the Uganda SAGE evaluation one of the key problems was that it turned out that the programme assignment variable was not systematically correlated with the key outcome indicators of interest. Ultimately this meant that the RD approach was dropped for the final analysis in favour of a PSM-based approach.[17] For this evaluation, however, the PMT component of the assignment process appears more systematically correlated with key outcome measures (based on our analysis of programme MIS data presented in Annex A). We thus believe that our proposed RD approach offers the best option for rigorous impact analysis. Furthermore, the Uganda SAGE experience –where the PSM approach that was eventually used was not part of the original evaluation design– has directly influenced our decision to build into our quantitative survey design the option for conducting a robust PSM-based impact analysis as a back-up (set out in detail in Annex A.2).

In relation to the Tigray evaluation, our understanding is that the problems were primarily an issue of inadequate sample power, which meant that measures of programme impact were not statistically significant. This has therefore informed our decision to implement just one single round of quantitative survey data collection with a larger sample size, instead of two survey rounds, each with a much smaller sample size (see Sub-section 0 below for further discussion). We present more detail on the sample size calculations conducted for the survey in Section 0 below.

Finally, and in acknowledgement of the risks associated with the RD approach, a PSM approach is proposed as a back-up methodology. This is set out in detail in Annex 0. By matching treatment and control households that are sufficiently similar on the basis of their observable characteristics, the PSM method attempts to mimic an experimental design by avoiding systematic differences between the two groups and therefore obtaining robust impact estimates. Although the external validity of the PSM approach is higher than that of the RD design, since the analysed households are not limited

[17] It should be noted that the Uganda SAGE evaluation team warned against the use of RDD from the outset. However, for a number of reasons beyond the evaluation team's control, the decision to adopt RDD as the basis for the study was taken by a multi-stakeholder group. The evaluation team then proposed a risk mitigation strategy (effectively increasing the sample size to generate a representative sample of the whole population – not required by RDD) in order to be better placed to implement a back-up methodology based on PSM should the RDD fail. However, the risk mitigation strategy was declined. In the event, the RDD did fail, but the evaluation team was still able to implement a PSM (albeit, in the absence of a sample fit for purpose, the ability to construct a viable matching model relied somewhat on chance and had limitations, in particular with regard to some individual-level indicators).

to the group of observations around the PMT cut-off, the procedure to reliably match treatments and controls can be challenging.

The rest of this section outlines the key risks associated with our proposed RD approach.

Risks around the fuzzy RD IV approach

As set out in the previous sub-section, our assignment variable (the PMT score) does not perfectly correspond to beneficiary status. This, combined with apparent pre-intervention discontinuities in a range of covariates, requires us to use a fuzzy RD IV approach that accounts for both this imperfect prediction of treatment status and any pre-intervention discontinuities. However, despite the mitigation potential provided by this IV approach, there exists a small possibility that the fuzzy RD estimator will be unable to generate satisfactorily robust and estimates if we are unable to identify suitable instruments, i.e. measures that are systematically related to the explanatory variable, but are not correlated with any unobservable factors influencing the outcome variable.

Risk that other social protection and emergency interventions may confound the impact analysis

The HSNP now operates alongside a number of other social protection and livelihood support interventions that were not present during Phase 1. The household survey that will be conducted to collect quantitative data for both the IE and LEWIE will also include questions on any other type of support received. These data can be cross-checked with the HSNP MIS data. This will enable us to determine how many different types of assistance the households in our sample are receiving, including emergency CTs among our control households. Also in this case, descriptive as well as regression analyses could be carried out to investigate potential correlations between household outcome levels and the receipt of financial assistance different from the HSNP transfers.

In the case of the HSNP emergency payments mentioned above, we plan to control for their presence by including receipt of these payments in the specification of the impact estimation model in the form of covariates.[18] As other interventions, including, for instance, the other NSNP programmes, have a less defined relation with the probability of treatment under HSNP and/or our outcome indicators of interest, the sensitivity analysis of their distribution on our selected sample of households will help determine the best way of controlling for their confounding influence. In any case, the fuzzy RD design itself assuages the problem in relation to NSNP interventions to a certain extent, as some households below the PMT threshold may receive these interventions, and some above may not, which reduces the disparity between our treatment and control groups in receiving the HSNP.

Risk that spill-over effects may confound the impact analysis

Spill-overs are a risk for any IE, including those based on an RCT design. However, they are likely to be more pronounced when the counterfactual (control group) is situated in the same communities as the treatment group, as is the case here. This is unavoidable in this case because Phase 2 covers all communities, so there is no scope for identifying non-programme communities to act as control areas. This implies that, should spill-overs prove a confounding factor, the impact estimates derived will be lower-bound estimates of impact (i.e. are likely to underestimate impact to some degree). On the other hand, the LEWIE study will provide an estimate of the extent and magnitude of any spill-over effects, and hence provide an insight into this issue.

2.3.4.7 Sampling considerations and external validity limitations

Our sample for the quantitative survey is structured in order to provide a basis for the multiple components of analysis under the IE: (i) the LEWIE; (ii) the RD impact analysis (including the scope for robust heterogeneity analysis); and (iii) the option of using a PSM-based IE approach as a back-up. This design setup was discussed extensively with DFID and

[18] Although control variables associated with emergency payments cannot be included as instruments, given their expected high correlation with our outcome indicators, they may be used as exogenous covariates, conditional on which treatment is predicted in the first stage and the outcome is explained in the second stage of the proposed IV estimation.

programme staff as part of the inception phase, and in particular the trade-off between the advantages and disadvantages of implementing two rounds of a longitudinal survey with a smaller sample size, versus the single large survey that is currently proposed.

Our power calculations are explained below in Section 2.4. We have determined our sub-group sample sizes in order to pick up a minimum detectable effect (MDE) of 8 percentage points. These calculations were informed by the results of the Phase 1 HSNP IE study (including MDE, intra-cluster correlation and non-response rates). They incorporate conservative estimates of design effects informed by other evaluations we have conducted using RD design.

Our sample frame is the programme's own MIS dataset, which provides pre-intervention registration information on all beneficiary and non-beneficiary households across the four programme counties. The MIS registration data also provide us with an effective baseline dataset, since they relate to key pre-intervention characteristics of beneficiary and non-beneficiary households that can be linked (by design) to our sampled households. (As mentioned in Section 0 above, the MIS data were used to scope the feasibility of our proposed RD design approach.)

The RD approach is focused on households in relatively close proximity to the eligibility threshold (in this case the county-level PMT cut-off). The estimate of the impact therefore represents a local average treatment effect (LATE). This means that while the RD approach has strong internal validity, in that it provides robust estimates of HSNP impact for the set of evaluation households on which it is implemented, it has weaker external validity, in terms of its applicability to households further away from the eligibility threshold. In other words, observations located at the two extremities of the PMT distribution are not included in our impact estimation. The width of the range of observations that should be used for an RD analysis in the vicinity of the cut-off point is determined by the choice of a bandwidth, which is a smoothing parameter intended to minimise the expected squared error of the estimation (i.e. trade-off between bias and variance). The rules on bandwidth restriction are less stringent for a fuzzy RD design compared to a sharp RD design, as the functional form of a fuzzy RD is less likely to be close to linear in the vicinity of the cut-off point. This means that external validity constraints are less binding for a fuzzy RD than a sharp RD.

The quantitative IE only requires households within the RD bandwidth to produce a robust impact estimate. However, there are a number of reasons that a broader sample – that is, a sample that is representative of the whole population, and not just the population that falls within the bandwidth around the eligibility cut-off necessary for the RD design – is desirable. These include: the LEWIE study, which requires a representative sample of the whole population; the back-up PSM methodology, which requires a broad sample across all household types to maximise the chances of constructing a viable matching model; and the inherent value of the dataset, which is vital for future policy research purposes. A fully representative sample will also be useful for providing descriptive statistics to help explain how relevant the LATE estimator is for the entire population of beneficiary households, as well as a source of information for programme specific micro-simulations (see Section 4 below).

For these reasons, and to **ensure value for money** in relation to the cost of constructing this dataset, we propose to sample some additional households outside of the required RD bandwidth in order to obtain a representative sample of all beneficiary and non-beneficiary households. The exact numbers of beneficiary and non-beneficiary households that we are proposing to sample for the IE and LEWIE are reported and explained in Section 0 on the sampling strategy.

2.3.4.8 Why linking to the Phase 1 evaluation sample is not appropriate

During the inception phase the evaluation teams was asked to consider ways of linking to the Phase 1 evaluation sample in order to assess the impact of Phase 2 households that also benefitted from Phase 1. Below we describe our consideration of this issue and why we ultimately felt it was not feasible.

Under the Phase 1 evaluation there were 1,440 households in the treatment group, and 1440 in the control group. In Phase 2 programme coverage is on average around 20%, so one should expect to find about 288 of our original treatment households still treated, and 288 of our original control households now treated. Because the Phase 1 evaluation was based on an RCT, theoretically one might suppose that we could thus compare these two groups, based

on new data from an additional standalone survey (i.e. by re-interviewing these 576 households). However, although making use of the original randomised design is appealing, this approach would not yield robust measures of Phase 2 programme impact. The reason for this is that selection into Phase 2 will be affected by beneficiary status in Phase 1: the Phase 1 IE showed that the HSNP transfers had a positive impact on household welfare, and we would therefore expect the treatment group to be less likely to be selected for the programme compared to the control group, all else being equal. In other words, there will be systematic differences between the Phase 1 treatment households selected under Phase 2 and the Phase 1 control households selected under phase 2, which would thereby introduce selection bias into any impact analysis based on comparing these two groups. Furthermore, in addition to this potential bias, the diminished sample size would imply that we could only detect very large differential marginal impacts (that is, the marginal difference between benefitting from both Phase 1 and Phase 2 of the HSNP and just benefitting from Phase 2).

We therefore do not feel that the added cost of any such standalone survey represents value for money, especially given that results from the first phase evaluation already indicated an accumulating effect of the programme over time, i.e. given the way the impact on consumption expenditure (including health expenditure) and poverty evolved and consolidated between the first and second rounds of the Phase 1 evaluation survey.

2.3.5 Output

The output from the quantitative impact analysis will be a report detailing the methodology and findings from the study and representing the impact estimates of the HSNP in terms of the various outcome indicators described above (Section 0). The report will also incorporate findings from the first two rounds of qualitative research (see Section 0 below).

2.4 QUANTITATIVE SURVEY DESIGN AND SAMPLING STRATEGY

We propose a single round of quantitative data collection based on a household survey that will underpin both the LEWIE and the quantitative IE. The survey, which will comprise a household and a business instrument, is discussed in Section 2.2.4 on the LEWIE.

The sampling procedure envisaged for this quantitative survey is intended to cover the different sample requirements of the IE approaches, including the LEWIE, the quantitative IE based on the RD approach, the PSM back-up, and the requirements of the simulation analysis to be completed as part of the policy analysis under Workstream 3 (see Section 4). Due to the data quality issues affecting the village categorisation (see footnote 14 above), we are proposing to use the sub-location (a sub-county administrative unit) as our primary sampling unit (PSU). Within each of the four counties, we will perform a first stratification based on sub-counties. Sub-locations will then be selected within each sub-county. An explicit stratification of sub-locations will be carried out with the aim of identifying sub-locations that can be defined as towns, nearby villages and remote areas. This will be guided by the settlement type classification that has already been made by the programme as part of its sub-location mapping exercise. As explained in Section 2.2 LEWIE above, this categorisation is required for the analysis of the impact of HSNP on the local economy. A pre-set number of households of different types will then be sampled from each of the three strata (see Table 4).

This strategy can thus be defined as a multi-stage cluster sample design, with explicit sub-cluster stratification and reweighting at the household level. The HSNP MIS will be used as a sample frame to select 198 PSUs (sub-locations) from across all 21 sub-counties in each of the four counties. The PSUs will be stratified as explained above, and a pre-set number of towns (defined as being sub-county centres), villages and remote areas will be selected within each stratum. Specifically, we will select one town, four villages and four remote areas from each sub-county. This categorisation adopts a LEWIE-specific terminology and reflects the programme's own classification of community types. The number of towns and related villages and remote areas is determined, on the one hand, by the total number of urban sub-locations (towns) in the 21 sub-counties and, on the other hand, by the associated location pattern required for the LEWIE analysis. The sampling technique that we envisage for the selection of these PSUs is probability proportional to size (PPS). The PPS selection method entails that the probability of selection will be directly proportional to the size of each PSU and this will ensure that every household in the target population has an equal chance of being selected.

Specifically, households will be randomly selected from each stratum, thus representing our secondary sampling unit. Given the household information contained in the MIS (names, geographic information, identity card numbers, etc.), our experience in the area and our understanding of the northern Kenyan context, we are confident that the identification of sampled households from the MIS will be feasible.

A different number of households will be sampled from the various strata in line with both the methodological requirements of the LEWIE and the expected dispersion of households across the three groups. As presented in Table 4, we propose to randomly sample 54 households (27 treatment and 27 control) in each town, 40 households (20 treatment and 20 control) in villages and 20 households (10 treatment and 10 control) in the remote areas, giving a total of 114 households per PSU. Given the distribution of towns, villages and remote areas in each sub-county, this gives a weighted average of 12 households per PSU in each treatment arm. The total sample size will be 6,468 households (3,234 treatment and 3,234 control households), as shown in Table 5.

TABLE 4 HOUSEHOLD SAMPLE SIZE PER PSU BY TYPE

BENEFICIARY STATUS	IE – LEWIE SAMPLE RANGE	TOWNS	VILLAGES	REMOTE AREAS
Non-beneficiary	Without IE (Treatment)	7	4	3
Non-beneficiary	Within IE (Treatment)	20	16	7
Beneficiary	Within IE (Control)	20	16	7
Beneficiary	Without IE (Control)	7	4	3
Total		54	40	20

TABLE 5 TOTAL SAMPLE SIZE BY HOUSEHOLD TYPE AND PSU TYPE

BENEFICIARY STATUS	IE – LEWIE SAMPLE RANGE	TOWNS	VILLAGES	REMOTE AREAS	TOTAL
Non-beneficiary	Without IE (Treatment)	154	352	264	770
Non-beneficiary	Within IE (Treatment)	440	1408	616	2,464
Beneficiary	Within IE (Control)	440	1408	616	2464
Beneficiary	Without IE (Control)	154	352	264	770
Total		1,188	3,520	1,760	6,468

In order to adjust for the sampling imbalance created by the explicit stratification of the sub-locations, at the analysis stage, households will be weighted depending on their probability of selection within towns, villages or remote areas. The weight will be based on the actual proportion of households living in towns, villages and remote areas within the total population of the sub-location. This will ensure that the sample is representative of the real distribution of the population across towns, villages and remote areas. The use of PPS for the selection of PSUs helps address the potential issue emerging from the fact that some of the sub-counties from which the sub-locations are selected may be disproportionately characterised by the presence of towns, villages or remote areas.

The RD design IE component will be based on a smaller sample of 4,928 households (2,464 treatment and 2,464 control households). This will provide our estimation model with enough power to confidently detect an 8% change in outcome

indicators of interest (0.16 standardised effect size – the effect size expressed as the mean difference between two groups in standard deviation units). On the basis of the results obtained with the IE of HSNP Phase 1, we are confident that this 8% level of change can be expected for primary impact indicators such as proportion of households in the bottom poverty decile, other poverty-specific indicators (e.g. poverty gap and severity of poverty) and some of the indicators pertaining to the ownership of assets. Although the selection of the sample for the quantitative IE will be guided by considerations pertaining to the optimal bandwidth of the RD design, as mentioned in Section 0 above, the proposed fuzzy RD design with IVs does not present the same bandwidth constraints of a sharp RD design approach. This allows us to ensure that the selected sample size is large enough to provide our estimation model with sufficient power to detect the HSNP programme impact.

TABLE 6 POWER CALCULATIONS AND SAMPLE SIZE

NUMBER OF PSUS	WEIGHED NUMBER OF HOUSEHOLDS	TREATMENT SAMPLE	CONTROL SAMPLE	TOTAL SAMPLE	DESIGN EFFECT (DEFF)	MDE	STANDARDISED EFFECT SIZE (SES)
198	12	2,464	2,464	4,928	4.26	8%	0.16
198	6	1,232	1,232	2,464	3.44	10%	0.21

Table 6 above presents the technical parameters that were used in the power calculations and demonstrates how the proposed total sample size for the IE of 4,928 households across 198 PSUs will allow us to confidently detect an increase of at least 8% in one of our outcome indicators of interest discussed above.[19]

The last row in Table 6 also shows that a sub-sample of half the households derived from the 4,928-household sample would still have enough power to detect a minimal programme impact of 10% (0.21 SES). This is important because it shows that we will be able to analyse heterogeneity of the HSNP impact across a number of relevant categories of households. For example, we can split the household sample in half, e.g. between large and small households or poor and less poor households, to compare the level of impact across those different dimensions.

However, as discussed in Section 0 above, the same type of heterogeneity analysis cannot be performed across sub-samples that are not close to equal in size, and do not therefore represent half of the total IE sample. This is because the smaller of the two unbalanced sub-samples would not be sufficiently large to provide the estimation model with enough power to confidently detect the impact of the programme. Once our sample of households is selected, we will determine whether some unbalanced sub-groups of households are however similar enough in size to allow for a heterogeneity analysis.

Finally, despite the precise figures presented in Table 5 (upon which our power calculations were based), we have budgeted for a sample size of 6,600 households overall. This will help us deal with potential sampling issues (e.g. rate of non-response[20]) and will provide a slightly larger pool of households for matching treatment and control observations in case we have to adopt the back-up PSM approach.

[19] These calculations assume a power of 80% and are underpinned by a set of technical parameters that include a deff of 4.26 (i.e. RDD design effect of 2.75 and intra-cluster correlation of 0.05, which represents the average degree of homogeneity across households within the towns, villages and remote areas where the sampled households are located). As this is a single cross-section in a post-treatment setting, no inter-temporal correlation is included.

[20] We assume a non-response rate of up to 3%, based on previous experience from the HSNP Phase 1 evaluation.

2.5 QUALITATIVE ANALYSIS

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DATA ANALYSIS METHODS	DAC CRITERIA	EVALUABILITY ISSUES
a. What are the overall effects of the CTs on household and individual welfare, comparing beneficiaries and non-beneficiaries?	What are the effects on subjective wellbeing, social networks, and conflict/social tension?	Perception of wellbeing among household and community members; different dimensions of wellbeing in the community; influence of social norms on control of resources and decision making; factors affecting levels of social cohesion within the community	HSNP beneficiaries and non-beneficiaries, with information collected from different sub-groups of the population (e.g. youth, pastoralists, female household heads, ethnic minorities, local leaders, etc.)	FGDs, KIIs and observation, with the use of a range of participatory tools including participatory photography	Interviews will be voice recorded and subsequently transcribed for coding and analysis software; daily analyses during fieldwork and a summary analysis at the end of the fieldwork will be performed.	Impact	Medium to high. The ability of the qualitative study to achieve its aims is related to availability and cooperation of the study's respondents.
b. For which sub-groups are effects most pronounced?	Are vulnerable groups less or more affected?	Main characteristics of the different wellbeing categories; distribution of wellbeing within the households and in the communities; distribution of changes in wellbeing across categories	As above	As above	As above	Impact	Medium to high. Qualitative analysis will focus on dynamics affecting sub-groups, including vulnerable groups.
c. Do CTs impact on women's control of cash within their (often polygamous) households and their wider empowerment?	Are gender relations affected by the HSNP?	Women's wellbeing conditions; economic power within the household and the community; changing livelihoods; resilience to shocks; access to financial services	HSNP female beneficiaries and non-beneficiaries, including different sub-groups such as female household heads and female youth	FFGDs and one-to-one interviews, with the use of a range of participatory tools	Interviews will be voice recorded and subsequently transcribed for coding and analysis software; daily analyses during fieldwork and a summary analysis at the end of the fieldwork will be performed.	Impact	Medium to high. This evaluation question will be answered with qualitative information, depending on women's willingness to be interviewed.
d. How do the larger one-off transfers some households will receive due to the later than	Do beneficiaries report differences in impact over transfer size	Reported expenditures of HSNP transfer; Reported impacts	HSNP beneficiaries from different sub-groups in terms of early and late joiners	FFGDs and one-to-one interviews, with the use of a range of	Interviews will be voice recorded and subsequently transcribed for coding and	Impact	As above, but also depends on the extent to which we capture sufficient

anticipated start of the programme impact on those households?	and frequency?		to the programme	participatory tools	analysis software; daily analyses during fieldwork and a summary analysis at the end of the fieldwork will be performed.		variation in transfer start date in our sample.
e. Does the combination of CTs and wider livelihoods activities open up new livelihoods opportunities/ income-generating activities for poor households?	Through which channels are new opportunities generated and how do they develop?	Main livelihood activities undertaken within the community; livelihood activities associated with or done by different social groups; new/changing sources of livelihoods	HSNP beneficiaries and non-beneficiaries including those benefitting from other livelihood interventions	FFGDs and one-to-one interviews, with the use of a range of participatory tools	Interviews will be voice recorded and subsequently transcribed for coding and analysis software; daily analyses during fieldwork and a summary analysis at the end of the fieldwork will be performed.	Impact	Medium to high. This will be analysed qualitatively as far as possible depending on the prevalence of other livelihood support interventions and their overlap with HSNP.
g. Is there evidence of the HSNP having an impact on community relations?	Is the impact observable both within and between communities ?	Forms and sources of disputes and tension between and within communities	Household and key informant community members (i.e. teachers, religious and local leaders)	FFGDs and KIIs	Recorded and transcribed interviews; daily fieldwork analysis and final summary analysis.	Impact	High. Qualitative analysis of communities is feasible.
k. Do the new payment platform and expansion of financial services provide benefits for beneficiaries and non-beneficiaries?	Are beneficiaries able to access formal financial services as a result of the programme? What are the benefits of this? How does this compare to non-beneficiaries ?	Access to formal financial services; usage of formal financial services; benefits of formal financial services	As above	As above	As above	Impact	Medium to high. The impact of increased access and usage of formal financial services may be hard to capture accurately.
s. Do the reliable CTs and emergency payments build people's resilience to climate variability?	How can beneficiaries build resilience to climate variability? How is this affected by the HSNP? How does this compare for non-beneficiaries ?	Main climate variability response mechanisms undertaken	As above	As above	As above	Impact	Medium to high. The impact on increased resilience to climate variability may be hard to capture accurately.

2.5.1 Rationale

The qualitative research provides an assessment of the impact of both the routine HSNP and emergency payments. It also incorporates an understanding of the interaction between the HSNP and other livelihoods support programmes such as the ASP.

Taken together, these two objectives of the qualitative research seek to complement the quantitative research to provide a more complete and nuanced assessment of programme impact. The qualitative research will build on the contextual knowledge gained about the four counties from the HSNP Phase 1 evaluation and provide new contextual information where relevant. This will help to avoid duplication and enable the qualitative research to be more responsive and focus on areas where in-depth research is required in order either to better understand the quantitative data, or to generate data on areas not covered by the quantitative analysis. In particular, it will fulfil the following four objectives:

1. It will provide an **understanding of the context** within which the programme is operating, and how this affects and is affected by the CT programme.
2. It will help to **understand experiences and processes that produce outcomes of interest** in the evaluation. This is complementary to quantitative research, which will provide a measure of the outcomes and impacts.
3. It will enable an **assessment of impacts that are (methodologically) difficult to cover** completely and sensitively via the quantitative survey (such as social cohesion and inter- and intra-household relations) and provide nuanced data that explains findings where useful.
4. It will provide complementary data on some of the topics covered by the household surveys, thereby **triangulating, validating and providing depth** to the quantitative findings.

2.5.2 Timeframe

The qualitative research will take place over three rounds. Each round will be preceded by a design phase in which the methodology and research plan are refined with respect to the goals of that particular round and the tools and fieldwork protocols are developed. The fieldwork will take place in all four counties simultaneously over a two- to three-week period. Within each county the research will be undertaken in three treatment sub-locations in the evaluation areas. The three rounds of research will take place in August 2015, September 2016, and June 2017. This scheduling will enable an unpacking of both short-term and longer-term impacts. It also facilitates triangulation between the quantitative and the qualitative analysis. After fieldwork at each round the qualitative data will be processed and analysed over the following two-month period with the aim of producing a report for each round of research in the following month: October 2015, November 2016 and August 2017. Note that findings on the impact of the emergency payments from the first round of the qualitative results will be analysed as a special priority in order to be reported in early September. At the end of each round of fieldwork, a short fieldwork implementation note will be produced, summarising some preliminary findings.

The timing of the qualitative and quantitative data collection activities are sequenced in such a way as to inform each other's design and maximise synergy during analysis (see Section 2.6). The first round of qualitative research takes place during the design phase of the quantitative survey, such that preliminary findings can feed into the design of the quantitative study. The second phase of the qualitative research takes place after the results from the quantitative study have been produced, enabling us to investigate the findings that emerge from the LEWIE and quantitative IE. The final round of the qualitative research focuses on longer-term impacts, which are not covered by the quantitative study (due to the absence of a second longitudinal survey).

2.5.3 Research questions

We have regrouped the impact-related research questions specified in the ToR into four broad areas that the qualitative research will focus on. These are:

- perceptions of wellbeing at individual, household and community levels;
- vulnerability and resilience;
- livelihoods and local markets; and
- informal institutions and social relations (inter- and intra-household relations, gender relations, social cohesion).

Table 7 maps the more detailed research questions against these four research areas.

TABLE 7 MATRIX OF KEY QUALITATIVE RESEARCH AREAS AND QUESTIONS

RESEARCH AREA	KEY RESEARCH QUESTIONS
Dimensions and perceptions of wellbeing (levels or categories, distribution and trends of wellbeing)	How do community members perceive wellbeing? What are the different dimensions of wellbeing in the community? What different wellbeing categories exist within different communities? How have these categories changed over time? What are the main characteristics of the different wellbeing categories (e.g. social characteristics, assets, coping strategies, power and influence, etc.)? How are households in the community distributed among these categories? How does this distribution change over time? What is the distribution of wellbeing *within* households (e.g. between old and young, male and female, etc.)? How has this changed over time? How has the HSNP affected wellbeing among different social groups?
Vulnerability and resilience	What are main shocks faced by individuals, households and different social groups? How do shocks affect different individuals, households and social groups? Have the main shocks faced by households changed over time? How and why? How are these shocks categorised (e.g. long-term trends, seasonal)? What determines different levels of vulnerability and resilience to these shocks? What effects do these shocks have if they occur? What strategies are adopted to reduce, mitigate and/or cope with vulnerability to and the effects of these? How do these strategies differ within and between different households and social groups? How have these changed over time? How do the HSNP and emergency payments affect the ability to reduce, mitigate and cope with different stresses and shocks at the individual, household and community levels? How do they help to build households resilience?
Livelihoods and local markets	What are the main livelihood activities undertaken within the community? What livelihood activities are associated with or done by different social groups? Why is this? How and why have these changed in recent years? How and why do people move between different livelihood activities? What are the preferred sources of livelihoods and why? What are the constraints and challenges to participating in these forms of livelihoods? How do participation and forms of livelihood activities vary within households? How do people save and/or invest in assets? How has this changed over time? Why? How and why have asset levels changed over time? What is the level of trade and business activity in the community? How and why has this changed in recent years? What livelihood support programmes are available in the community? How do these affect the wellbeing of community members? How has the HSNP CT affected livelihood choices and options?

	How have the HSNP and emergency payments affected the local markets for goods, services and labour markets? What are the interactions among the HSNP, emergency payments and other social assistance or livelihood support programmes (e.g. ASP)? Has the HSNP affected access to these programmes? What are the impacts of this on households?
Social relations	How are households defined within the community? What social structures and institutions exist at community level? What influence do social norms based on gender, age, ethnicity, etc. have on an individuals' and households' capacities and entitlements? How do social norms affect control over resources and decision making? What factors affect levels of social cohesion within the community? What are the forms and sources of disputes and tension between and within households? How have the HSNP and emergency payments affected, or been affected by, informal institutions, social relations and cohesion? Why? What are the effects of these changes?

2.5.4 Method

The qualitative research is designed to move beyond the traditional way of undertaking qualitative IEs for CT programmes in four ways. First, we propose to assess impacts and outcomes of the HSNP through a broader wellbeing framework, in which development interventions have not just material, but also relational and subjective, effects (White and Ellison, 2006). Within this framework, recipients and research participants are involved in the articulation of how they define wellbeing, including 'psychosocial' wellbeing. This framework will enable research participants to assert their agency and voice in the research process, allowing them to articulate change in their own terms, and raise underlying or less obvious issues that researchers may otherwise not have anticipated.

Secondly, the qualitative research will focus on less tangible or quantifiable dimensions of interest. It will seek to uncover the complex range of interrelated social processes that can be either positively or negatively affected by the introduction of CTs.

Thirdly, the qualitative research will be designed to understand the reciprocal interaction between programme processes and impacts. In other words, it will seek to unpack how HSNP operational processes, such as targeting and grievances, or regular and full receipt of payments, mediate programme outcomes and social processes. Conversely, it will seek to comprehend how social processes might impact on and/or explain the effectiveness of programme operations.

Finally, the qualitative research is sequenced in such a way as to both inform the design of the quantitative studies (LEWIE and IE) and to respond to findings and issues raised from previous rounds of both the quantitative survey and the qualitative research itself

and capture longer-term impacts. This will enable the evaluation to drill down into complex or unexplained results from the research, as well as respond flexibly to new research questions that might emerge over the evaluation period.

2.5.4.1 Research methods

For the qualitative research we will use a combination of panel studies, FGDs, KIIs and observation.

Qualitative household panel studies are a form of data collection in which the same households are interviewed using semi-structured instruments over multiple rounds of research. This gives an in-depth understanding of household behaviours and dynamics in the face of the contexts in which they are embedded. In this case, the qualitative research will be conducted over three rounds of fieldwork, spread evenly across an 18-month period.

Using qualitative household panel studies is innovative within CT evaluations and will offer several advantages. It will offer a temporal dimension of change, as well as an understanding of causality, how and why change occurred, and how

aspects of social, cultural and contextual processes interact to produce different individual outcomes. It will also allow for flexibility in the research design. For example, the repeat household visits will allow for an iterative and continuous evolution of theorisation and research questions.

The household panel studies will be complemented with **FGDs and KIIs** to explore specific issues identified through the research, as well as any emerging from the panel studies that need group level validation. We will explore the use of **participatory tools** during the discussions such as **social mapping**, **wellbeing ranking**, **household income and expenditure analysis**, and **livelihood ranking**. The final choice of tools will be made at the design phase, but will be linked to key research areas that have been identified in Table 7. Box 2 elaborates on a list of participatory tools that could be used.

We will also incorporate the use of **structured observations and informal conversations** at community level. For example, it should be possible to time the fieldwork in order to observe community dynamics on payment days. Similarly, engaging in informal conversations will provide spontaneous information and generate data that community members may not divulge in a more formal discussion context.

Finally, we will also explore the use of **participatory photography** where beneficiaries will be supported to capture **photographic images** that reflect their experiences and lived realities as HSNP beneficiaries. These images can then be used to inform and structure the respondents' discussions during fieldwork.

Box 2 Examples of participatory tools

Vulnerability mapping aims to identify different types and levels of vulnerability within different social groups. Important issues include how people's vulnerability to poverty and social exclusion is manifested; whether it is changing over time; and if so, how and why. While no one tool will provide the answer, taking time to look at vulnerability in terms of what it means to people in different interest groups – who they think are the most vulnerable in their area; how this has changed over time; and what they think are the causes of vulnerability – helps build up a picture of how effectively social assistance grants and other mechanisms are working to reduce people's vulnerability. A vulnerability map is a visual tool used to initiate these types of discussions with different social or interest groups. Capturing responses from a range of interest groups enables an analysis of commonalities and differences, as well as trends in risk and vulnerability.

Institutional mapping is a visual method of identifying and representing perceptions of key formal and informal institutions and individuals within and outside a community. It can focus, for instance, on the availability of formal and informal social protection systems and perceptions of their effectiveness. Institutional mapping produces a picture of each institution's perceived relevance and importance to the group. This enables an understanding of how different community members or social groups perceive and relate to the institutions that can influence or affect their level of vulnerability to different shocks, trends or seasonal variations.

Ranking/scoring can be used to identify and explore participants' preferences for different social protection mechanisms, livelihoods and coping strategies, among other issues, and articulate why they prefer them over others. Building on the institutional mapping, for example, a ranking exercise could be undertaken in order to understand the preferences of different social groups for different systems of social protection and what makes one more effective than another. Alternatively, ranking can be used to explore which types of risk and shock are most serious for different groups and why. As well as qualitative data, this produces contextual quantitative data in a participatory manner.

2.5.4.2 Sampling

Selection of evaluation areas

Within each county, the qualitative research will take place in three treatment sub-locations. These sub-locations will be chosen purposively. This is because, in contrast to quantitative research, qualitative research aims to substitute breadth of coverage and statistical generalisability with the use of contextual methods to explore complex issues in depth. The sub-location sample frame for the qualitative research will be determined by the programme MIS as that covers all sub-locations in four programme counties.

The exact sub-locations to be selected will be determined in the qualitative research design phase in collaboration with the key programme stakeholders. **Sub-locations will be purposively selected** to cover a range of contexts, including geographical (e.g. urban and rural), socioeconomic (livelihood, poverty level, etc.) and programme-related issues (e.g. remote). This selection will therefore be designed to include sub-location characteristics that are expected to have differentiated responses to the HSNP transfers. In addition, for looking at the emergency payment, we will work with the NDMA to select sub-locations that have a high susceptibility of facing future droughts, so that the assessment is able to capture the impact of more than just one round of payment.

Selection of respondents

The sampling framework we propose will ensure that we capture some level of social and economic differences and diversity within the selected communities, and of household types, while at the same time enabling a level of consistency in groups across counties.

The **households for the qualitative panel survey** will be selected using random stratified sampling. For example, after beneficiary status, households will be further stratified according to pertinent characteristics, based on an analysis of the HSNP MIS data. These could include household size, welfare status (based on the programme PMT), recipient status under HSNP Phase 1, and gender of household head. The criteria for stratification will be refined during the design phase.

A total of 10 households (six routine beneficiaries, two emergency beneficiaries and two non-beneficiaries[21]) will be selected in each sub-location for the panel studies. This means that a total of **120 households across the four counties** will be tracked over the duration of the evaluation.[22] This sample provides a decent basis for making inferences from the data, including enough variety at the county level and enabling us to triangulate with information gleaned from other sources such as the FGDs and KIIs.

Households that are sampled for the panel study will not be part of the quantitative interviews (i.e. no household will be in both the quantitative and qualitative surveys).

Sampling for the FGDs will be undertaken in the field, but will also be purposive and stratified, in that we will target particular population groups consistently across all sub-locations (male and female beneficiaries [including recipients of both routine HSNP payments and emergency payments], male and female non-beneficiaries, and female-headed households), as well as ensuring that other social or economic groups particularly relevant in specific sub-locations are captured (e.g. young women, female casual workers, female traders, ethnic minorities, child-headed households, etc.). These specific groups will be identified and targeted by the research during the early stages of fieldwork as the research teams are introducing and orienting themselves in the study locations.

KIIs will be undertaken with relevant people in communities, service-providing institutions (e.g. schools, health centres), national government and civil society. In some situations, it may be more useful to conduct a group discussion with several key informants (e.g. a group of teachers from a school) rather than an individual interview, and this will be assessed during fieldwork. Key informants have different positions and perspectives and bring their own sets of interpretive biases to their analysis of the impact of HSNP. In the first round of fieldwork we will aim to hold discussions with:

- sub-location chief;
- elder(s)/community leader(s);
- local trader(s);
- teacher(s)/health worker(s);
- religious leader(s); and
- local NGO worker(s).

In the follow-up rounds, as with the FGDs, discussions with key informants will be determined by findings emerging from the panel studies and other research activities.

[21] These non-beneficiaries will not be beneficiaries of the routine payments or emergency payments. However, they will be selected to be just above the PMT threshold.

[22] In the baseline we will interview an additional four households (two beneficiaries and two non-beneficiaries) as a reserve list in each sub-location.

We will incorporate observations to help make 'thick description' of particular themes. Researchers will spend time at pay points and in markets to observe community interactions and the nature of economic activity following payments. In each round of the research, an attempt will be made to observe a payment day in at least one sub-location in each county. For this purpose paypoints will be selected purposefully based on both community characteristics and logistical circumstances.

2.5.4.3 Recording the data

All interviews will be voice recorded and subsequently transcribed for coding and analysis (see below). However, researchers will also take comprehensive field notes. The research team will record the discussions among the participants as they speak, using the words they use and noting occasions when participants disagree or when one participant's opinion is particularly strong. Where possible, they will include any thoughts on why differences are emerging (often a reflection of the personal experiences, aspirations and world views of the different participants).

The research team will also record why the group came to a decision, answer or agreement. Where disagreements within the group occur, these should also be recorded accurately. If particular individuals try to dominate the discussion, it will be explicitly reflected in the notes so that their views do not distort the representation of the discussion and the views expressed. Finally, the team will accurately capture recording of the diagrams produced by participants, by digital photograph if appropriate.

2.5.4.4 Analysis of the data

Analysis of the data will start in the field. Researchers will be trained to confer with each other on the highlights for each research area and major points and issues raised during the discussion. Such discussion will form the basis of the daily team debrief.

At the end of each day, the team will meet to debrief and discuss key issues arising from the research that day. This will be done in two parts:

Firstly, about 20–30 minutes will be taken to ask each other the following questions:

- What went well and why?
- What did not work so well and why?
- What can we do differently tomorrow?

Secondly, using cards and flip chart papers, the team will record the key issues as they emerged from the interviews and FGDs. The team will use pieces of the flip chart for each of the components of the agreed analytical framework.

On cards, team members will write up the key issues that they identified during the day, one issue per card. They will record where the issue came from – e.g. from one interview or FGD, two or several, and which ones they were.

If more than one household raised the issue, for example, the team will be able to start identifying commonalities, whereas if only one household or group raised the issue, there may be differences between different groups which can be identified. Once all the issues have been written on cards, the cards will then be ordered under the different components of the analytical framework and in terms of commonalities and differences. Quotes and stories will also be written on cards and placed under the themes next to the issue that they are helping to illustrate.

While it is very likely that some of the issues emerging each day will be the same, it is important to continue to record on the cards which groups were raising the same issue. There may be some issues emerging that do not fit easily into the analytical framework. These will be recorded on a separate flip chart and documented.

This process of daily debriefing and analysis will enable the quick and timely summation of data, and increase the ability of the research teams to respond to new information or, if necessary, areas where insufficient data are being collected (e.g. because the questions are not relevant to a particular location or group, or because participants do not feel able to

respond to questions in a certain way). This will help teams adapt and improve the research process where needed as they go along.

At the end of the research in each county, time will be taken to go through all the data and analysis undertaken to produce a short report that analyses the data recorded during the discussions. This will again be in line with the components in the analytical framework. The report for each community should also include as much contextual information as possible about the community.

All voice recoded discussions will be transcribed into English and analysed using NVIVO.

2.5.4.5 Ethical considerations in carrying out research with vulnerable groups

We adopt an ethical approach to the research. There are a number of ethical issues to consider in planning and facilitating the participation of people in research, including:

- How are participants being selected? Is there any deliberate exclusion on the basis of, for example, access or stigma? Have cultural and community norms been understood and considered in the selection process? People should be offered the opportunity to participate and, ideally, be invited to volunteer rather than be asked, to avoid any pressure. People have the right not to participate and to pull out at any stage.
- Ensuring that permission is sought for the research to go ahead, through consultation with both the direct participants and local community officials.
- Setting and communicating clear parameters for the research – this means clearly stating the purpose, the limits and what the follow-up will entail. It also means ensuring that demands on participants' time are not excessive and that they are aware of their right to not participate or withdraw at any time.
- Recognising that participants are possibly vulnerable and that the exercise is carried out with full respect. Power differentials will exist between community members and researchers: for example, people can easily be treated as inferior and such power dynamics need to be understood and purposefully mitigated in planning and implementation by researchers and facilitators.
- Ensuring the safety and protection of participants – this means ensuring the environment is physically safe, that there are at least two facilitators present at all times and, if possible, that a local stakeholder group is involved in monitoring activities. Facilitators should also be supervised.
- Ensuring that people understand what is happening at all time, for instance that appropriate language is used (language, dialect, community terminology, etc.).
- Ensuring the right to privacy – this includes ensuring anonymity and confidentiality in record keeping and report writing, and making sure participants understand that what they do and say in the group session will remain anonymous.
- Ensuring that plans are in place for dealing with unexpected or adverse consequences, such as an urgent physical protection issue. Provision of immediate support needs to be planned for in advance.

2.5.5 Output

The output from the qualitative research will be a series of three qualitative research reports, one from each round of data collection, which will provide a comprehensive presentation and analysis of the research findings and methodology. In addition, findings from the qualitative research will feed into the outputs from the LEWIE and quantitative impact analyses (see Sections 0 and 0 above).

2.6 MIXING QUALITATIVE AND QUANTITATIVE EVALUATION APPROACHES

The analysis of quantitative and qualitative data will be combined and integrated to enrich the overall results of our evaluation. While the quantitative analysis will provide us with an aggregated measure of the average programme impact on our treatment group of households for a range of key welfare indicators, the qualitative analysis will shed light

on the wellbeing of different groups within the household and the wider community, with the particular aim of being responsive to emerging findings and explaining causal pathways of impacts.

The mixed-methods approach will be centred on three integrated analytical levels of the core impact areas:

1. The analytical concept of wellbeing: while the quantitative analysis will make use of a set of outcome indicators designed and developed to reflect the level of individual and household wellbeing (e.g. food security, nutrition, and consumption or ownership of assets, livestock and land), the qualitative analysis will gather information from the interviewees on what they themselves consider the most central aspects defining their general wellbeing, including psychological wellbeing. This will help us understand how to interpret the findings associated with quantitative indicators, including the degree of significance to assign to the different indicators.

2. Household-level analysis: As discussed in Sections 0 and 0, the quantitative evaluation will include a heterogeneity analysis of the impact on particular sub-groups of households (noting that due to sample size issues, only a descriptive statistics analysis will be feasible for unbalanced sub-groups of households). Qualitative analysis can assist in validating and better interpreting the findings emerging from this analysis. Such analysis can provide us, for instance, with information on gender dynamics within the household, in an attempt to assess how the HSNP transfers have specifically benefitted women and their level of empowerment. Other household dynamics, including across age groups for example, will be investigated by comparing relevant descriptive statistics with these qualitative insights.

3. Community-level analysis: The quantitative household survey will collect data on a range of variables concerned with the characteristics of the communities and local markets to which the respondents belong. This will be used for the construction of our impact estimation models, descriptive analysis and LEWIE. Qualitative data will also focus on this analytical level, by looking at the social effects produced by the CTs. Potential conflicts stemming from the division between HSNP beneficiaries and non-beneficiaries or, on the contrary, solidarity groups and other forms of cooperation between the two categories will help us understand how the programme is shaping interpersonal and social relationships in the towns and villages where HSNP households live and work.

The integration of quantitative and qualitative analysis will be based on a dynamic process that will evolve during the course of the evaluation to greatly increase our ability to assess the overall impact of HSNP. Two subsequent rounds of qualitative research will first follow up on immediate quantitative findings emerging from the quantitative survey and then return later to look at longer-term impact (i.e. through our qualitative panel studies).

In order to ensure that our quantitative and qualitative teams are aware of each other's findings and interact in the interpretation of the results, we are planning to set up a cross-analysis of the quantitative and qualitative draft and final reports. Quantitative researchers will provide feedback on qualitative sections of the evaluation that could have an influence on the interpretation of quantitative findings, while qualitative researchers will read the more quantitative sections to ensure that the insights emerging from the qualitative analysis are taken on board. This procedure ensures a properly integrated mixed-methods evaluation approach.

2.7 STUDY ON THE ASP

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DATA ANALYSIS METHODS	DAC CRITERIA	EVALUABILITY ISSUES
e. Does the combination of CTs and wider livelihoods activities open up new livelihoods opportunities / income-generating activities for poor households? How?	What are the enabling and constraining factors to undertaking range of livelihood activities? Are there other livelihood support programmes? What is their impact on households' wellbeing and resilience? Are these impacts differentiated at household/community levels? How do these compare with the benefits of the CTs? How do households respond to multiple interventions? How do they choose between them?	Livelihood activities of ASP; livelihood activities of HSNP CT; benefits of combined ASP and HSNP CT	FGDs with beneficiaries and non-beneficiaries; quantitative data from the LEWIE survey	Qualitative FGDs; quantitative household survey	Qualitative analysis of FGD data triangulated with descriptive statistics from the quantitative survey	Impact	Medium: Combination of respondents' testimony and potential descriptive statistics should inform the interaction between ASP and HSNP CT in relation to livelihood opportunities and benefits.
m. How effective is the work with NGOs to understand how communities respond to shocks under ASP? Does this get reflected in county plans?	What is the role of NGOs in the ASP? What are their challenges and constraints? How do NGOs coordinate with county officials on resilience-building activities? To what extent are their efforts reflected in county plans and resource allocation processes? What	Contribution of the ASP to county plans and budgeting processes	County officers and NGOs	KIIs	Data will be analysed to establish the contribution of the ASP	Effectiveness	Medium. A selection of interviewees will be able to express opinions about how far the ASP has influenced county planning and budgeting processes, which can then be triangulated.

lessons can be learnt in terms of NGOs engaging in resilience-building activities?

2.7.1 Rationale

In some areas, the HSNP is complemented by the ASP, which aims to build longer-term resilience to drought and shocks by enhancing community adaptation strategies, livestock insurance, livelihood diversification and emergency drought response. To enable an assessment of the ASP, we propose a special study for the ASP. This will be complemented by findings from the qualitative IE work stream, which looks broadly at the interaction between HSNP and other livelihood support programmes (potentially including the ASP[23]).

The results from this study will help assess whether the components and activities of the ASP are leading toward or are contributing to the programme outcomes of the GoK, community and non-state actors working together to create resilient livelihoods and adaptive capacity in households and communities, and, if so, how. The study will assess the degree to which the ASP contributes to county-level planning and budgeting processes. It will also provide some evidence on the extent to which the combination of cash transfers and other livelihood support programmes enhances longer-term resilience among households.

2.7.2 Timeframe

Detailed design of the ASP study will be undertaken in December 2015. Data collection will take place in January 2016. Data analysis will take place in February 2015. An individual report for the ASP will be submitted in March 2016.

2.7.3 Research questions

In addition to the research questions outlined in Section 0 above, a number of research questions aimed more specifically at ASP beneficiaries and programme implementers will be answered:

- What are the merits of combining the CT and other livelihood support programmes?
  - What are the enabling and constraining factors to undertaking the range of livelihood activities identified?
  - What activities would households invest in that they currently do not have funds for to enhance their resilience?
  - What other livelihood support programmes exist? What is their impact on households' wellbeing and resilience? Are these impacts differentiated at household and community levels?
  - How do these compare with the benefits of the CTs? What is the impact on household vulnerability and resilience?
  - How do the benefits of combining programmes that support livelihoods compare to just CT?
  - How do households respond to these multiple interventions? How do they choose between CT and other interventions?
  - To what extent, if at all, do the businesses stimulated by ASP funding result in a loss of household mobility?
- What is the role of NGOs in building community resilience?
  - What is the role of NGOs in the ASP programme?

[23] This depends on the degree to which HSNP beneficiaries overlap with ASP beneficiaries. Due to the minimal overlap of the ASP with the much larger HSNP coverage area, as well as the methodological challenge of constructing a viable counterfactual for ASP beneficiaries and the implications on sample size, it is not possible to include a rigorous quantitative assessment of impact for the ASP.

- What are the challenges and constraints in their role? How do NGOs coordinate with county officials on resilience-building activities?

- **Is this combined approach to livelihood support appropriate in the four counties?**
  - What are the sources of risks and vulnerability?
  - How do community members reduce, mitigate and cope with different stresses and shocks?

- **To what extent is there a consensus on the path toward resilience and the place of livelihood-enhancing actions in the county development plans?**
  - What do the county development strategies say about the pathway toward resilience, and what are the views of county councillors, civil servants and NGO ASP implementers on the country strategy? What else should be in the county plans?
  - To what extent are activities funded by ASP likely to be central to a path toward universal resilience in each of the four counties?

- **What are the views of county development plan managers, fund allocators and NGO officials on the value for money of ASP activities?**
  - What proportion of county-level budgets is allocated to ASP-funded activities in the 2013/14 financial year?
  - What can be gleaned from the country records regarding expenditure trends on these activities?
  - What are the views of country-level officials regarding trends in such allocations over the next five years?
  - What are the views of NGO implementers on the trend in county allocations toward the actions of ASP funds? What else could be done, if appropriate, to convince county governments to allocate more resources to ASP-like activities?

2.7.4 Method

The assessment of the ASP will focus on two aspects: the degree to which the ASP contributes to county-level planning and budgeting processes and the extent to which the combination of CTs and other livelihood support programmes enhances resilience among households. These impacts will be assessed via the analysis of data collected through the following means:

- **FGDs:** The experience of change by a number of programme beneficiaries as a result of receiving support from the ASP will be explored in FGDs with ASP beneficiaries. Where possible, FGDs will include the views of those who also benefit from the HSNP. Up to four FGDs and four household case studies will he held with ASP beneficiaries in sampled communities (depending on the variety and numbers of different projects run by the NGOs). These interviews will be supplemented with the data from the qualitative household panel studies where available.
- **Document review:** county government development plans, budgets and expenditures will be reviewed (assuming availability) to track trends in planned and actual expenditures on activities also supported by ASP.
- **KIIs:** KIIs with county officials and NGO ASP implementers will be conducted, focusing on:
  - county officials' sentiments regarding trends in ASP-like investments using GoK funds;
  - county officials' satisfaction with the quality of county development strategies, in respect of the extent to which the strategies will succeed in increasing the proportion of households considered resilient; and
  - NGO ASP implementers' experience of supporting community resilience programmes, eliciting their views on trends in the focus on broad resilience programming in county development plans and the execution of those plans.[24]

[24] These latter interviews may take place in the course of quarterly review meetings when ASP NGO implementers gather together, rather than in the field at the same time as the other KIIs and FGDs.

Roger Pearson will provide technical leadership on the design and analysis of the study. Ramlatu Attah, the IE qualitative research component lead researcher, will provide oversight in the implementation of the fieldwork.

2.7.5 Output

The output from the special study on ASP will be a standalone report detailing the results and methodology of the study. The report will produced in March 2016 and will incorporate the analysis of the interaction of the HSNP with other livelihood support programmes from the first two rounds of the qualitative research component from the IE.

3. Workstream 2: Operational M&E

3.1 OBJECTIVES OF THE WORKSTREAM

The operational M&E workstream will assess how well the HSNP is being managed and implemented by its various implementing partners, highlight best practice, offer recommendations for improvement, and consider the implications of operational processes for the programme's overall efficiency and value for money. The outputs of this workstream take two forms:

1. **One-off, standalone studies on particular topics of operational interest.** These may be retrospective, providing accountability to funders on how HSNP funding has been spent, or may contribute to planning future phases of the HSNP after the end of Phase 2.

2. **Regular (bi-monthly) monitoring reports**, generating independent evidence on operational processes such as registration, payment and case management, using a series of questions that remain broadly consistent, to permit the implementation team to take stock of progress and make timely adjustments to its operations if required. The regular monitoring reports will also include scope for particular individual focus areas that are of interest to the implementation team on a bi-monthly basis. This will allow the routine monitoring to react flexibly to any special issues and concerns arising during the regular course of programme operations, and to provide information on a timely basis to allow the programme to understand and respond to such issues.

In line with the ToR, and following discussions with the PILU and DFID in the inception visit, we propose the following outputs from this workstream, comprising three standalone studies besides the bi-monthly monitoring:

1. process and institutional capacity assessment;
2. costing study;
3. emergency payments process review; and
4. bi-monthly monitoring reports (through to early 2017), plus an endline synthesis report.

The remainder of this section outlines, for each proposed activity, the timeframe, the research questions it proposes to answer, its intended audience and use, and the research method.

3.2 PROCESS AND INSTITUTIONAL CAPACITY ASSESSMENT

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DATA ANALYSIS METHODS	DAC CRITERIA	EVALUABILITY ISSUES
l. What is the effectiveness of HSNP management structures and processes?	How is the HSNP intended to operate under Phase 2? How, where and why do operations vary in practice? How does the human resourcing of the programme (both staff and non-staff) reflect the tasks to be delivered? How do the levels of physical and financial resources reflect the tasks to be delivered? How are programme operations affected by organisational capacity?	N/A	Interviews with programme implementers and other stakeholders; HSNP operations manual; human resource records; budget and financial records	Qualitative KIIs; quantitative data analysis	Meeting notes will be collated and analysed systematically.	Effectiveness	Medium. The operations manual is already known not to fully reflect day-to-day practice, which may complicate the benchmark against which capacity to deliver is measured. But willingness to discuss operational issues and challenges is mostly good. Good opportunity for coordination of research with a broader capacity review being undertaken separately for the HSNP
q. What is the capacity of the NDMA to deliver the HSNP in the future?	As above	As above	As above	As above	As above	Sustainability	As above

3.2.1 Rationale

In May–June 2014 OPM supported the NSNP with a functional review of the agencies involved in delivering all five government-run CT programmes. This included elements of a capacity assessment, since it compared the stated and actual functions of the agencies with their capacity to carry out those functions in terms of human resources, financing, and equipment and materials. The function and capacity of the agencies delivering the HSNP –the HSNP secretariat, the NDMA, FSD, Equity Bank and HAI, at central and local levels– were a core part of this review.

Since that date the resourcing of the HSNP has changed considerably. In June 2014 the HSNP secretariat was replaced by the PILU. The NDMA county drought coordinators are now supported by a dedicated HSNP programme manager and four programme officers in every county, reporting into the county drought coordinator but contracted by the PILU. Equity Bank is expanding its network of local agents, and even opening new bank branches (e.g. in Kakuma, Turkana County).

With the new structures now in place, it has been agreed that it would be valuable to assess whether the resources now employed match the required task, and where capacity could usefully be strengthened. The assessment will relate specifically to the long-term social protection component of the HSNP (i.e. the regular support to the 100,000 'Group 1' beneficiary households), not the emergency payments component, which will be looked at separately (see Section 0). The activity will encompass KIIs with implementing partners in Nairobi and at least two of the four counties, as well as FGDs at the community level with volunteers and committees.

3.2.2 Timeframe

With a new HSNP operations manual having been introduced in 2014, it was agreed with the PILU that the research team should allow time for the revised processes to become familiar to the implementers, so that we did not assess the capacity of staff members who were only just learning their role and whose skills would very rapidly improve. At the same time, we should not wait too long, since the PILU wishes to have the results in time to make adjustments to its programme. This points to a proposed date of around October 2015 for the fieldwork. The report will be produced in November 2015 and submitted in December.

In early 2015 a risk was expressed by the PILU that the 16 newly recruited HSNP programme officers might have their contracts terminated in May 2015 owing to a change in funding arrangements for the posts, in which case the people occupying those posts might have just been newly appointed and not fully up to speed around October. However, in discussion with the PILU's operations manager it was felt that the research would cover enough other posts that this should not delay fieldwork; and subsequent advice from the PILU indicates that the programme officers' contracts have been extended, so this should not influence the timing.

3.2.3 Research questions

The focus will be on exploring the capacity of different implementing partners to deliver the long-term CT component of the HSNP in its current form.[25] The reference for how the HSNP is expected to operate will be the operations manual, which we understand will be updated in mid-2015 as improvements to processes have already been identified by the PILU.

Any capacity assessment requires the research team first to answer the question, 'Capacity to deliver what?' The team can then analyse the resources that are available to deliver the programme, and consider whether they are appropriate and sufficient for the task. The main research questions are therefore likely to cover the following:

1. **How is the HSNP intended to operate during Phase 2?** The team will summarise its understanding of HSNP processes for the Group 1 households, focusing particularly on day-to-day operations once households are registered, i.e. the processes of payment, case management and monitoring. It will also review briefly the processes that have been used for targeting and registering households, activities that will have largely been completed by the time of the assessment, and will look ahead to any likely changes in those targeting processes for the possible recertification exercise in 2016.

[25] The HSNP is contracting a separate consultancy, led by Ursula Blackshaw, to conduct a forward-looking assessment of the capacity of various partners to deliver a potential Phase 3 of the programme from 2017. This is likely to take a differing form with increased government ownership. We will liaise with the consultant to ensure that the respective assessments are complementary rather than duplicating.

2. **How, where and why do operations vary in practice?** This will explore whether counties, communities or individuals are implementing the HSNP according to the written guidelines and any other training or advice issued; and, if their practices differ, whether this is because they have developed improved ways of operating that could be shared with others in the programme, or whether there are capacity gaps in terms of understanding or resourcing.

3. **How does the human resourcing of the programme (both staff and non-staff) reflect the tasks to be delivered?** This is likely to include both organisational and individual-level aspects:

    - number of programme posts and their role levels, vacancies and turnover;
    - quantity of other resources available to the HSNP (NGOs, bank agents, committees, volunteers);
    - the workload of programme implementers in terms of time allocation between HSNP and non-HSNP activities; and
    - individuals' qualifications, skills, training and knowledge (staff and non-staff).

4. **How do the levels of physical and financial resources at national and county levels reflect the tasks to be delivered?** This includes access to office space, electricity, computers and internet, mobile phones, vehicles and fuel. The financial component will not be a detailed budget analysis but rather a qualitative assessment of e.g. the sufficiency and timeliness of disbursements.

5. **How are programme operations affected by organisational capacity?** This includes the structures, working groups and committees that deliver the programme, and their effectiveness; and the effectiveness of coordination with the NSNP and its SPS.

3.2.4 Method

Clare O'Brien, the workstream leader, and Caroline Riungu, the national coordinator, propose to undertake two–three weeks of field research, with part of the time to be spent conducting interviews and collecting data in Nairobi and part in two counties.

Interviews in Nairobi will examine the capacity of central-level agencies to carry out their functions, as well as draw on their overarching programme perspective to learn their views on the capacity of county- and sub-county-level partners to implement the programme. At local level the interviews will include both people who have responsibility for a whole county and those with responsibilities at sub-county or community level, including volunteers and committee members as well as programme staff. Respondents are likely to be sought from among those indicated in Box 3.

Box 3 Respondents for the capacity assessment

NAIROBI

- NDMA
- PILU
- DFID
- SPS
- FSD
- HAI
- Equity Bank

LOCAL LEVEL

- NDMA staff at county level (drought resilience officer/county drought coordinator)
- HSNP programme manager/programme officers
- Equity Bank staff responsible for HSNP at the branch, and their local agents
- a selection of rights committees
- programme volunteers
- local chiefs

Where possible, for consistency the team will draw on the research questions posed for the capacity assessment undertaken as part of the functional review of the NSNP.

3.2.5 Output

The output of the research will be a report on processes and institutional capacity reflecting the research questions and methodology articulated above. It will be presented at one of the programme learning events, to be held at an appropriate time that will be agreed with the PILU and DFID (see Sections 5 and 6 below).

3.3 COSTING STUDY

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DATA ANALYSIS METHODS	DAC CRITERIA	EVALUABILITY ISSUES
n. Are triggered payments more/less cost-effective than longer-term predictable CT payments?	Are triggered payments more/less cost-*efficient* than longer-term predictable CT payments?	Alpha ratio (share of transfers in total expenditure) Cost-transfer ratio (amount spent on administration for every $1 disbursed)	Interviews with programme implementers and other stakeholders; financial records; timesheets	Qualitative KIIs; quantitative data analysis	Meeting notes will be collated and analysed systematically.	Efficiency	Low. We can only provide a detailed breakdown of costs for Group 2 households between start-up, roll-out and ongoing operations, but not any sub-categories.
o. How much has it cost the HSNP to deliver Phase 2 of its programme, for regular beneficiaries and for emergency payments?	How much money has been disbursed to Group 1 and Group 2 beneficiaries? What has been the pattern of expenditure over time/by funding source/by funding agent (who spends the money)/ by line item/ by activity? (Group 1 households only)	AS above	As above	As above	As above	As above	Medium. Prospects for ascertaining expenditure by PILU are high, provided that accounts are kept as agreed in a manner that disaggregates expenditure by the key dimensions. Prospects for analysing government expenditure on the programme are dependent on government willingness and ability to share financial data.

3.3.1 Rationale

Expenditure on Phase 2 of the HSNP dates back to late 2012, when NGOs began incurring costs on the targeting and registration of beneficiaries. Since that date programme funding has passed through two channels: the GoK's own accounts and DFID. With Phase 2 of the HSNP having introduced many new activities that incur high up-front investment, in particular the opening of bank accounts for all households in the programme area, it is useful to take stock of this expenditure and consider how much has been spent, on what items and on what activities. The costing study will aim to provide this retrospective analysis, also offering accountability to programme funders as to how their resources have been spent, which is useful in the light of the emphasis on value for money.

3.3.2 Timeframe

CT programmes have costs, first, of setup and roll-out –i.e. the design, along with the identification and enrolment of beneficiaries– and then recurrent costs of regular disbursement once the transfer is in place. Until now a large proportion of the administrative costs have been spent on setup and roll-out, owing to the continued work of helping beneficiary households obtain their national identity cards and obtain a bank account, so it has not had the chance to settle into routine disbursement. In order for the costing study to capture the costs of regular disbursement, and not just the early-stage activities of registering beneficiaries, we advise that the study be completed toward the end of the evaluation period. However, we recognise that in order to feed into the business case for HSNP Phase 3, a preliminary report is requested sooner. We will aim to generate the analysis in two rounds, one around April–May 2016 that will capture costs associated with the registration and enrolment phase, and the other a year later.

The time period that can be covered by each round of analysis will depend on how soon after each accounting period the various parties are able to issue reports on their expenditure, e.g. whether they are aggregated monthly or quarterly. So, for example, a visit in April 2016 might be able to capture government expenditure for the first two quarters of financial year 2015–16 (Jul–Sep and Oct–Dec 2015) but may not yet be able to access information for the third quarter, Jan–Mar 2016. We await advice as to the time delay in the production of the accounts, and can arrange our two visits in line with when we are able to collect the information requested by DFID.

3.3.3 Research questions

The research will aim to analyse the expenditure of DFID and the GoK across several dimensions:

Time. What has been the pattern of expenditure between 2012 and 2016?

Accounting source. Who is the originator of the funds? This might include, for instance, the Australian aid agency DFAT and NGOs' own sources, as well as the funds of DFID and the GoK.

Accounting agent. Through whom have the funds been distributed, i.e. who has been responsible for making spending allocations? This is likely to show the distribution between GoK and DFID funds. It should be possible to disaggregate the DFID funds further to show the amounts allocated to its partners, e.g. FSD and Equity Bank.

Line item. What have the funds been spent on? This will disaggregate by budget line, e.g. salaries, transport, communications etc.

Activity. This last dimension is the most critical for enabling policymakers to understand which aspects of their programme are most resource-intensive. It might reveal how much money has been spent on, for instance, registering households, paying them or dealing with complaints. The activities to be analysed are shown in

Table 8 below.

Note that the analysis of GoK funds, and the breakdown of expenditure by other partners whose accounts do not pass through DFID's or the PILU's accounting system, is dependent on the interest of those organisations in participating in the study. We would be grateful for the PILU's assistance in facilitating access to the relevant people.

We propose to do a financial, rather than economic, costing. In other words we will account for what has actually been spent, not for the value of goods and services donated in kind (such as the hypothetical cost of the time spent by volunteers and committee members) or opportunity costs such as working hours lost by beneficiaries due to time spent registering on the programme. An economic costing would entail a lot of assumptions about the value of time spent by participants who do not have regular incomes, and we believe it would not add to the accuracy of the analysis.

Box 4 A note on the questions proposed in the ToR

The research questions presented here represent the outcome of discussions with the PILU during the inception phase, and differ somewhat from those that were proposed in the ToR. The reason for this adjustment is that questions n, o and p in the ToR, cited at the start of this sub-section, require elements additional to a costing study that are beyond the scope of the evaluation agreed with DFID:

Question 'n' aspires to a direct comparison of the *impact* of emergency versus non-emergency payments, but such a comparison is not possible because the payments are targeted to different types of households, for different lengths of time and under different circumstances. The costing analysis will endeavour to assess the total amount spent on running the emergency payments, if they are triggered. It has been agreed with the PILU that this will be calculated as a lump sum, not broken down into different activities of monitoring, case management, etc., in order to limit the quantity of analysis codes to a feasible number (see the method section below).

Question 'o' seeks to explore the feasibility of a CBA. It has been decided that this is not appropriate as it requires the researchers to assign a monetary value to non-monetary programme benefits such as the value of a sense of dignity or security, which would not be part of a financial costing study. The 'benefit' component of a CBA is dependent on the IE, which will not be collecting data on e.g. the education benefits and changes in life expectancy referred to in the proposition. It has been agreed that the best value for money would be obtained by focusing the Phase 2 IE on some key indicators.

Question 'p' implies an analysis of the ASP, which we have agreed will not form part of the costing study.

3.3.4 Method

To produce a cost-efficiency analysis by activity, we need all expenditure to be coded against a consistent set of activities. These codes should be applicable to both salaries and non-salary expenditure, e.g. travel and accommodation in the field. We propose a set of nine codes, designed to be compatible with the DFID format for reporting on CT expenditure (start-up/rollout/operational/external M&E costs), and also with the HSNP operations manual. The codes agreed with the PILU are presented in

Table 8.

The PILU has agreed that these codes will be applied to its records of expenditure starting from March 2015. The records include:

- timesheets of HSNP implementers, in Nairobi and in the counties;

- journeys recorded in logbooks of vehicles owned by the programme; and
- expense claims for visits to the field (travel, accommodation, per diems, etc.).

The code '00', for non-HSNP activity, has been added to allow for the fact that vehicles purchased with HSNP funds may be used in the counties for non-HSNP purposes. We will consider whether and how to account for this at the analysis stage.

TABLE 8 PROPOSED ACTIVITY CODES, AND MAPPING TO DFID

PROPOSED CODE		DFID COST-REPORTING CATEGORY	STAGE OF THE HSNP OPERATIONS CYCLE AS PER THE OPERATIONS MANUAL
01	PROGRAMME DESIGN	Start-up	N/A
02	REGISTRATION	Roll-out	Step 1 – Registration
03	TARGETING	Roll-out	Step 2 – Beneficiary selection Step 3 – Community validation Step 6 – Notification, targeting of complaints
04	ENROLMENT	Roll-out	Step 4 – Identification of recipients Step 5 – Preparation of bank accounts Step 7 – Bank account opening and distribution
05	PAYMENT	Ongoing operations	Step 8 – Payment and reconciliation
06	CASE MANAGEMENT	Ongoing operations	N/A (Dealing with complaints and updates)
07	MONITORING/ REPORTING	Ongoing operations	N/A (Monitoring implementation)
08	MANAGEMENT/ COORDINATION	Ongoing operations	N/A (Linkages with other programmes and authorities, committee meetings, recruitment, general admin)
09	EMERGENCY SCALE-UP	Start-up	N/A (Design of the emergency scale-up mechanism)
		Ongoing operations	N/A (Implementation of the scaled-up mechanism in the event that it is triggered)
00	NON-HSNP ACTIVITY	N/A	N/A

Note: Expenditure under the DFID category of 'External M&E' does not need its own code as this is simply the value of the OPM contract.

During 2015 we will consult periodically with DFID to check that the coding is easy to apply and being used consistently: this preparatory work greatly facilitates the eventual analysis, ensuring that the breakdown of cost is based as much as possible on accurate data about the distribution of expenditure by activity rather than being reconstructed a year or two later. We will also need to approach other partners who spend money under the HSNP –including the government– to see to what extent they are able and willing to classify their expenditure using the same method.

Once the systems for coding expenditure are set up, understood and in regular use, the costing analysis itself can then take place in the two rounds planned to be undertaken over the life of the project by the costing expert.

Box 5 Feasibility of comparisons with other programmes

Several CT programmes have begun to publish general figures relating to their administrative expenditure. These figures can be noted to provide a context for the figures generated under the HSNP Phase 2. However, we caution against any expectation that they will be directly comparable. Programmes have undertaken their costing analyses at different stages of development. Many long-term programmes necessarily spend large amounts of money in setup costs in the first year or so of implementation, and gradually offset these costs against more transfers as time passes, which makes the overall outlay appear increasingly cost-efficient over time. Moreover, different costing studies include different components and assumptions; and in any case can vary enormously in their efficiencies of scale (duration and number of beneficiaries) and implementing context. We will therefore not be able to say that the HSNP is more or less cost-efficient than another programme. We do have information on the cost-efficiency of the Kenya CT-OVC programme during its second phase, as well as recent analysis of three emergency CT programmes in Kenya, which we can report on and discuss.

3.3.5 Output

The output of the costing study will be an analytical report accompanied by an Excel spreadsheet containing the analysis of costs by several different dimensions. A preliminary report just reporting on costs of the registration and enrolment process will be produced early on. The final report will be presented at one of the programme learning events, to be held at an appropriate time that will be agreed with the PILU and DFID (nominally February 2017; see Sections 5 and 6 below).

3.4 EMERGENCY PAYMENTS PROCESS REVIEW

3.4.1 Rationale

Scaling up the HSNP transfers during droughts is one of the objectives of HSNP 2. Emergency payments are triggered once a sub-county reaches severe or extreme drought threshold, as measured by the Vegetation Condition Index (VCI). County officials are also given the opportunity to add additional sub-locations in sub-counties where the VCI threshold has not been reached but are equally affected by drought.

Following poor rainfall in 2014, food security in the four HSNP counties worsened. In February 2015, 10 counties reached severe drought status and two were in extreme drought status. A single scale-up payment was made in April 2015 to 90,000 non-routine HSNP beneficiaries. This payment covered the period between January and March 2015. Further payments will be made.

Evidence on impact of the emergency payments needs to be documented for accountability purposes and to serve the needs of donors who may contribute funds in the future. An impact assessment of the emergency programme has been incorporated in the qualitative component of the IE, as discussed above in Section 2. This will provide a detailed account of how the emergency transfers support or affect households' ability to build resilience (e.g. how they influence the coping strategies they adopt in the face of drought). Such impacts derive from households changing their behaviours in

response to the new transfers and over a relatively long period of time. The longitudinal design of the qualitative IE is thus well suited to explore these dynamics.

In the meantime, given that the first payment was executed as a pilot, the lessons learnt during the first round of payments need to be documented to help refine some of the operational processes and the overall strategy, as well as help in the development of implementation guidelines that will guide the NDMA when making subsequent payments. Preliminary evidence from the field so far has identified three main operational challenges:

- The use of a single remotely sensed indicator (VCI) to generate county CT allocations is justified. However, using the VCI status to re-allocate these CT quotas within the county does not properly reflect the vulnerabilities of these populations.
- The reallocation of CT quotas by the Counties does not guarantee a fair and systematic approach.
- The preselection of beneficiaries using the CBWR/PMT ranking is causing confusion among CT recipients who are not aware of their ranking on the MIS.
- This emergency payment process review aims to provide an independent and systematic account of implementers' experiences in the implementation of the emergency payments. The comments of beneficiaries about the implementation process will be captured as part of the impact assessment above.

3.4.2 Timeframe

The emergency payments study will rely on a number of KIIs and will be led by Roger Pearson. A preliminary desk-based design of the study will be done in May 2015. The study will be undertaken in June–July 2015. A report will be produced in August 2015.

3.4.3 Research questions

This review will rely on a **process tracing exercise** with a number of national and county-level stakeholders. The process tracing exercise will help to map out all activities undertaken as well as to understand decisions and actions taken at key stages and the outcome of these. The exercise will also be used to identify bottlenecks and challenges at each key stage and identify solutions for improving them. From a review of available documentation, Table 9 maps out key activities that were associated with the recent payments, along with accompanying research questions. Beneficiaries' perspectives on the research questions will be gleaned from the IE.

TABLE 9 RESEARCH QUESTIONS

PROCESS	RESEARCH QUESTIONS
Communication and sensitisation	To what extent is there buy-in from the humanitarian sector and media of the emergency payment strategy? Have humanitarian actors and media representatives disseminated the HSNP emergency payment strategy? How do stakeholders, and more specifically implementers, perceive the use of the VCI as a means of measuring drought-related vulnerabilities? What communication approaches were used during the emergency payments? How effective were these approaches? How informed are different stakeholders about the processes involved in the emergency payments? How knowledgeable are different stakeholders about the processes involved in the emergency payments?
Targeting and verification	How were the sub-county/county budgets decided upon? Was the communication regarding the sub-countries/county budget allocation to County Drought Coordinators (CDCs) clear and easy to understand? What was the process in identifying additional sub-locations? Was this done in a fair and transparent manner? What was the process involved in selecting beneficiary households? What were the roles and responsibilities of different stakeholders? What went well? Why? What challenges were faced? Why?

	What are different stakeholders' views on the targeting process used for the emergency payments? What are their perceptions on the preselection of households prior to droughts? Was it fair? How effective was the targeting mechanism perceived to be in reaching vulnerable and drought-prone households? In what way can the targeting process be improved?
Payments	What are different stakeholders' experiences of the payment process? What are community members' experiences of opening and activating bank cards? Were there any problems during the payment process and how could this process be improved? Are beneficiaries satisfied with the attitude of the pay agents? Was the amount received by beneficiaries as expected? Do they think the allocation is a suitable amount? How does this payment value compare with other payments that they receive? What are their views on the frequency of payment? What are the benefits/drawbacks of a system that pays into bank accounts? What were the challenges and constraints that pay agents faced during the first payment? What are the benefits and costs to pay agents in participating in HSNP emergency transfers?
Grievances	What is the process of initiating grievances? How effective is it? What sort of complaints did beneficiaries have? Have these been adequately addressed?

3.4.4 Methods

This process review will rely mainly on KIIs. These will be structured around the key questions highlighted above, but will also allow for discussions around unanticipated results and emergent findings during the period of research.

We will visit aim to cover three of the four affected counties. Box 6 gives a proposed list of interview participants.

The interviews at national level will take place over a period of one week. County-level discussions will then be undertaken simultaneously across the four districts.

Box 6 Respondents for the emergency payments study

NAIROBI

- PILU
- NDMA
- DFID
- Catherine Fitzgibbon
- FSD
- Equity Bank
- NGOs and other donors in the humanitarian and arid and semi-arid land (ASAL) sectors

LOCAL LEVEL

- NDMA staff at county level (drought resilience officer/county drought coordinator)
- HSNP programme managers and officers

3.4.5 Outputs

The output of the research will be a draft emergency payments process review report, detailing the method and findings from the study in answer to the research questions above, along with implications for future programme design. The draft report will be delivered in July 2015. Following this the report will be finalised during a bespoke workshop to be held as part of the review of the emergency scale-up payments, to be facilitated by the evaluation team (see Section 0 below).

3.5 REVIEW OF THE EMERGENCY SCALE-UP PAYMENTS

3.5.1 Rationale

In September HSNP will hold a meeting to review and reflect on the emergency CTs made so far in 2015. The event will be facilitated by the management agent for the HSNP independent evaluation, OPM. The event aims to reach a consensus on what worked well and what can be improved and how to improve the process for future payments. The event includes a wide group of parties interested in the success of HSNP, including both central- and county-level stakeholders.

3.5.2 Timeframe

At the time of writing it is anticipated that emergency CTs will have been received in April, May and possibly July/August. The workshop to review the experience of HSNP emergency payments is thus expected to take place in September 2015.

3.5.3 Methods

The emergency payments process review report (see Section 0 above) will serve as the key background document feeding into the review. The theme of the workshop will be to validate the findings of report and to agree on how to improve the emergency CT process in the future.

The core of the review will be the workshop facilitated by the OPM team. Different stakeholder groups will be brought together to discuss the emergency payments action for a period of one and a half days in a residential location. The following groups and indicative numbers would be brought together in the workshop:

- HSNP PILU, programme managers and programme officers (max. 10);
- HSNP Steering Committee, including: DFID, NSNP-SPS, World Bank, FSD, Equity Bank and HAI (max. six);
- civil servants with fund allocation responsibilities at national and county levels (max. five);
- members of parliament and county councillors from the four counties (max. eight);
- members of the international humanitarian network active in Kenya (e.g. UN Office for the Coordination of Humanitarian Affairs, European Commission Humanitarian Aid and Civil Protection Department, OFDA, WFP, UNICEF, the larger NGOs) (max. 10);
- a selection of village leadership chosen by county PILU staff (max. 10); and
- a selection of the Kenyan media specialised in development issues in the arid counties (max. 4); media participation will be restricted to a briefing and journalists will not be invited to report on the event itself.

The agenda for the workshop will be jointly finalised by PILU and OPM along the following lines:

1. introductions and presentation of the findings of the July process review;
2. group work to discuss those findings and report back on concurrence or suggested modifications to those findings;
3. plenary discussion to reach a consensus on findings;
4. presentation of the draft recommendations;
5. group work to discuss draft recommendations and report back with suggested modifications; and
6. plenary discussion to reach a consensus on recommendations.

3.5.4 Outputs

This work will lead to a redrafted report on the emergency scale-up, with findings and recommendations as agreed with the participants of the review workshop, containing clear discussion of any dissenting views. The report will be shared with workshop participants within five days of the end of the workshop.

There will also be a final report taking on board key observations and suggestions made by participants of the workshop. Participants will be given two weeks to send in comments and a final report will be delivered by OPM one month after the end of the workshop.

3.6 BI-MONTHLY OPERATIONAL MONITORING AND ENDLINE REPORTS

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DATA ANALYSIS METHODS	DAC CRITERIA	EVALUABILITY ISSUES
i. How well does the system based on use of HSNP registration data work as a basis for targeting the transfers and selecting households for a triggered drought response payment?	Are implementing partners fulfilling the requirements of their service-level delivery agreements? What are beneficiary perceptions of service delivery? What are implementing agents' perceptions of programme operations?	TBC during design phase	Interviews with beneficiaries and payment agents; interviews with key informants	Qualitative KIIs, beneficiary and payment agent interviews; quantitative analysis of brief questions to beneficiaries	Meeting notes will be collated and analysed systematically.	Effectiveness	Medium. The ability to reach payment agents and beneficiaries in Wajir and Mandera may be limited by the security situation.
j. How well does the rights and grievances process work?	As above	As above	As above	As above	As above	Effectiveness	As above
k. Do the new payment platform and expansion of financial services provide benefits for beneficiaries and non-beneficiaries?	As above	As above	As above	As above	As above	Effectiveness	As above

3.6.1 Rationale

Too often, independent evaluations that are based only on single end-of-programme assessments generate findings too late for recommendations to be fed into improving programme implementation. A core part of the present assignment is to remedy this by providing timely, ongoing monitoring of routing programme performance, to highlight trends in the effectiveness and efficiency of operations and to enable the programme to update its procedures where opportunities are identified. This regular monitoring will complement the monitoring undertaken by the programme itself, providing an independent perspective on successes and challenges.

3.6.2 Timeframe

We propose to start the design and testing of the research instruments very early in the evaluation implementation phase, i.e. around the time of the July 2015 pay cycle, with the first bi-monthly assessment beginning from around the September 2015 pay cycle. Thereafter results will be updated after each payment round and continuing to early 2017.[26]

3.6.3 Research questions

During OPM's evaluation of HSNP Phase 1, the team focused its quarterly operational monitoring on quantitative surveys of beneficiary households. For Phase 2 we propose that, in light of the considerable operational changes since the previous phase, it will be important to place an emphasis on monitoring the achievements and challenges of the implementing partners, such as payment agents. This will be of particular importance to the PILU because, for Phase 2, Equity Bank is accountable to FSD for a service-level agreement that sets minimum standards regarding, for example, the maximum distance between a beneficiary household and a paypoint, or the time between payments. DFID also has memoranda of understanding (MOUs) with FSD and with HAI. Regular monitoring of processes can help the PILU identify whether the standards of the service-level agreement and the MOUs are being reached. At the same time, regular tracking of progress against key outcomes can provide contextual background to inform programme design on a more continuous basis throughout the evaluation period.

The bi-monthly monitoring will have quantitative and qualitative components. The quantitative component will be relatively consistent from one reporting period to the next and will provide the trend analysis and generate figures on e.g. the number of beneficiaries served, the size of withdrawals requested, the service provided by pay agents, etc. The qualitative component may have some variation each reporting period, depending on the priorities of the PILU and ongoing operational issues that arise.

3.6.4 Method

Quantitative operational monitoring

We propose to sample 24 paypoints per payment round, i.e. roughly six per county.[27] If the number of payment agents does not change much from its mid-January 2015 figure of about 380 recruits, this could mean that by early 2017 we should have reached about two-thirds of agents at least once. The sampling strategy will be finalised during the design phase based on an assessment of the type and quality of paypoint data available. Some paypoints could be sampled purposively if difficulties are identified that merit guaranteed inclusion in the sample (e.g. to ensure paypoints serving emergency payments recipients are included). We plan to cover both urban and remote paypoints, reflecting the geographic stratification used in the IE, to help explore whether different impacts are associated with different operational contexts.

Fieldworkers will visit the sampled paypoints on payment days. They will interview the payment agent, and use stepwise sampling to interview eight beneficiaries of the regular HSNP payments per paypoint. This will provide a sample of 24 payment agents and 192 beneficiaries each payment round. This approach will provide an indication across the four counties of the experiences of all routine HSNP beneficiaries collecting their payments on paydays. We do not necessarily propose to weight the observations to be statistically representative, since the primary purpose of the exercise is to provide usable feedback to the programme; for that it may be more useful to focus attention on a particular sub-group of beneficiaries (e.g. those in a specific location) or to make repeat visits to certain areas. These issues will be discussed further with the PILU during the design phase and the sampling design and presentation of the

[26] The precise sampling approach needs to be finalised during the design phase. Depending on the number of paypoints and beneficiaries required each payment cycle, it is envisaged that the operational monitoring survey will cover somewhere between 10 and 12 payment cycles, thus ending in March, May or July 2017.

[27] The exact distribution may vary to reflect coverage: as of January 2015 Turkana had recruited about twice as many payment agents as Wajir and Mandera.

paypoint survey estimates refined accordingly. The selection of respondents can be discussed with the PILU as each period's fieldwork is planned.

Two features of the programme present challenges to the interview schedule: the regularity of the payment date and the mobility of payment agents:

The programme intends to issue payment on the fifth day of every other month; beneficiaries can collect it any time thereafter. Beneficiaries tend to arrive on the first day of payment if they can. If payments do become regular, this means that much of the quarterly monitoring will have to take place on the same day. But it will also be important for the team to try to interview some beneficiaries who do not arrive at the paypoint on the day when the bulk of payments are made, to understand their reasons for selecting a different day, and whether they have a different perception of the efficiency of programme operations. If payments remain irregular, the team will be prepared to respond at short notice to announcements of when the transfers will be issued.

We understand that some payment agents operate from a fixed location, such as a kiosk or shop, while others are mobile, travelling around remote areas. We will aim to capture the experiences of beneficiaries representing both types of agent. In order to interview the mobile agents we will need to obtain their itinerary from the PILU or Equity Bank.

Examples of indicators for payment agents

We will seek to include indicators that shed light on the programme's ability to meet the targets under its service-level agreements with its partners, if these are available. Typical indicators might include:

- number of beneficiary households served;
- extent to which households come on same day;
- size of withdrawal typically requested (full or partial?);
- ease of maintaining liquidity;
- strategies to ration cash if short; and
- questions on recertification of households, the change in workload and the effect on familiarity of beneficiaries with the programme (if recertification takes place in 2016).

Examples of indicators for beneficiaries

Typical indicators might include:

- distance travelled;
- time spent waiting;
- attitude of payment agents;
- ability to withdraw desired amounts of cash;
- knowledge of, and interaction with, rights committee;
- planned spending patterns; and
- basic outcome indicators (subjective wellbeing).

The method for the quantitative research depends on the team being able to obtain lists of current payment agents from the PILU and/or Equity Bank. Some of this information is available from the programme dashboard, but it will be useful to get an updated list every two months. Equity Bank may also be able to provide more comprehensive data on the indicators shown for all paypoints.

Qualitative research

In each reporting period we will also conduct interviews with committees and volunteers, as well as at sub-county, county and national levels with implementing agencies –subject to their agreement– to understand their experience of

challenges in implementation, and to highlight and share effective solutions they may have developed to resolve them. We would aim to reach two counties per reporting period. Respondents would be drawn from the same groups that were presented in the box 6 above in relation to the capacity assessment; questions could also follow up any findings from that assessment to monitor improvements.

We also suggest that each bi-monthly monitoring report, besides having some standard general questions about programme operations, could have one special theme responding to a major concern of the implementing partners. Our national consultant, Caroline Riungu, could select the special theme each quarter in consultation with the PILU during her attendance at its regular technical working group meetings. Discussions with the HSNP team in the inception phase highlighted the following issues as prospective themes for special attention in the first three monitoring rounds:

1. the relationship between the HSNP and the administrative bodies at county level, including the county governor, county commissioner and county technical working groups, and prospects for closer coordination;
2. strategies for dealing with the response to liquidity shortages; and
3. the role of the rights committees.

3.6.5 Output

We will deliver short bi-monthly monitoring reports with accompanying PowerPoint slides, nine over the lifetime of the project, to provide a timely snapshot of programme operations at all levels. The primary audience of these reports will be the PILU and implementing partners. The form of the report will thus be designed and agreed such that it reflects the needs of that audience in terms of providing both routine data and ad hoc information on particular identified areas in a practical format to aid operational decision making. A one-off task toward the end of the evaluation will be a synthesis of the findings into a single overarching report that analyses trends in programme operations. This is in order to gather the evidence into a single summary document for absorption by a broader range of stakeholders. This latter output will be presented at one of the programme learning events, to be held at an appropriate time that will be agreed with the PILU and DFID (nominally July 2017; see Sections 5 and 6 below).

4. Workstream 3: Policy analysis

4.1 OBJECTIVES OF THE WORKSTREAM

As well as assessing the impact, operational effectiveness and efficiency of the programme, it is important for the evaluation to consider questions of **sustainability** and **relevance**. For this we propose a separate 'Policy analysis' workstream, which will focus on these 'bigger picture' issues. The policy analysis workstream will deliver a targeting study, micro-simulation analysis, and what we refer to as a 'HSNP strategic review'. The targeting study will assess the combined effectiveness of the registration process and selection criteria (i.e. combined PMT/CBWR methodology) in identifying the poorest households for the programme (evaluation question i).The micro-simulation analysis will focus on the potential impact of varying the transfer value, coverage rates and/or eligibility criteria in the future. We will also explore the possibility of analysing the potential impact of varying payment frequency (i.e. larger 'lumpy' vs. smaller frequent payments). In addition we will simulate the impact of the emergency payments. Finally, the strategic review will address key questions of sustainability and relevance of the HSNP, as well as covering its interaction with external institutions and the broader social protection policy context in Kenya.

This section of our inception report sets out our proposed approach to each of these three elements in turn:

1. targeting study;
2. simulation analysis; and
3. strategic review of HSNP (Phase 2 retrospective; Phase 3 prospective).

4.2 TARGETING STUDY

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DATA ANALYSIS METHODS	DAC CRITERIA	EVALUABILITY ISSUES
i. How well does the system based on use of HSNP registration data work as a basis for targeting the transfers and selecting households for a triggered drought response payment?	What are the characteristics of beneficiary households from a poverty and welfare perspective? How do beneficiary households compare to non-beneficiary households in this regard? How do Phase 2 beneficiaries compare to Phase 1 beneficiaries? To what extent did the PMT and	Proportion of beneficiary households above the poverty line (inclusion error); proportion of non-beneficiary households below the poverty line (exclusion error); ratio of beneficiary and non-beneficiary poverty levels; proportion	HSNP beneficiaries and non-beneficiaries with data collected at household and individual levels	HSNP MIS data (Phases 1 and 2) Quantitative IE survey (Phase 2 evaluation) Quantitative IE survey (Phase 1 evaluation)	Descriptive analysis of comparative characteristics of beneficiary and non-beneficiary households.	Effectiveness	Medium. The analysis is limited by the fact that we do not have any data on consumption expenditure for beneficiaries and non-beneficiaries at the point of targeting. Analysis of comparative poverty rates among beneficiaries and non-beneficiaries based on the Phase 2 IE survey data will

	CBWR elements of the Phase 2 targeting process select the same households, and to what extent were the poorest and/or most food insecure selected?	of poor households covered by HSNP; proportion of non-poor households covered by HSNP	yield results that are problematic to interpret. The viability of some of our proposed analysis depends upon how successfully we can link survey data from the Phase 1 evaluation to the current MIS data.

4.2.1 Rationale

Targeting describes a range of mechanisms for identifying households or individuals who are defined as eligible for resource transfers, and simultaneously screening out those who are defined as ineligible. Achieving this seemingly simple objective is one of the most challenging aspects of implementing social transfer programmes, and typically requires trade-offs to be made between targeting accuracy and targeting costs. With its overall objectives of reducing poverty and food insecurity, it is therefore critical that the HSNP evaluates the effectiveness and efficiency of its targeting procedures, especially given changes to the targeting methodology and lessons learnt during transition from the independent targeting review (Fitzgibbon, 2014 and Njagi, 2014).

For the Phase 1 evaluation, we used data from the IE baseline survey to conduct analysis of the targeting effectiveness of the three different mechanisms employed by the programme at the time: CBT, dependency ratio (DR), and an old age 'social' pension (SP). Using the baseline survey data, conducted after the beneficiaries were identified but before they received any cash, we analysed the comparative performance of the three targeting mechanisms in terms of: (1) the ability of each mechanism to correctly identify the poorest in practice, i.e. whether the selected households were in fact among the poorest; (2) the coincidence of selected beneficiary and eligibility criteria (targeting implementation performance); and (3) the coincidence of poverty status and eligibility criteria (targeting design performance). The findings revealed that CBT was the best performing of the three mechanisms, although micro-simulation analysis suggested a simple PMT could potentially have performed even better. These findings influenced the selection of the targeting methodology for Phase 2 (intended to be a combination of CBT (or CBWR) and PMT).

While not such a core aspect of the evaluation as it was for Phase 1, since a new targeting mechanism has been adopted, understanding the targeting performance of the programme is still a key evaluation question for Phase 2. Although somewhat limited by available data (see below), the aim of our targeting study is to assess the targeting effectiveness of the current system, with the ultimate aim to support the HSNP and its stakeholders in assessing options for revising the targeting approach in the future. By analysing the distribution of the beneficiary allocation between and, crucially, within counties, we could help devise an approach to targeting that exploits the combined strengths of poverty-focused resource allocation provided by the PMT, community participation and ownership that comes from CBT, and political acceptability that comes from a rational distribution of benefits within and between counties. This will also be complemented by the work that will take place under the operational M&E (Workstream 2), which will assess the effectiveness of the management and implementation of the targeting process.

4.2.2 Timeframe

We propose conducting the targeting analysis in March 2016 after the IE survey has been conducted and the data processed and analysed. Results from the targeting analysis are proposed to be presented at a learning event (nominally) in May 2016 alongside the results of the micro-simulation analysis described below (see Section 0).

4.2.3 Research questions

The targeting analysis will seek to answer the following research questions:

- What are the characteristics of HSNP (Group 1) beneficiary households from a poverty and welfare perspective?
- How do beneficiary households compare to non-beneficiary households in this regard?
- How do Phase 2 beneficiaries compare to Phase 1 beneficiaries in terms of welfare status?
- To what extent did the PMT and CBWR elements of the Phase 2 targeting process select the same households, and to what extent were the poorest and/or most food insecure selected?

4.2.4 Method

In light of the scoping conducted during the January 2015 inception mission, it is proposed that the Phase 2 targeting study will be based on data from the programme's own MIS data, complemented by analysis of the IE/LEWIE survey (which will cover both beneficiaries and non-beneficiaries). We will also explore the potential for exploiting the HSNP Phase 1 evaluation survey data. The various analyses we propose for each of the data sources are outlined below.

A key limitation that our analysis will seek to address, as far as possible, is that we will not have consumption expenditure data (our primary measure of wellbeing and poverty) for a representative sample of all Phase 2 HSNP beneficiaries and non-beneficiaries from just *before* the point of targeting. Such data is required for a 'classic' targeting analysis that assesses the average pre-intervention consumption expenditure and poverty status among those selected for HSNP Phase 2 (programme beneficiaries) compared to those who were not selected (non-beneficiaries). By contrast, our first round of IE survey data will be collected at some point in mid- to late-2015, sometime after the Phase 2 targeting was implemented (October 2012 to June 2013).

We will undertake the following analyses:

Analysis of programme's MIS data

- What are the characteristics of HSNP Phase 2 beneficiaries?
- How do they compare to non-beneficiaries? Analysing descriptive statistics of household characteristics will show which households are over represented among beneficiaries.
- How do they compare to Phase 1 HSNP beneficiaries? The characteristics of Phase 1 HSNP beneficiaries can be assessed by looking at: Phase 2 MIS (there is a flag to identify Phase 1 beneficiaries in the data); Phase 1 MIS data; and Phase 1 evaluation survey data.
- Was there any variation in household structure between beneficiary and non-beneficiary households? Were polygamous household structures accurately captured in the Phase 2 targeting process?
- To what extent did the PMT and CBWR elements of the Phase 2 targeting process select the same households, and to what extent were the poorest and/or most food insecure selected?
- Assess the degree to which CBWR and PMT scores overlap at the household level using the MIS data.
- Conduct descriptive analysis of the comparative characteristics of households by (i) CBWR wealth groups, and (ii) the PMT-based groupings.
- Conduct comparative analysis of the correlation between poverty indicators available in the MIS data: (i) the PMT score and (ii) the CBWR wealth groups.

- Use the food consumption information collected for at least 50% of the MIS sample to create a simple FCS, based on the classic WFP/Food and Agriculture Organization (FAO) approach. Use this to analyse the extent to which the lowest ranked households using the PMT, CBWR and combined score coincide with the most food insecure households (i.e. those at the bottom of the FCS distribution).

Analysis of Phase 2 IE survey data

The Phase 2 IE survey will provide consumption expenditure and poverty data for a representative sample of all households (beneficiaries and non-beneficiaries) across the four HSNP counties. The sample will be drawn from the programme's own Phase 2 MIS system, which aimed to capture every household in the HSNP areas (i.e. a census).

To what extent were HSNP beneficiaries poorer than non-beneficiaries at the time of targeting? Ideally we would answer this question by comparing pre-intervention average consumption expenditure and poverty status between beneficiary and non-beneficiary households. However, a complication arises because we do not have consumption data corresponding to the time when targeting took place. This is because the MIS does not contain consumption expenditure data and the IE survey will come some one–year years after targeting took place; the programme will have been operating for much of that time.

The consequence is that any analysis of comparative poverty rates among beneficiaries and non-beneficiaries based on the Phase 2 IE survey data will yield results that are problematic to interpret. If beneficiaries appear to be significantly better off than non-beneficiaries, therefore, it may indicate that the programme has reduced poverty among beneficiaries (a key Phase 1 evaluation result) such that they are now generally better off than those not selected. Alternatively, it may mean that the programme did not reduce poverty among beneficiary households, so the results indicate that the targeting process selected better-off households on average. Conversely, if beneficiary households are poorer on average, then this would suggest that targeting was effective in selecting poorer households, because we are reasonably confident (given the Phase 1 evaluation findings) that HSNP CTs will not have made beneficiaries worse off, or non-beneficiaries substantially better off (via spill-over effects). The hardest results to interpret would be if there is no significant (or very small) difference in average consumption expenditure and poverty levels among beneficiary and non-beneficiary households, since this could indicate either the programme's poverty reduction effect or poor targeting performance, both of which are plausible.

Analysis of Phase 1 evaluation survey data combined with the programme's MIS data

Analysing the Phase 1 evaluation data in combination with the Phase 2 MIS data potentially provides a solution to the problem posed above, i.e. the lack of pre-targeting consumption data. Here we would try to perform a 'classical' targeting analysis and determine inclusion and exclusion errors by using Phase 1 evaluation follow-up survey 1 data, which relates to 2010–11.

The most recent Phase 1 evaluation survey that provides representative data of all households in Phase 1 evaluation areas is from 2010–11. By linking these data to the programme's MIS data, which indicate current beneficiary status, we can assess the average consumption expenditure and poverty status in 2010–11 among current beneficiary and non-beneficiary households. Although households' average expenditure and poverty status will not have remained the same between the data reference period (2010–11) and the point at which Phase 2 targeting took place (October 2012 to June 2013), many of those households who were poor in 2010–11 would still be expected to be poor at the time of Phase 2 targeting. Therefore if Phase 2 targeting was effective at selecting poorer households, we would expect current Phase 2 beneficiary households to have lower average 2010–11 consumption expenditure and higher 2010–11 poverty rates, compared to current non-beneficiaries.

The viability of this analysis depends upon how successfully we can link Phase 1 evaluation follow-up 1 survey data to the current MIS data.

4.2.5 Output

The output from the targeting analysis would be an analytical report detailing the results from the above set of analyses, relating these findings to the overarching policy debate around social protection in Kenya and the design of possible future phases of the HSNP. The output will be presented at one of the programme learning events, to be held at an appropriate time that will be agreed with the PILU and DFID (see Sections 5 and 6 below).

4.3 REGISTRATION INSTRUMENT DESIGN

4.3.1 Rationale

Registration for HSNP Phase 2 took place between October 2012 and June 2013. Its purpose was to populate an MIS that requires data for (nominally) every single household in the four counties covered by the HSNP. The MIS requires these data for three reasons:

1. the need to identify the specific subset of the population that is eligible to receive HSNP routine payments (i.e. the poorest households as identified via a combination of PMT and CBWR);
2. the need to identify an additional segment of the population eligible to receive HSNP emergency scale-up payments; and
3. the ambition to utilise the MIS (and accompanying payments infrastructure) to target other interventions, including the other national CT programmes under the NSNP.

The HSNP and its stakeholders are now considering the registration requirements for Phase 3 and beyond. Due to its size and complexity, the registration exercise has a number of significant challenges. And because the data will be used to target a variety of interventions for poor and vulnerable households, the need to ensure sufficient data quality is emphasised. In addition, because the registration exercise is effectively a census, there is a high premium attached to every variable in the dataset.

As part of the Phase 1 evaluation, the evaluation team conducted a review of the PMT to be used for Phase 2 and identified a number of ways in which it could potentially be improved and made more efficient.[28] DFID and the PILU also commissioned a study to draw lessons from the Phase 2 registration and targeting process (Fitzgibbon, 2014), and are now in the process of designing a revised approach to targeting a potential Phase 3, in order to eliminate some of the problems experienced (Njagi, 2014). The HSNP and its stakeholders also have ambitions that the registration data may be utilised for future M&E purposes. They are conscious that registration for HSNP and other programmes will be an ongoing consideration and thus there is a need to think now about how the registration process can become sustainably institutionalised and whether this will affect the design of the instrument for Phase 3.

Moving forward, there is a need to ensure that the registration instrument captures only the data that are absolutely necessary, but at the same time are fit for its multiple purposes and robust within the evolving context of the NSNP.[29] Under this component we thus propose to:

1. conduct a data quality assessment of the current MIS data with a view to identifying any gaps in the data and data quality;
2. redesign the PMT to optimise its efficiency and effectiveness for future rounds of HSNP and other programme targeting;

[28] The World Bank also reviewed the PMT and found that it could be improved.

[29] Terms of Reference Technical Assistance for Harmonized Targeting and Recertification and Consolidation Strategy for the National Safety Net Program. 28 April 2015.

3. design the registration instrument for Phase 3 based on the above to ensure it is fit for purpose (including targeting of HSNP routine and emergency payments plus other NSNP programmes, as well as optimised for future M&E).

4.3.2 Timeframe

The design of the Phase 3 registration instrument will be undertaken between September and November 2015, with the final results submitted in December 2015. This timing is driven by consultation with DFID and the PILU around the next round of programme registration. Any revisions to the timeframe for this activity (for example due to changes to the HSNP and/or wider NSNP planning processes) will be discussed with all key stakeholders.

4.3.3 Research questions

This activity will seek to answer the following questions:

- What is the quality of the current registration data and how can it be improved in the next registration exercise?
- How can the PMT be improved to ensure it maximises its efficiency and effectiveness at identifying poor households in the context of the four programme counties and considering the challenges to data quality?
- How can the registration instrument be optimised for future M&E purposes?

4.3.4 Method

The objective of this exercise it to design a PMT that (a) accurately identifies the poor, (b) is practical and reliable, given existing data and administrative constraints, and (c) is integrated with the overarching M&E system and NSNP framework.

The first step of this exercise will be to assess the quality of available MIS data. This will be crucial in determining which indicators should be prioritised when constructing the PMT. An indicator may be very strongly correlated with poverty, but may be very difficult and costly to collect, making it impractical for large-scale targeting exercises. This is particularly true when using MIS data, which are not normally collected by trained enumerators or subjected to the same level of quality control as survey data.

In order to assess the reliability of the MIS indicators, we will start by performing standard QA tests, such as looking for outliers and patterns of missing observations. In addition, we will compare the distribution of the indicators against comparable indicators in a reference survey, such as the KHIBS, which is known to have reliable indicators. An MIS indicator will be considered to be reliable if it displays a 'plausible' distribution – that is, a distribution that mimics the one found in the reference dataset for a comparable indicator. Based on the findings of this analysis, we will formulate succinct recommendations on ways to improve the quality of MIS data. The specific reference dataset to be used will be selected based on the availability of comparable indicators, the known quality of the reference data, and the level of statistical representativeness of the reference dataset. The reference dataset should be statistically representative at least down to the level at which the MIS provides continuous territorial coverage.

The second step of the registration instrument design will be the construction of the PMT index. This will be preceded by a consultation with key stakeholders to the NSNP (e.g. the SPS and World Bank), in order to ensure compatibility with the ongoing NSNP consolidation agenda.

When constructing a PMT, there is always a trade-off between targeting accuracy and simplicity. To increase the targeting accuracy of the PMT, it is usually necessary to include more and more sophisticated indicators. However, this increases the cost of the registration exercise, and can also increase the likelihood of error at the data collection stage. For these reasons, we will seek to construct a PMT using as few indicators as possible, prioritising robust (i.e. simple, verifiable and objective) indicators that have a low likelihood of being distorted at the data collection stage and that have been validated through the quality assessment on the MIS data (Step 1 above).

The PMT will be constructed in a dataset, such as the KHIBS, that contains both the proxy indicators contained in the MIS and a full consumption module that can be used to accurately identify poor households using standard poverty measures. The validity of the PMT will be assessed using various tests of targeting accuracy, such as inclusion/exclusion errors and a Coady-Grosh-Hoddinott index. In addition to the PMT index based on indicators contained in the MIS, we propose to construct a reference PMT index, which is based on a comprehensive set of proxy indicators contained in the KHIBS. This reference PMT will be used to see how precise the targeting could have been under 'ideal' circumstances – that is, without data quality constraints. The targeting accuracy of the reference PMT will be compared to the targeting accuracy of the actual PMT to assess the reliability of the latter and to estimate the loss of accuracy incurred as a result of MIS data limitations.

The third step of the registration instrument design will involve the transfer of the PMT from the KHIBS to the MIS. This will be done by applying the formula developed in the KHIBS dataset to the MIS data. The reliability of the PMT will be checked by comparing the distribution of the PMT index in the MIS data to the one it had in the KHIBS data. We will also compare the demographic characteristics of households targeted by the PMT in the MIS with those targeted in the KHIBS. The MIS will be considered reliable if the selected households display similar characteristics in both datasets.

In order to reduce redundancy and duplication of data collection exercises, it is important to ensure that the MIS data are aligned, and, if possible, integrated with the M&E framework of the HSNP. Consequently, when assessing and selecting indicators from the MIS, we will consider targets and indicators defined in the M&E framework to see if there are any MIS indicators that can be used to reliably track M&E outputs and outcomes. We will also formulate recommendations for possible additional indicators that could be collected through the MIS system to monitor M&E indicators. In so doing, we need to take into account the quality constraints inherent in MIS-type data and focus on low-cost and reliable indicators that are easily verifiable. To the extent possible, these should be indicators that are required for the registration process and PMT design and that can be collected routinely as part of the normal operations of the HSNP, without imposing an additional burden on staff and beneficiaries.

4.3.5 Output

The output from this activity will be a new draft registration instrument fit for purpose for HSNP Phase 3 and an input into the ongoing NSNP consolidation agenda. It will take the form of a soft-copy instrument that can then be utilised in either hard or soft format for the data collection exercise.

4.4 SIMULATION ANALYSIS

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DATA ANALYSIS METHODS	DAC CRITERIA	EVALUABILITY ISSUES
r. Can we learn anything about the transfer value for future transfers? Could a lower amount have a similar impact for some households? Could more have sufficient impact to justify the extra cost?	What are the costs associated with different programme design options and coverage scenarios? What are the poverty impacts associated with different programme design options and coverage scenarios? What are the local economy impacts (multiplier effects) associated with different programme design options and coverage scenarios?	Programme costs per beneficiary household Coverage: proportion of households receiving HSNP transfers Poverty impact: proportion of households that are below poverty line; proportion of poor households receiving HSNP transfers Local economy impacts: estimated average multiplier coefficient	Quantitative IE survey (Phase 2 evaluation) Programme cost data	Representative household survey covering the four HSNP counties	Simulation analysis	Effectiveness	High

4.4.1 Rationale

It is critical for policy makers to be able to assess, prior to implementation, the likely outcomes and impacts of alternative programme options, their costs and potential affordability, and their cost-effectiveness (the relationship between programme expenditures and outcomes). Micro-simulation analysis refers to the use of micro-data – usually at the household or individual level – to assess the potential costs and impacts of hypothetical policy changes. In the case of CTs and other social assistance programmes, the impacts under focus are normally monetary poverty (or more generally the distribution of gains and losses by socioeconomic status) and possibly key human development indicators. Usually the policy changes under analysis relate to modifications to the eligibility criteria, coverage assumptions, value, payment structure and frequency of the transfers.[30]

[30] Commonly this analysis assumes no behavioural change on the part of households or individuals in response to the policy change, which is a limitation but reasonable given the analytical challenges involved in trying to model and predict these behavioural responses.

Assumptions on the administrative and operational costs are normally estimated on fixed percentages based on international evidence, but in this case could be based on more detailed costing analysis that will be performed as part of Workstream 2.[31]

For the Phase 1 evaluation, we conducted two rounds of micro-simulation analysis. Firstly, it was as part of the targeting study (see Section 0 above). Secondly, in response to requests from DFID and the HSNP, we conducted analysis using official nationally representative survey data (KIHBS), to assess how well the provisional PMT part of the Phase 2 targeting methodology managed to identify the poorest households (OPM, 2013).

The January 2015 inception mission helped to clarify the key areas of policy design on which the simulation analysis should focus. In particular it became clear that this work should cover both the design of the core programme going forward (i.e. the targeting of and payments to Group 1 households) and the ongoing design of the 'scalability' component of the programme.

4.4.2 Timeframe

The micro-simulation analysis will draw on the first round of IE survey data. It is thus envisioned that the micro-simulation analysis for the emergency payments component will take place in October–November 2016, with the results to be submitted in December. The micro-simulation analysis for the core HSNP programme Phase 3 will take place in January-February 2017, with results to be submitted in March. Because there is some flexibility in the timing of these two activities (after the completion of the quantitative survey data processing) they may be subject to change in order to ensure they feed appropriately into DFID and NDMA planning processes. Any changes to the timeframe described here will be agreed with DFID and the PILU.

4.4.3 Research questions

- What are the costs associated with different programme design options and coverage scenarios?
- What are the poverty impacts associated with different programme design options and coverage scenarios?
- What are the local economy impacts (multiplier effects) associated with different programme design options and coverage scenarios?

4.4.4 Method

For both the core programme and the 'scalability' component, the simulation analysis will assess the impact of alternative design scenarios in relation to:

1. cost;
2. poverty impact; and
3. local economy impacts (multiplier effects).

The micro-simulation analysis of the core HSNP component will be conducted using data from the first round of the IE survey. A simulation model will be constructed to predict the cost and impact of different programme design options on the relevant population groups and the local economy. The assessment of the local economy impacts will build directly on the LEWIE study, which is being implemented as part of the IE (Workstream 1; see Section 2 above).

The simulation analysis of the programme's emergency payments component was not originally anticipated, and is subject to additional resources being made available (as part of the budget extension requested to cover the assessment of emergency payments that was requested by DFID and the programme during the course of the inception

[31] This depends on the timing of the micro-simulation study. Currently, this is set to be conducted prior to the completion of the costing study. But we can be flexible in this regard and will seek to conduct the study at an appropriate moment useful for the HSNP in terms of its own policy development agenda.

mission). This analysis will complement and extend the work already conducted by DFID since it will be based on up-to-date data that are specifically representative of the four HSNP counties (our IE survey data). With our survey scheduled for September–November 2015, we hope to be able to conduct the simulation analysis of scalability design options in June 2016, and therefore feed into the process for fine-tuning the design of the emergency payments component under Phase 2, as well as for future phases of the programme. The policy design scenarios to be assessed will be those that are currently being proposed as part of the ongoing design work.

In contrast to the scalability design scenarios to be assessed, which are already determined, the analysis scenarios for the core programme design are not clear at this stage, since they should be determined by the design options being considered for possible HSNP Phase 3. Therefore the timing and focus of this component of simulation analysis work will be pinned down once the timeframe for the Phase 3 design process is established.

4.4.5 Output

There will be two outputs from the micro-simulation analysis: a report detailing the results of the micro-simulations analysing different design options for the core HSNP; and a report detailing the results of the analysis considering different design options of the emergency payments component of the HSNP. The former of these two reports may incorporate, or be later amended to incorporate (depending on the agreed timing of the initial analysis), the results from the costing study described in Section 0. Both outputs will be presented at one of the programme learning events, to be held at an appropriate time that will be agreed with the PILU and DFID (see Sections 5 and 6 below).

4.5 STRATEGIC REVIEW

EVALUATION QUESTIONS	SUB-QUESTIONS	MAIN INDICATORS	DATA SOURCES	DATA COLLECTION METHODS	DAC CRITERIA	EVALUABILITY ISSUES
Has the HSNP been successfully designed and implemented in line with its objectives, and to what degree to are these objectives aligned with the national social protection strategy?	What was the original objective of the HSNP? How has the objective changed over time and why? What were the enablers and constrainers that determined the scale-up? How does the HSNP relate to the broader social protection policy framework within the country, and particularly in pastoral areas of northern Kenya?		Survey of key stakeholders Official documentation (incl. minutes of key meetings) HSNP programme design and strategy documentation Survey of key stakeholders	Interviews during country visits Documentation review	Sustainability	Medium
h. To what extent does national, regional and institutional-level policy dialogue contribute to	What is the extent of national and regional policy dialogue (NDMA, development partners, and social protection sector)? Has the NDMA been	Evidence of move toward harmonisation of social protection programmes in the sector	Survey of key stakeholders National government budget data	Interviews during country visits Documentation review	Impact	Medium. Assessing NSNP budget allocations is straightforward, but attributing these to national and regional

changes in policy coherence, budget allocations of the GoK, and upscaling of the programme?	effective in relation to increasing the HSNP budget and building consensus on the way forward for HSNP and joined-up social protection policy? Has it had any effect on NSNP budget allocations at national and county levels? How about the national and regional-level dialogue? Has this contributed to changes in budget allocations?	Budget allocations to NSNP programmes Increased budget allocations at county levels	Official documentation (incl. minutes of key meetings)			policy dialogue is challenging.
t. In areas where the HSNP is linked to other activities, is the combination of a targeted CT, mechanism for additional triggered cash transfers and additional work on livelihoods an appropriate response to the situation in the four counties?	Is the combination of a targeted CT with a mechanism for additional triggered CTs an appropriate response to the situation in the four counties? To what extent is integrating this with additional work on livelihoods feasible and/or desirable?	Number of HSNP objectives that correspond directly to NSNP objectives	HSNP programme design and strategy documentation NSNP strategy and intervention scale-up plans HSNP Phase 2 evaluation outputs Survey of key stakeholders Evaluation and finance documentation relating to livelihoods projects	Interviews during country visits Documentation review	Sustainability	Medium. Assessing the degree to which the design of HSNP (plus trigger mechanism) is line with NSNP is straightforward, but assessment of the appropriateness of additional work on livelihoods depends on the extent to which these complementary interventions were implemented effectively.
u. Degree to which the cost-efficiency of the HSNP/ASP compares with other interventions (e.g. food aid) in the four HSNP counties	How does the cost of HSNP (cost of dollar transferred and cost per transfer) compare to that of WFP food aid (if those data are available from WFP)?	Programme costs per beneficiary household Average value of support provided per beneficiary (per month)	HSNP Phase 2 evaluation outputs (costing study) Evaluations and financial data relating to WFP food aid	Interviews during country visits Documentation review	Sustainability	Medium. Comparing unit costs between programmes is straightforward if we can access the data.

4.5.1 Rationale

As well as addressing the core evaluation questions related to how well the programme delivers its main objectives, there is a need to review the programme's underpinning objectives and the degree to which these are aligned with the broader social protection agenda in Kenya, and in particular with the national social protection strategy.

4.5.2 Timeframe

It is proposed that the strategic review will take place between October 2015 and March 2016. The report will be submitted in April 2016. This is in order that the strategic review may feed into the Phase 3 business case (see Sections 5 and 6 below).

4.5.3 Research questions

The strategic review will aim to answer the following questions:

1. Has the HSNP been successfully designed and implemented in line with its objectives, and to what degree are these objectives aligned with the national social protection strategy and other social protection programmes?
2. To what extent does national, regional and institutional-level policy dialogue contribute to changes in policy coherence, budget allocations of the GoK, and upscaling of the programme?
3. Has the combination of targeted CTs, additional emergency-triggered payments and support to livelihoods been an appropriate policy response in the four HSNP counties?
4. What is the cost-efficiency (cost of dollar transferred and cost per transfer) of HSNP vs. WFP food aid?
5. A prospective assessment of the extent to which the proposed design of Phase 3 is aligned with the NSNP, and in particular the plans for rolling out the other three key social protection programmes in the four HSNP counties.

We will consider a range of policy strategies recognised as being effective in order to build a framework that will inform our research on policy analysis. These strategies include: building policy coherence; consensus building; investing in relationships; coalition building; influencing strategies; evidence creation; and filling the knowledge gap. We will analyse these in relation to uptake and scale-up of the HSNP; targeting within the HSNP; and coherence and placement of HSNP within the broader social protection sector of Kenya (including, specifically, northern Kenya).

We will examine how the policy process has shaped the design and roll-out of the HSNP. We will investigate under what conditions the strategies listed above have been successful in producing commitment to the HSNP scale-up that is institutionalised and coordinated. We will look for evidence of the existence or future plans for the HSNP as an indicator of policy influence and sustainability.

A key focus will be on how the HSNP will interface with the other key social protection programmes currently operating in Kenya (CT-OVC, OPCT, CT-PWSD) as these other programmes expand into the four HSNP counties. It is critical for planning and budgeting to have a clear policy design for how multiple overlapping social assistance programmes interact. For example, if households are potentially eligible for more than one transfer scheme, and the choice is left to them which to choose (or simply gets determined by the sequencing of the roll-out of the various programmes), programme planning and budgeting and planning is complicated because it is not possible to estimate beneficiary numbers based on the number of eligible households/individuals.

4.5.4 Method

In the study we will use a range of methodologies, including KIIs, stakeholder analysis and desk-based review. The HSNP strategic review will be undertaken using the following methodology:

- Conduct a desk-based review of a wide range of documents on HSNP programme design; targeting studies; implementation documents; documents from the other main social protection programmers; and social protection strategy and policy and related guidelines.
- In collaboration with the GoK (including the NSNP, NDMA and HSNP PILU) and DFID, generate a list of key partners and stakeholders to be interviewed.
- Develop a short open-ended survey on issues described above, to conduct either by email or face-to-face with stakeholders. This will happen in the months of September–November. Results will be collated and analysed.
- Conduct KIIs with key officials and programme implementers, as well as donor partners in the social protection sector, to triangulate survey results and gain a deeper understanding of perceptions and future possible directions in relation to:
    - harmonisation of the different social protection programmes;
    - consolidation of targeting approaches and lessons from different targeting methods; and
    - views on the scale-up of HSNP and how this will relate to other programmes.
- Analyse the data from the survey and KIIs and synthesise them into a summary report.
- (If possible) in collaboration with WFP, acquire cost data on food aid so that we are able to compare them with cost-efficiency data for the HSNP generated by this evaluation (Section 0 above).
- Present the findings of the strategic review at a stakeholder workshop to be held at an agreed time (currently proposed to be in mid-2016).

4.5.5 Output

The outputs from the HSNP strategic review will be a report analysing the above research questions and a presentation of the findings to a key stakeholder group at a dissemination event planned in conjunction with DFID, NDMA and NSNP stakeholders (see Sections 5 and 6 below). A two-page brief will be produced on forward-looking opportunities and challenges for the HSNP.

5. Workstream 4: Communications and learning

5.1 OBJECTIVES AND RATIONALE OF THE WORKSTREAM

This workstream will ensure that all the stakeholders in relation to the HSNP are aware of the evaluation and understand the approach. It will make sure that the results and information produced by the various evaluation activities are disseminated to the relevant audiences using appropriate media. It will work in partnership with the PILU to develop and implement communications and learning with the overall objectives of supporting the PILU and NDMA to understand, accept and integrate M&E tools and approaches into their day-to-day evaluation and monitoring.

To ensure ownership of results, we propose a three-pronged strategy focused on programme learning, wider communications and support to the HSNP's existing monitoring framework. This workstream will thus ensure an inclusive learning environment and that all key stakeholders are kept informed and supportive of the evaluation process and its outputs.

Through strategic communication with targeted stakeholders and in partnership with PILU, OPM will stimulate dialogue at both the national and county levels in order to further support the HSNP and wider NSNP.

In partnership with the PILU, the outputs of the communications and learning workstream will be the development of a detailed, actionable and measurable communications and learning strategy document, and the implementation of the communications and learning strategy over the evaluation period from June 2015 to September 2017.

5.2 TIMEFRAME

A stakeholder analysis is key to determining the development of tactical elements within the strategy, such as communication activity, message, medium and measurement. The evaluation team will undertake a stakeholder analysis that will be validated with input from the NDMA, the PILU, DFID and other key NSNP stakeholders.

With this information, and in line the communication objectives articulated above, the communications and learning strategy document will be developed in partnership with the PILU by end of August 2015. However, as some outputs such as the presentation of the inception report may be required prior to August, ad hoc communication activities will be implemented prior to finalisation of the communications and learning strategy document.

Throughout the remainder of 2015, 2016 and 2017, the communications and learning strategy will be implemented. The strategy and its action plan are not static documents, but rather 'living documents' with ongoing inputs, including measurement of results achieved, such that they will evolve over time.

5.3 METHOD

In partnership with the PILU, OPM will develop the communications and learning strategy document that will become part of the PILU communication strategy. It will outline and direct communication activities through a detailed and calendared action plan. Consultation with NDMA, SPS and DFID working groups will be scheduled to leverage existing materials and media and reduce duplication. Communication activities include elements such as learning events, customised briefs and digital messages.

5.3.1 Communications and learning strategy document

The communications and learning strategy document will support the overall programme evaluation's objectives by taking the outputs from the evaluation's assessment and analysis and presenting them to targeted stakeholders. The strategy will act as a conduit between OPM's workstream outputs and the stakeholders of the HSNP in order to keep them involved and informed, with the ultimate aim of stimulating dialogue at national and county levels on how to further develop safety nets in pastoralist areas and the goals of the HSNP.

The objective of the communications and learning strategy is to enable the PILU and NDMA to understand, accept and integrate M&E tools and approaches into their day-to-day evaluation and monitoring (for example IE and operational monitoring). This will be achieved through such activities as 'learning events' in which stakeholders will engage with the insights from the evaluation's outputs, as well as be provided with opportunities for input into those outputs, and be assisted in how best to make use of them.

An indicative content summary for the strategy document is given by Box 7.

Box 7 Examples of communications and learning strategy document contents

1. Communication objectives
2. Target stakeholders (defined, mapped and prioritised)
3. Messages (customised for each stakeholder group based on their needs)
4. Media (customised for each stakeholder group based on its situation)
5. Results measure (pre- and post-benchmarks)
6. Action plan (calendared and responsibility assigned)

Key to developing the strategy is a strong understanding of the key stakeholders. Through a mapping exercise, stakeholders will be plotted on the stakeholder map. Understanding where each stakeholder group falls on the map allows the communication teams at OPM and PILU to determine the message and prioritise the level of focus and effort, as well as determine the type and frequency of communication.

The messages and content for each targeted stakeholder group will be identified and developed based on the stakeholder's existing position and the desired outcome or movement of that stakeholder on the stakeholder map. In general, the message content will be focused on the outputs of the other workstreams and tied to the overall objectives of the HSNP. For example, Workstream 3: Policy analysis includes an analysis of HSNP targeting performance and a redesign of the HSNP registration instrument. The results of these activities will inform the design of the possible HSNP Phase 3. We will make sure that the information these analyses produce is understood by the necessary stakeholders and provided in an appropriate format such that it can and will be used for that purpose.

The communications and learning strategy document will detail a number of communication activities such as tailored 'learning events', which will include the presentation of key evaluation outputs. It will also include the development and dissemination of materials produced in various formats appropriate to the key audiences, such as but not limited to, hard and soft-copy reports, presentations, briefs, summary notes and brochures, social media messages and web site content – all with the goal of increasing demand for evidence-led policy and reaching closer consensus on the future evolution of the HSNP. To this end, the strategy document will include a methodology by which the evaluation team can measure its results and progress against the aims of the communications and learning workstream.

5.3.2 Implementation of the communications and learning strategy

Key to the success of implementing the communications and learning strategy is a close partnership with the PILU. Here OPM will support the PILU as a champion of the strategy.

The action plan of the communications and learning strategy document plan will outline the various activities to be undertaken, including the 'learning events', when and how they will be implemented, who the key audiences are, and what the key messages/content are.

The communications and learning strategy will also seek to measure the success of its activities, as this will be instrumental in ensuring we continuously improve our communication in order to meet the goals of the evaluation and those of the HSNP. The timeframe for the various outputs from the evaluation is nominally specified in Section 6. The precise timing for each of these activities and outputs will be refined and agreed during the communications and

learning strategy development phase, with any subsequent alterations in timings made in conjunction with the PILU based on the evolving national context.

5.4 OUTPUT

The output from this workstream will be a detailed, actionable and measurable communications and learning strategy document that will be integrated into the PILU communication strategy and championed by the PILU, with the implementation of that strategy supported by the evaluation team over the course of the evaluation period. Many of the other evaluation outputs will also be released via the HSNP website (after discussion and agreement with the PILU and DFID).

6. Work plan

Figure 1 Timetable for activities by workstream

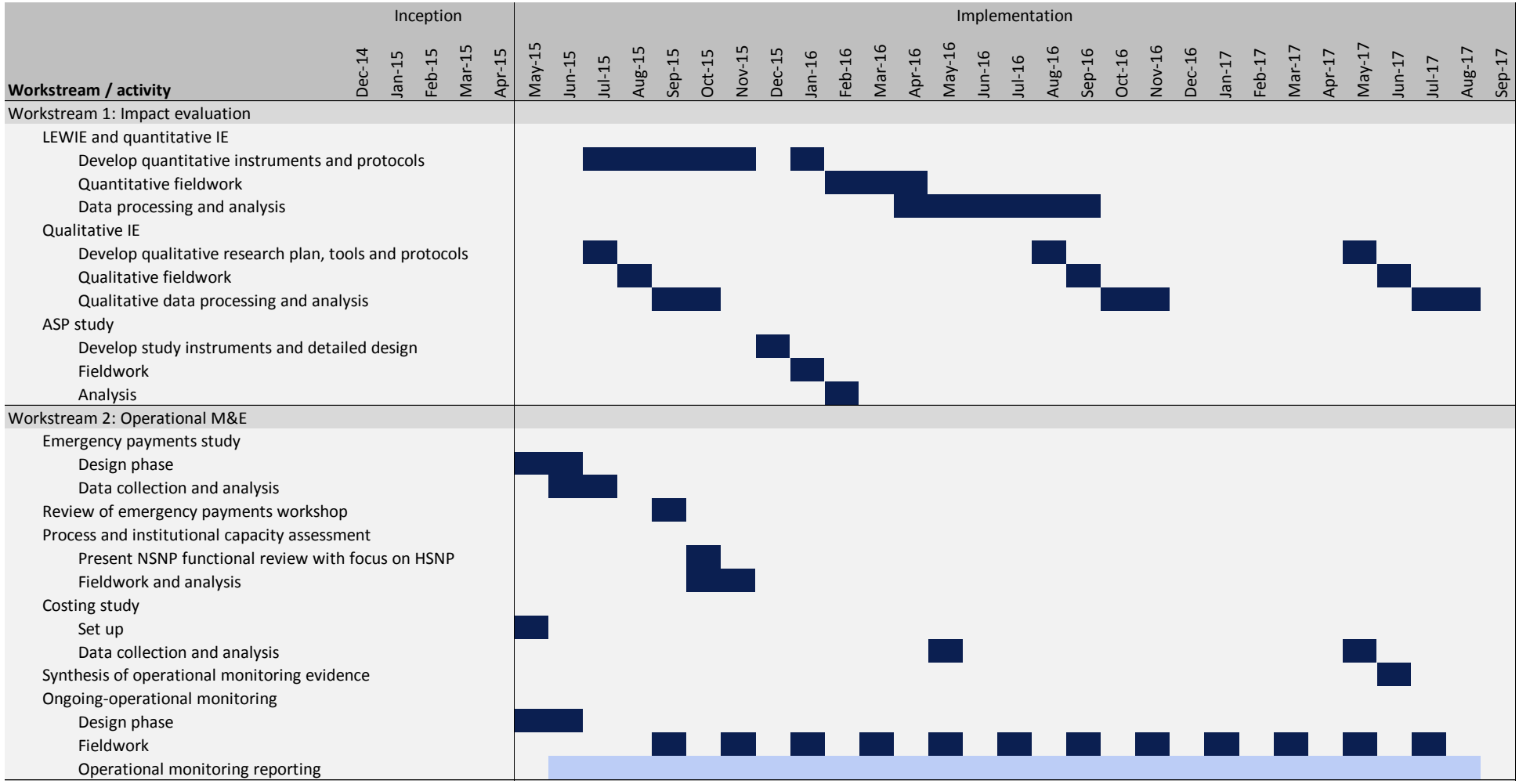
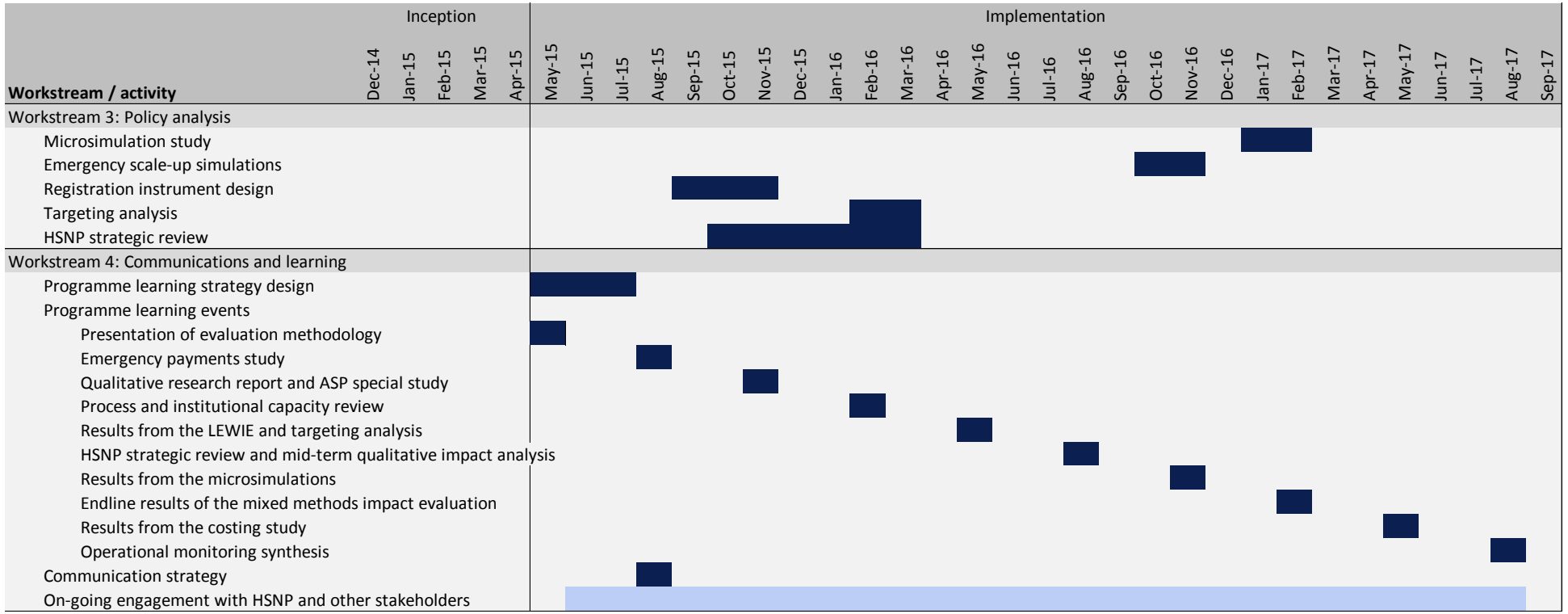
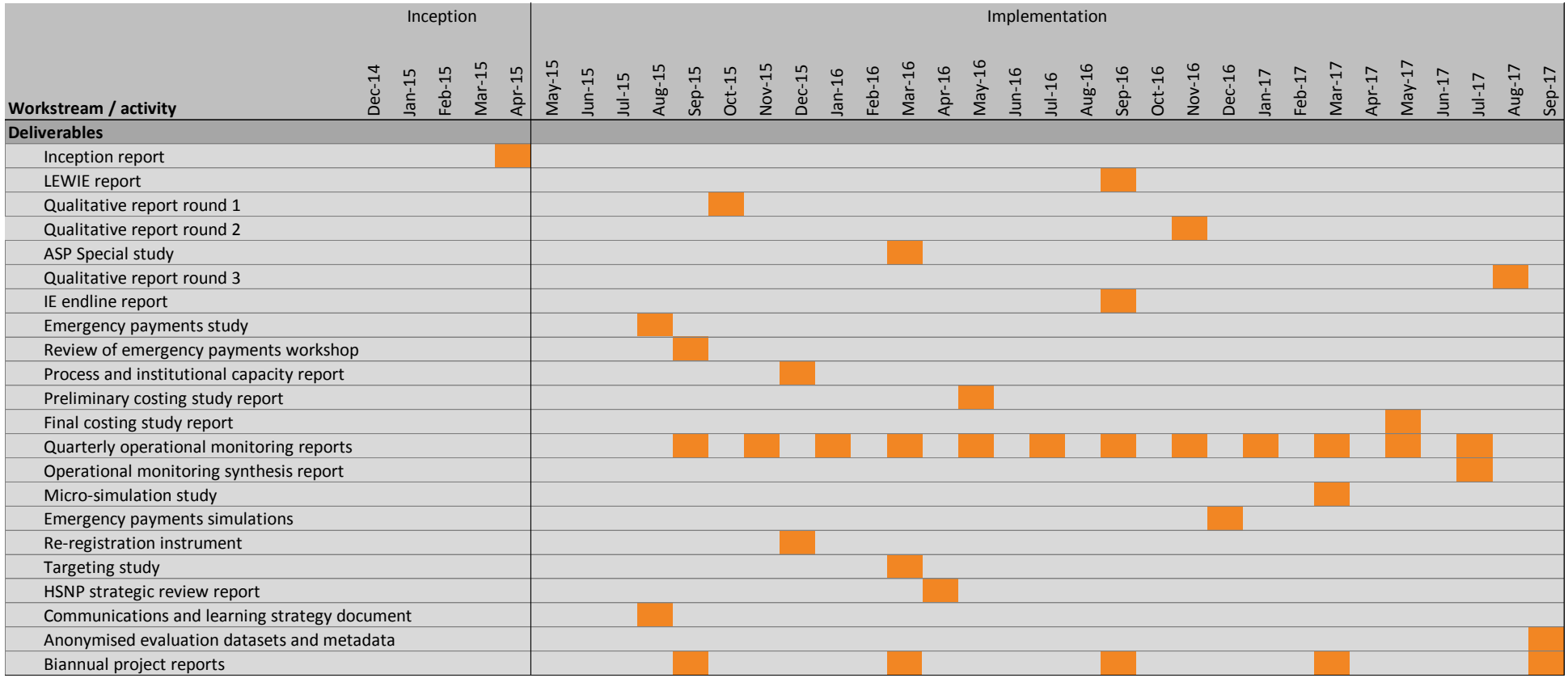


Figure 2 Timetable for activities by workstream (continued)



The precise timing and arrangement of learning events will be laid out in the communications and learning strategy. These may shift and move depending on agreement with the PILU and DFID based on consideration of the most apt time for each event to be held.

Figure 3 Timetable for project deliverables



7. Technical team, management and governance

7.1 TECHNICAL TEAM

This section sets out the technical team. The profile of the TL is provided in 0. The profiles of the other long-term key personnel are summarised in 0. The team has four pertinent characteristics:

Our TL has a proven track record of successfully guiding complex evaluations and providing strategic direction to social protection policy and programming in many developing countries.

Our TL is supported by a full-time KNPC who has extensive expertise in M&E and experience working in northern Kenya.

We have a core evaluation team of 12 experts, five of whom were involved in the Phase 1 HSNP evaluation. This team has the contextual understanding to navigate the cultural, political and practical complexities of northern Kenya, coupled with a global knowledge of best practice in M&E and social protection.

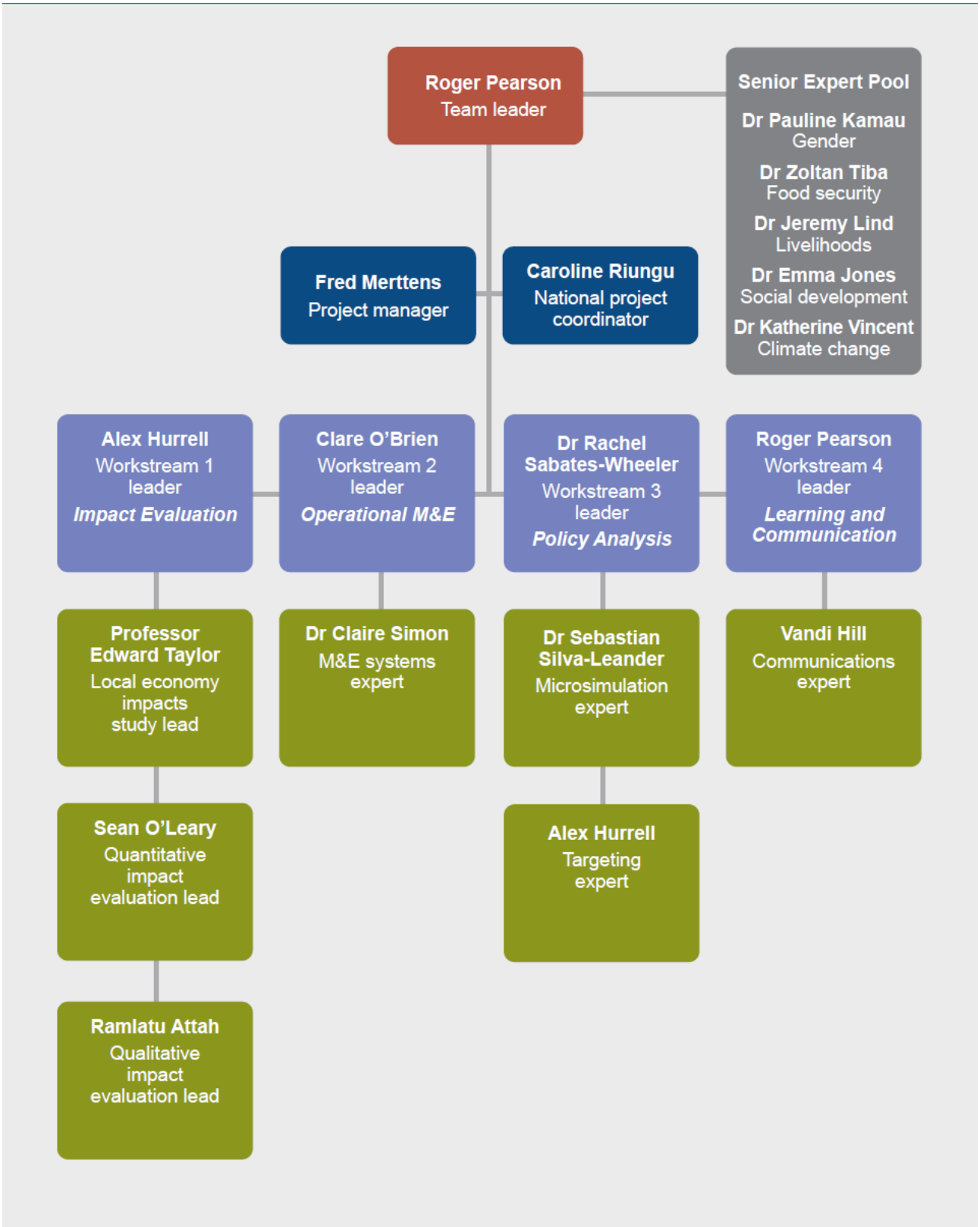
Supporting the core team we have an experienced pool of senior experts who, combined, give us first-class technical expertise in all the key areas required to deliver the evaluation.

The team structure is designed to respond to the needs of this programme. It comprises a mix of leading international expertise and strong Kenyan knowledge and experience. This is embodied in the pairing of an international TL and full-time KNPC, both based in Nairobi for the duration of the project. All of the field research will be conducted by leading Kenyan field research company, RGA, with continuous involvement from relevant international team members under each of the four proposed workstreams.

Below in Section 0 we outline the requirements for each key long-term role in the core evaluation team, as well as a summary profile for each individual selected for that specific role.

The organisation of our technical team is shown in Figure 4. Note that this relates to the organisation and reporting lines for the technical team; the project management roles and relationships are presented separately in Section 0 below.

Figure 4 Team structure



7.1.1 Quality of TL

Our selected TL – Roger Pearson

The team is led by Roger Pearson, who has 30 years' relevant work experience. He is supported by Caroline Riungu as national project coordinator. Fred Merttens of OPM will undertake the day-to-day management and coordination to allow the TL to focus on the technical aspects of the work plan and thus keep costs down. The TL will help ensure consistency and coherence across all the workstreams, and provide strategic direction to the evaluation as a whole.

Roger has over 30 years of experience working on development programming in Africa, South Asia and headquarters locations. During his career he has managed complex evaluations including those with a focus on social protection issues from the point of view of an evaluator, evaluation facilitator and commissioner of evaluations. He has worked on data analysis, with three years as part of the core Demographic and Health Surveys team, and helped to pioneer the introduction of sentinel community surveillance in UNICEF in the early 1990s. He has also worked in senior management positions in UNICEF headquarters in South Asia, Kenya and Ethiopia. He has worked on the front line managing humanitarian service delivery, focused on resilience programming and linking up of emergency and development programmes.

As TL, Roger will provide overall technical leadership and have ultimate responsibility for ensuring the quality of all technical outputs (see Section 0 below for details of the project management structure and reporting lines). He will also lead Workstream 4 (communications and learning). Based full time in Nairobi, Roger's in-country presence will be crucial to ensuring that our evaluation team can be responsive to the evolving demands and priorities of the programme. Below we set out how Roger's profile satisfies our five TL selection criteria.

1. Ten years' experience working on complex evaluation assignments, specifically within social protection

Roger has almost 30 years of experience working on evaluations of social protection programmes. Early in his career he led the evaluation of the 1984 famine relief programme in northern Sudan. In the early 1990s he managed a global evaluation of UNICEF's response to emergency situations with the results reviewed by the UNICEF executive board. More recently, he conceived and designed the evaluation of the OVC CT programme in Kenya, preparing the ToR together with the GoK, and helping the GoK to manage the process until his departure to Ethiopia in 2008. In 2010 he conceived and designed the evaluation of the Tigray OVC CT programme by helping the regional government finalise its ToR, managing a tender for consultants and then working closely with the evaluators on the implementation until his departure in 2014. Other notable evaluations of social protection programming Roger has been involved in recently include: an evaluation of the Ministry of Health's mobile health team strategy in the Somali Region of Ethiopia; the evaluation of the efforts to eliminate female genital mutilation/cutting in Ethiopia; an evaluation of the health extension worker programme in Ethiopia; and an evaluation of the impact of the interactions between community nutrition and water, sanitation and hygiene programmes in stunting reduction in Ethiopia.

2. Ten years of professional expertise in social protection policy, research and influencing policy and strategy

Most recently Roger spent six years as UNICEF Ethiopia's chief of research, evaluation, policy and monitoring. One of his responsibilities was leading UNICEF's contribution to social protection policy. In 2008 he supported the government of Ethiopia's participation in the African Union deliberations leading up to the adoption of the African Union Social Policy Framework. He worked closely with the African Union commissioner for social affairs in the drafting of the chapter. In 2009 he played a key role in coordinating international development partners' responses to the prime minister's request to strengthen the articulation of national social protection policy in the national development plan. The government accepted the suggestions and this led to the creation of the National Social Protection Platform, the key national-level forum through which the national policy, strategy and action plan has been developed. Roger was a consistent supporter of the platform, funding its secretariat and providing consistent support over five years in supporting the public dialogue and eventual drafting of the policy and strategy. This dialogue included efforts to bridge the work of the

humanitarian responses coordinated by government and the implementation of the mainstream national development programme in pastoralist areas.

3. Excellent communication skills and proven ability to use complex evaluations to bring about change

Roger is noted for his excellent written, verbal, and presentational skills. He has and made numerous presentations to senior government and development partner officials over the last 20 years, and authored reports including evaluations and management responses to evaluations. In evaluation settings Roger is keen to listen to the clients and clearly understand what their key evaluation questions are, as well as the detail they require in terms of evidence to help them reach decisions on actions stemming from the evidence. These communication skills, coupled with proven experience in using complex evaluations to bring about change, will be crucial to the success of the communications and learning workstream, which Roger will lead. This workstream underpins our vision of an evaluation that will not only provide programme implementers, DFID and the GoK with the evidence on programme performance required for future decision making and accountability for funding, but also lay the groundwork for a mature HSNP with the capacity to implement its own effective and efficient evaluation, monitoring and policy analysis activities.

4. Substantial leadership and senior management experience, including of large multidisciplinary teams

From 2008 to 2014 Roger was the chief of research, evaluation, policy and monitoring for UNICEF Ethiopia. He was a member of the senior management team, managing a programme of $170 million per year and a staff of 450. From 2001 to 2008 Roger was the deputy representative for UNICEF Kenya. By the time he left he was responsible for a $40 million per year operation and a multidisciplinary team of 150 staff. Prior to this, from 1994 to 2001 Roger was one of the secretaries to the UNICEF South Asia Management Committee. He had responsibilities relating to reviewing the strategic coherence of UNICEF programming across South Asia, focusing on budgeting for research, evaluation and monitoring, human resource deployment and budgets by country programmes in these areas. In the early 1990s Roger spent six years reporting, via the director of the Evaluation Office, to the executive director of UNICEF on the commissioning of evaluations of UNICEF's work at the global level, answering evaluation questions posed by the executive director.

5. Substantial knowledge of CT programmes in sub-Saharan Africa

Roger facilitated the creation of the Kenya OVC CT programme, starting in 2002 with the lobbying of parliamentary candidates on OVC issues, through to the advisory work with the parliamentary OVC committee leading to supporting it in the creation of a pilot CT programme. Apart from South Africa this was the first child-focused cash transfer programme in sub-Saharan Africa and led to hosting several visits, taking part in network meetings and subsequently visiting several other CT programmes. Starting in 2009, Roger spearheaded the support to regional governments in Ethiopia in helping them conceive and create their own pilot CT programmes managed by the Bureaus of Labour and Social Affairs. Notably Roger helped the Tigray regional government discuss its strategy and design its CT programme, including the evaluation design in 2009/10, and supported implementation starting in August 2011. The design of the pilot was based on testing the extent and capacities of the Bureau of Labour and Social Affairs to take over responsibilities for transfers to the welfare cases currently managed by the Productive Safety Net Programme (PSNP). Starting in 2013 Roger was part of the design team for the next generation of rural safety net programmes.

Roger is part of a group writing a book to be published by Oxford University Press in 2015. The subject is the evolution of CT programmes, focused on child grants, over the last 10 years in Africa and the influence of these programmes' IEs on CT policies.

7.1.2 Quality of technical team

Our team has the collective knowledge and experience required to deliver a project of this size and complexity. In addition it is a particular strength that half of our core evaluation team members (including the project manager and

three of the four workstream leaders) were involved in the Phase 1 HSNP evaluation. The combined expertise of the team is summarised in the skills matrix table presented below (Figure 5).

The profiles of the core project team are summarised below. Full CVs were presented in Annex C of the technical proposal. In addition to the core team, we have a senior expert pool to provide specific technical or sectoral expertise and advice as required. The profiles of our proposed pool were presented in Annex A of the technical proposal.

Figure 5 Overview of project team's skills and experience

Role	Name	Hunger Safety Net programme	Social assistance/ Social safety nets	Social protection policy/ strategy	Managing complex evaluations	M&E design and implementation	Experimental and quasi-experimental impact design	Quantitative data analysis	Qualitative data analysis	Participatory research tools	Livelihoods, resilience and income generation	Institutional/operational reviews	Political economy analysis	Targeting analysis	Gender	Social development (empowerment, social inclusion analysis)	Nutrition/food security	Climate change, resilience, disaster management and relief	Communications strategy	Local Economy Wide Impact Analysis (including social accounting matrix)	Micro-simulation	Cost-benefit/cost-effectiveness analysis; Value for money analysis	Skill development, training and capacity-building	Kenya experience	East Africa experience	Working with high-level government officials
Core team																										
Team Leader	Roger Pearson		●	●	●	●	●	●	●		●				●		●	●					●	●	●	●
National Project Coordinator	Caroline Riungu					●		●			●												●	●	●	
Project Manager	Fred Merttens	●	●	●	●		●	●						●										●	●	●
Impact Evaluation Workstream 1 Leader and Targeting Expert	Alex Hurrell	●	●	●	●	●	●	●						●							●			●	●	●
Operational M&E Workstream 2 Leader	Clare O'Brien	●	●	●	●		●	●	●	●		●		●							●	●		●	●	●
Policy Analysis Workstream 3 Leader	Dr Rachel Sabates-Wheeler	●	●	●	●	●	●	●	●		●			●						●			●	●	●	●
Local Economy Impacts Study Lead Researcher	Professor Edward Taylor		●				●	●												●	●				●	●
Quantitative Impact Evaluation Lead Researcher	Sean O'Leary		●	●			●	●						●												
Qualitative Impact Evaluation Lead Researcher	Ramlatu Attah	●	●	●		●			●	●		●	●	●	●	●							●	●	●	●
M&E Systems Expert	Dr Claire Simon		●	●		●		●	●			●		●									●	●	●	●
Microsimulation Expert	Dr Sebastian Silva-Leander	●	●	●			●	●				●	●	●							●	●		●	●	●
Communications Expert	Vandi Hill							●	●										●					●	●	●
Senior experts																										
Gender expert	Dr Pauline Kamau								●	●	●				●	●								●	●	●
Food security expert	Dr Zoltan Tiba		●					●						●			●							●	●	●
Livelihoods expert	Dr Jeremy Lind		●	●					●	●	●						●	●						●	●	●
Social development expert	Dr Emma Jones		●						●	●			●		●	●								●	●	●
Climate change expert	Dr Katherine Vincent		●								●				●			●							●	●
Quality Assurance Panel																										
External quality assurance	Patrick Nolen		●		●		●	●							●										●	●
Internal quality assurance	Patrick Ward		●		●	●	●	●																●	●	●
Internal quality assurance	Luca Pellerano	●	●	●	●	●	●	●						●										●	●	●

Project manager: Fred Merttens

Our chosen project manager is Fred Merttens. The project manager has overall responsibility for project delivery and reporting to the PILU for all issues concerned with management and delivery of the contract. Fred's key activities will include coordinating the inception phase, providing ongoing project management, and ensuring all elements of the project are delivered on time and to the required specifications, with all QA processes in place.

Fred managed the Phase 1 HSNP evaluation, and so has a firm understanding of the programme and the requirements for its evaluation. As for Phase 1 evaluation – in addition to his project management responsibilities – Fred will continue to play a technical role, providing review of all outputs across each workstream.

Fred is a senior consultant to OPM's Social Policy Programme with a large array of experience from a number of different projects and contexts, where he has performed a variety of managerial, analytical and research roles. He specialises in the evaluation of social protection, social policy, and poverty reduction programmes, including quantitative and qualitative research, survey design and management, and M&E. Currently Fred is managing an evaluation of the Uganda SAGE Programme, in addition to the quantitative component of an evaluation of a health vouchers programme in Mozambique. He has also been involved in a number of other large-scale evaluations of social protection programmes in sub-Saharan Africa and beyond, including the Kenya OVC-CT, the Lesotho Child Grant Programme, and the BOTA conditional CT programme in Kazakhstan. These projects have all involved implementing large quantitative household surveys and mixed-methods research approaches in complex environments.

KNPC: Caroline Riungu

Our chosen KNPC is Caroline Riungu. The KNPC is a full-time role, based in Nairobi. Reporting to and working closely with the TL, Caroline will be the day-to-day contact and coordination point for the evaluation team. Externally, she will support the TL in ongoing effective and responsive engagement with the PILU and NDMA. Internally, she will support the TL to liaise effectively with the workstream leaders in order to coordinate the delivery of all technical project outputs. In addition to supporting the TL, Caroline will support the project manager with various project management functions. She will also provide extensive technical inputs, in particular for Workstream 2 (Operational M&E). The coordinator will not just be responsible for producing the bi-monthly operational monitoring reports, but will also be heavily involved in the collection of the data on which these will be based, overseeing the operational monitoring survey fieldwork conducted by RGA and making regular monitoring visits to the four HSNP counties. Caroline will then be responsible for the synthesis of all the ongoing operational monitoring evidence, which will form part of the endline operational evaluation report. She will also provide technical inputs to Workstream 4 (Communications and learning), supporting the TL and communications expert with the communications and learning strategies, drafting synthesis reports and summaries, delivering learning events and reviewing the programme's own M&E processes. Finally, Caroline will be on hand to provide day-to-day oversight of the Workstream 1 IE and LEWIE surveys and qualitative fieldwork implementation.

Caroline's profile is ideally suited to this role that combines practical coordination support with technical M&E inputs. She is an M&E specialist with six years of experience in this field, predominantly in Kenya. She also has a wide array of project coordination skills and experience, and an extensive knowledge of quality control, data validation, statistical methods and IT applications for projects. Caroline worked for Save the Children as M&E officer in Wajir South, where she supported the development and implementation of clear, practical M&E plans. She supported the team collating and analysing monitoring data and reporting on the findings. She provided technical support on qualitative and quantitative data analysis and coordinated and shared learning related to M&E practices across offices. As M&E manager and later regional M&E technical advisor at Education Development Centre Inc, Caroline designed and implemented an M&E system and tracked the quarterly progress of programme activities. She was responsible for training data collectors and building the capacity of M&E staff. She ensured compliance with United States Agency for International Development procedures and agreements and prepared regular communications for stakeholders. Caroline has excellent verbal and written communication skills, is efficient, organised and detail-orientated. She is

currently pursuing a Master's in Research at the University of London International Programmes and is an MBA graduate.

IE workstream leader (Workstream 1) and targeting expert: Alex Hurrell

The IE workstream is led by Alex Hurrell, who was the original project manager for the Phase 1 evaluation. He will be responsible for ensuring delivery of the LEWIE and the mixed-methods IE over two rounds (midline and endline). In addition, as the lead author of the Phase 1 targeting analysis report, Alex will deliver the targeting analysis for HSNP Phase 2. Reporting to the Workstream 3 leader in this function, he will be responsible for delivering a targeting study that will assess the effectiveness of the registration process and selection criteria in identifying the poorest households for the programme.

Alex Hurrell is a senior consultant who leads OPM's M&E Portfolio. He is an economist specialising in M&E, poverty analysis and targeting, social protection policy, and quantitative survey design and analysis. He has extensive specific experience in IE and targeting analysis. In addition to his role on the HSNP pilot study he was lead researcher for the targeting analysis that was conducted as part of OPM's evaluation of the Kenya CT-OVC programme. More broadly, Alex has led or been involved in IEs in Kazakhstan (BOTA conditional CT), Kenya (HSNP; OVC-CT), Lesotho (Child Grant Programme), Pakistan (Lady Health Worker Programme), South Africa (Child Support Grant), Uganda (SAGE), and Zambia (Fertiliser Subsidy Programme).

Operational M&E workstream leader (Workstream 2): Clare O'Brien

Our chosen operational M&E workstream leader is Clare O'Brien. Reporting to the TL, the operational M&E workstream leader is responsible for leading the process and institutional evaluation assessments of the HSNP and ASP programmes, a costing study, ongoing operational monitoring of the programme performance, and an endline operational evaluation report that synthesises the findings of all these activities. In addition she will support the TL and KNPC to facilitate the various learning events relating to operational M&E outputs.

Clare O'Brien is a senior consultant in the Poverty and Social Protection team at OPM. Her recent assignments have focused on analytical studies of CT programmes – such as feasibility studies, capacity assessments, impact and operational evaluation and costing analyses (including in Kenya, Malawi, Benin and Kazakhstan). She has an in-depth understanding of the HSNP and of the social protection policy context in Kenya more broadly, as she is currently a social protection expert for the World Bank's functional review of agencies implementing Kenya's NSNP. In 2013–14 she managed a study of the cost-efficiency of emergency CT interventions by NGOs in Kenya and Somalia. Other work in Kenya includes the costing study of the CT-OVC programme in 2010 and support to the mid-term review of the Education for All Fast Track Initiative in 2009.

Policy analysis workstream leader (Workstream 3): Dr Rachel Sabates-Wheeler

Our chosen policy analysis workstream leader is Dr Rachel Sabates- Wheeler. Reporting to the TL, she will be responsible for delivering micro-simulation analyses, an assessment of the effectiveness of the programme's targeting process and a strategic review of the programme. In addition she will support the TL and KNPC to facilitate the various learning events relating to these policy analysis outputs.

Dr Rachel Sabates-Wheeler is a development economist with extensive experience in rural development, social policy and social protection. Rachel has been a research fellow at the Institute of Development Studies (IDS) since 2001, and was a director of the Centre for Social Protection between 2006 and 2012. At IDS she has worked on areas of poverty analysis, social protection, food security and migration in over 12 African countries. She has published on the topics of rural institutions, social protection in Africa, migration and poverty, and graduation and targeting in social protection programmes. Rachel has led and been involved in a number of studies that explore understandings of risk and vulnerability both conceptually and empirically. These studies include: the PSNP, Ethiopia; the HSNP, Kenya; the Child Support Grant, South Africa; and a number of studies on home-grown school feeding programmes in Africa. Recently,

Rachel took a 2.5 year leave of absence from IDS to work with UNICEF in Rwanda as chief of social policy and research. She led a range of operational research on poverty and vulnerability, and she used this work to influence policy and programme design. Rachel co-chaired the social protection sector committee on M&E and she provided QA on a range of national policy documents. Her work also covered areas of targeting and graduation policy analysis, along with work on child-sensitive social protection and early childhood development.

LEWIE study lead researcher: Professor Edward Taylor

Our chosen LEWIE Study Lead Researcher is Professor Edward Taylor, an internationally renowned expert in this field. Reporting to the Workstream 1 leader, he will provide a technical lead on this element of the IE. Edward is professor of agricultural and resource economics at the University of California, Davis, where he teaches courses on international development economics and econometric methods. He is also co-editor of the *American Journal of Agricultural Economics* and founder of the alternative textbook initiative RebelText.org. Edward has written extensively on the economy-wide impacts of development projects and policies. He carried out the first social accounting matrix (SAM) multiplier analysis of a village economy (1988), the first computable general equilibrium modelling of village (2006) and rural (2005) economies, and the first use of SAM and general equilibrium analysis of project impacts (2014). He directs the local economy component of FAO's Protection to Production project in seven African countries. His forthcoming book *Beyond Experiments in Development Economics: Local Economy-Wide Impact Evaluation* (Oxford University Press) presents the methodology for carrying out general equilibrium micro-simulations to evaluate the impacts of policies, projects, and market shocks in local economies. He has published more than 75 publications involving the use of quantitative data analysis in development economics and co-authored the book *RebelText: Essentials of Econometrics*. He is listed in *Who's Who in Economics* and has advised a number of foreign governments and international development agencies on matters related to social development programmes.

Quantitative IE lead researcher: Sean O'Leary

Our chosen quantitative IE lead researcher is Sean O'Leary. Reporting to the Workstream 1 leader and working closely with our fieldwork research company (RGA), he will lead the design, implementation and analysis of the quantitative IE survey. Sean is an IE expert, with particularly relevant experience in the evaluation of CT programmes using quasi-experimental techniques. Sean is a senior consultant in the Poverty and Social Protection team of OPM as well as leading OPM's cross-cutting portfolio. Sean specialises in economic and econometric analysis of policies aimed at alleviating poverty. He holds an MSc in Economics for Development from the University of Oxford and a BSc in Economics from University College London. Recently Sean has designed and managed a range of experimental and quasi-experimental evaluations in Pakistan involving large-scale surveys, including the Benazir Income Support Programme; the Citizen's Damage Compensation Programme; and the Education Fund for Sindh. Elsewhere Sean is leading the quantitative IE component of a cash transfer aimed at alleviating chronic poverty in Uganda (SAGE), as well as an assessment of alternative payment mechanisms for the delivery of the social CT in Malawi.

Qualitative IE leader researcher: Ramlatu Attah

Our chosen qualitative IE lead researcher is Ramlatu Attah. Reporting to the Workstream 1 leader and working closely with our fieldwork research company (RGA), she will lead the design, implementation and analysis of the qualitative IE research. Ramlatu was extensively involved in the Phase 1 HSNP evaluation, where she worked on the design, implementation and analysis for all three rounds of the qualitative evaluation. Ramlatu Attah is a consultant in OPM's Poverty and Social Protection portfolio. Ramlatu specialises in qualitative and participatory research on social protection programmes, and has led a number of qualitative evaluations of CT programmes. In Zambia, Ramlatu led the qualitative component of the assessment of targeting mechanisms of the Social Cash Transfer programme. In Ghana, she was the lead country researcher for qualitative and participatory research on the economic impacts of the Livelihood Empowerment Against Poverty (LEAP) programme, as part of a six-country case study funded by FAO. She was also the lead country researcher on an FAO study in northern Ghana to look at synergies between the LEAP and other social protection programmes. She is also currently working on the evaluation of the Uganda SAGE programme.

M&E systems expert: Dr Claire Simon

Our chosen M&E systems expert is Dr Claire Simon. Reporting to the Workstream 2 Leader and supported by the KNPC, she will lead the design of the data collection, analysis and reporting processes for monitoring the operational performance of the programme. Claire will also provide inputs to Workstream 4 (Communications and learning), to support the ongoing review and implementation of the PILU's M&E framework for the programme. Claire has over 10 years of experience in quantitative and qualitative evaluation research that includes formulating programme strategy, training and managing survey and research teams, analysing data, assessing outcomes, extracting lessons learned, and reporting to donors. She has worked extensively on multiple social protections programmes in Kenya and across sub-Saharan Africa. Her experience is diverse and includes both the design and the review of several large country-wide conditional CT programmes. In this capacity she has: a) designed and implemented closed-ended household questionnaires, semi-structured KIIs, and open-ended focus groups formats to identify potential gaps in programme delivery and triangulate data on beneficiary perceptions; b) developed the a logical framework and indicator set for monitoring programme processes such as targeting, registration, payments and exit; c) created programme operations manuals and training materials; and d) worked closely with teams to build capacity within government social protection/conditional CT implementation units. Before turning to the field of development assistance, Claire spent eight years working in software development, building client relationship management systems for large financial institutions and online retail organisations. In this capacity she specialised in gathering user requirements and transforming these requirements into technological solutions to streamline and improve business processes. She holds a PhD in Geography from the University of Colorado and an MA in International Economics and Finance from Brandeis University.

Micro-simulation expert: Dr Sebastian Silva-Leander

The micro-simulation analysis will be undertaken by Sebastian Silva-Leander, who performed this role toward the end of the Phase 1 evaluation. Reporting to the Workstream 3 leader, he will undertake micro-simulation analyses that will assess the potential impact of varying the transfer value, coverage rates and/or eligibility criteria in the future. As set out in Section 5.2.4, we do not envisage a specific report as such for the micro-simulation work, with this rather responding to the needs of the PILU and NDMA and providing outputs in an appropriate agreed form (which could be a report, a series of Excel files, etc.). Sebastian is a senior consultant in OPM's Poverty and Social Protection team and a research associate with the Oxford Poverty and Human Development Initiative. His main research is in poverty analysis and social protection, with special interest in fragile states. Sebastian has extensive experience in measuring poverty and inequality both in academic research, as well as in applied research related to IEs of public policies and poverty reduction strategies. Sebastian's applied work has focused mostly on the design and implementation of social protection programmes in sub-Saharan Africa, including the use of CT in humanitarian emergencies and the transition from humanitarian aid to nationally led social protection systems. Before joining OPM, Sebastian worked for seven years for the United Nations as an economist and adviser on strategic planning issues in New York, Rwanda, Afghanistan and the Democratic Republic of Congo. He holds a PhD in Economics from St Anthony's College, Oxford, and did a postdoctoral fellowship at Harvard University.

Communications expert: Vandi Hill

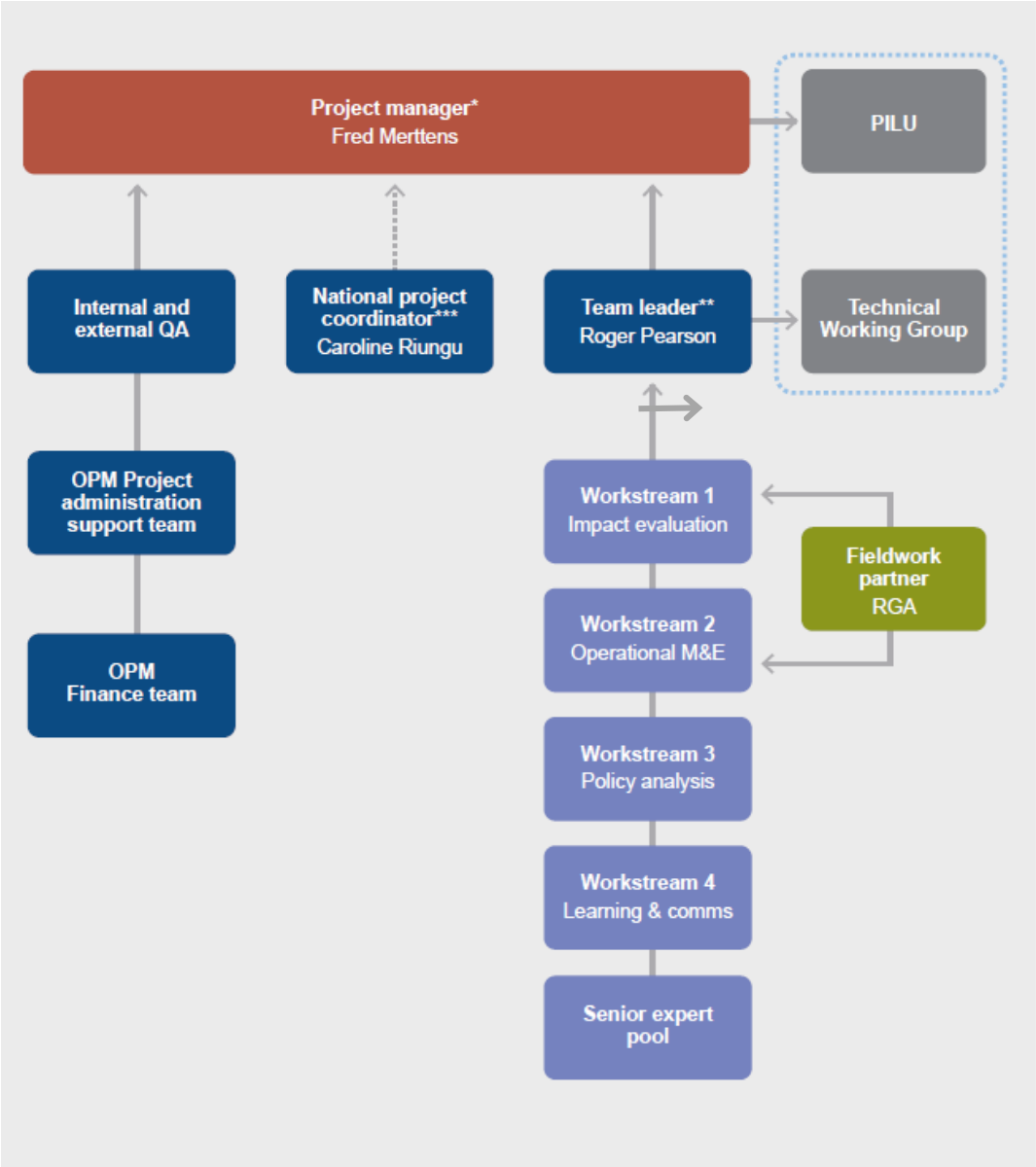
Reporting to the Workstream 4 leader (Roger Pearson), the communications expert will be responsible for design and implementing a comprehensive communications strategy to ensure the evaluation findings are effectively disseminated to all necessary stakeholders. In addition we are proposing that the communications expert will support the PILU with the programme's own broader communications strategy. Vandi Hill is a marketing and communications professional, experienced in developing and implementing integrated communications strategies and plans to meet stakeholder expectations, with a focus on behavioural change, training and information sharing. She is currently a consultant and co-founder at Ricochet Research and a marketing and communications consultant for Hill Consulting, where her regional focus is sub-Saharan Africa. Vandi is accustomed to leveraging the appropriate communication channels such as

internet, telephone, radio, TV and print in a creative and meaningful way for the Africa region. Vandi is consulting as director of communications at TradeMark East Africa on its communications strategy. She has also worked with DFID on its Accountability in Tanzania programme, where she supported the development of a two-year communications strategy. The strategy included building a website, Newsletter, style guide and media policies document for the programme as well as communicating results from the learning workstream to other stakeholders who would implement these findings to achieve greater success in their advocacy work. Vandi holds a BPE from McMaster University, Canada and an MBA from York University, Toronto.

7.2 EVALUATION MANAGEMENT AND GOVERNANCE

The project will be delivered by a team of experts who will have individual responsibilities for particular tasks and a collective responsibility for delivering all the project results. Our project delivery structure contains clear internal project management arrangements, as outlined in Figure 6.

Figure 6 Evaluation management and governance



Notes: *The project manager will report to the PILU for management of the contract. **The Team Leader will report to the Technical Working Group for technical outputs. ***The National project co-ordinator will support both the Project manager and the Team leader with their respective management functions.

The project manager will have overall responsibility for project delivery and resource management and will report to the PILU for all issues concerned with management and delivery of the contract. Technical inputs will be managed by the TL, who will be managed by and accountable to the project manager. Each workstream leader will be responsible for managing their workstream, and will report to the TL. Workstreams 1 and 2 require fieldwork, which will be implemented by RGA; RGA will report to the workstream leaders. The KNPC will report to the TL, providing coordination support and technical inputs across all workstreams, but will also support the project manager with various management functions. The scope of responsibility for each position above is outlined below:

The TL is responsible for all technical inputs on the project. He will ensure that the consortium provides the necessary resources for project delivery – senior expertise, management, lessons from other projects, etc. – and that they are used efficiently. He will have final technical oversight and quality control of all technical outputs, drawing upon the senior expert pool as needed. All workstream leaders will report to the TL. He will also act as the direct liaison for the PILU and Technical Working Group on all technical issues, working with the PILU on a day-to-day basis in his role as leader of the Communications and learning workstream, with support from the KNPC.

The project manager holds responsibility for the delivery of the project on time and within budget to the highest possible standard. He will manage all non-technical aspects of the project and support the TL to deliver the technical outputs. He will be responsive to any needs to develop project management procedures and systems, and will provide necessary related training and support to staff. He will identify and coordinate all project backstopping support, including technical quality control from both internal and external expertise as well as financial quality control and management.

The KNPC will report to the TL and will be the day-to-day contact and coordination point for the evaluation team. Externally, the coordinator will support the TL in ongoing effective and responsive engagement with the PILU and NDMA. Internally, the coordinator will support the TL to liaise effectively with the workstream leaders to coordinate the effective delivery of all technical project outputs. The coordinator will support the project manager with various project management functions, including preparation of the project's progress and financial reports. In her roles within the operational M&E workstream and the Communications and learning workstream she will work directly alongside the PILU on a day-to-day basis, supporting it both to interpret the data produced by the evaluation and utilise it in ways appropriate to the needs of the HSNP. The coordinator will both conduct and supervise much of the ongoing fieldwork planned as part of this evaluation.

Workstream leaders – The four workstream leaders will deliver the core project components and will report to the TL.

Senior expert pool – The senior expert pool comprises experts on relevant thematic areas who will provide the TL and project manager with support reviewing the project technical outputs as well as producing bespoke outputs as needed. The thematic areas already covered include social development, food security, livelihoods, gender and climate change. Additionally, with its extensive network of associates and trusted consultants and academics across multiple sectors and disciplines, OPM is well-placed to bring in any further expertise as required.

QA – Alongside the financial and project administration support teams, OPM has compiled a senior technical QA panel comprised of both internal OPM staff and external expertise in the field of quantitative evaluation methods: Patrick Nolen (external evaluation expert); Patrick Ward (director of OPM's Statistics, Evidence and Accountability Programme and previously project director for the Phase 1 HSNP evaluation); and Luca Pellerano (leader of OPM's Poverty and Social Protection team). CVs for the QA panel were presented in the Annexure of the technical proposal.

7.2.1 Presence in Kenya

Managing a large and complex evaluation will require a strong presence in Kenya. OPM and RGA have an established history of delivering high quality projects in Kenya, including large-scale surveys for evaluations of the HSNP and CT-OVC

programme. Drawing on lessons learned from these and other experiences, we will ensure the successful implementation of this project by: a) providing necessary support to enable our TL, who is based in Nairobi, to steer the project on the basis of ongoing in-country dialogue with the PLIU; b) mobilising a full-time KNPC who will also be based in Nairobi to support the TL, project manager and workstream leaders; and c) working with an established national survey firm with extensive in-country networks and a detailed understanding of the region in which fieldwork takes place.

TL and KNPC

Our TL, Roger Pearson, will be based full time in Nairobi for the duration of the project. Roger has an excellent understanding of the context in Kenya, having worked in UNICEF's Kenya office for eight years until 2008. The coordinator will also be based full time in Nairobi and will report directly to Roger and Fred Merttens as the project manager. She will work closely with RGA, providing field supervision support to the ongoing fieldwork, and with the other workstream leaders, providing a capable team member for other key project activities.

National survey firm

We believe that the extensive use of well-qualified local consultants is essential for successful survey implementation. We have chosen to work with RGA, which we have worked with successfully in the past on a number of projects, including the M&E component of the HSNP pilot phase and the recent functional review of the Kenya NSNP. RGA is based in Nairobi and brings a detailed understanding of the context and challenges in Kenya, and of northern Kenya in particular. RGA has a large network of experienced researchers with appropriate language skills and a sensitivity toward the cultural context. As far as possible it will deploy fieldworkers with experience from the HSNP Phase 1 evaluation.

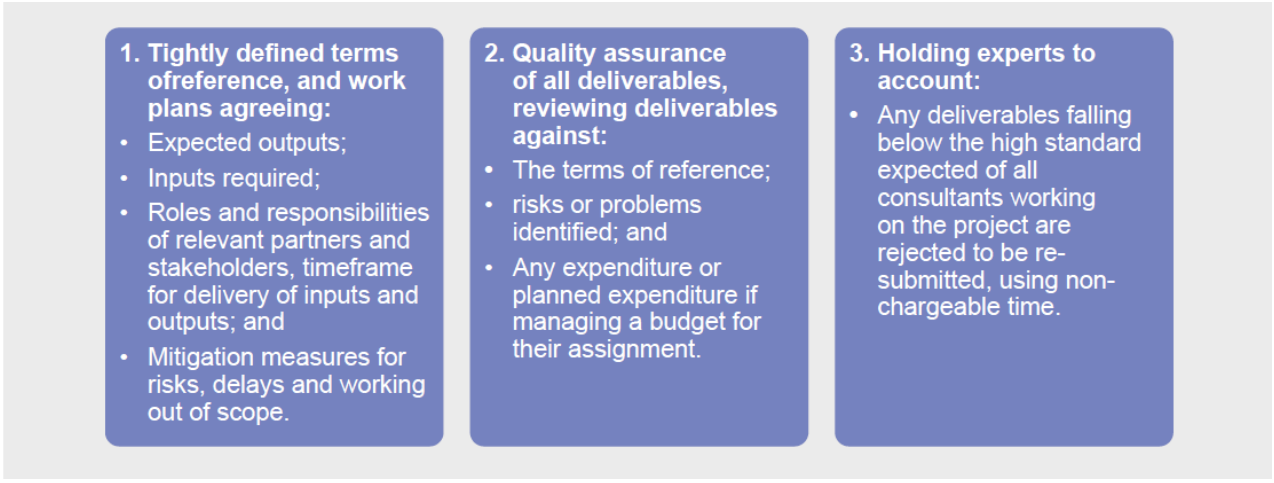
7.2.2 Team management

Ensuring experts' performance through project management

The core evaluation team is comprised of six experienced internal OPM staff members and six external consultants. Having a close relationship with the external experts helps us ensure the quality of their work and retain them for the life of their input. Effective management of their work will be supported by three key measures (see Figure 7):

- tightly defined ToR and work plans;
- QA of all deliverables (see 0 below); and
- holding experts accountable for their deliverables.

Figure 7 Measures to ensure expert performance



Staff retention

Many of our proposed team members are either long-standing, full-time staff members or long-term associates of OPM. This lies at the core of our approach to staff development. OPM has a strong record of retaining key staff within large and complex DFID projects.

We have, however, made further provisions to ensure we can retain their services throughout the life of the programme, by ensuring the following:

- We accommodate expectations of increases in salary in line with inflation over the life of this programme, which we will absorb as additional costs to ourselves (see the Commercial Proposal for further details).
- We will identify, understand and try to meet natural expectations for career progression and personal development by establishing an individual performance management system for every staff member. We will identify personal aspirations, and meet them through the project, by: (i) augmenting roles and responsibilities as the project expands; (ii) providing opportunities for training and international exposure for Kenyans; and (iii) providing opportunities for career development by encouraging internal applications for vacancies arising in the project team.
- Dissatisfaction is addressed: The TL and project manager will seek to anticipate any problems within the team, by staying in regular contact with consultants and encouraging the airing of problems. We operate an 'open door' policy. If any team member is dissatisfied for either a professional or personal reason, we will initiate a discussion in an effort to resolve the problem. Our preference is always to retain team members – and ensure the satisfaction of the team overall. If a problem cannot be resolved, we have procedures for managing personnel changes, as discussed below.

All our team members are fully committed to the project, many of them having been involved in the first phase evaluation and/or other recent projects in social protection in Kenya. The team is aware of the proposed schedule and is available to mobilise on the date required.

Managing changes to the team

For the reasons outlined above, we do not expect there to be significant changes to the team over the course of the project. However, in the event that changes to the team are required, we have established procedures for managing these changes swiftly, cost-effectively and compassionately so that the project is not disrupted. In all cases, DFID and the government counterparts will be fully informed and will have a say in any proposed replacement. Our key procedures include:

- **Inform DFID and the government counterparts immediately** and keep them up to date as circumstances change; and immediately initiate processes for identifying and recruiting appropriate replacements.
- **Recruit internally** – Our preference is to fill vacancies with someone from the existing team who has the necessary skills and ability to deliver the role. This will minimise the disruption to the programme and reduce transaction costs associated with bringing a new team member up to speed.
- **Second option, look more widely, including drawing on our networks** – OPM has extensive networks of international associates to draw upon. The search would be carried out against a clearly defined profile for the role being replaced.
- **Provide DFID and government counterparts with a choice of replacements for their approval** – We will always seek approval for any team changes from DFID and GoK counterparts. Offering a range of CVs can be supplemented by arranging a face-to-face interview if requested. Our preference will always be for a national, provided they are suitably qualified.

7.2.3 QA

OPM is committed to top quality and excellence in evaluations, and seeks to ensure that this e-Pact evaluation has a multi-layered QA system, which will address all dimensions of quality, including evaluation design, process, outputs, and teams and processes. Timeliness is also a quality dimension. The QA system includes the following elements:

- **In-house QA:** All deliverables, including methodologies and evaluation designs and reports, will be quality assured – employing checklists based on the OECD/DAC criteria – by the internal QA panel of Patrick Ward and Luca Pellerano before submission to DFID.
- **External QA panel:** All key deliverables will be quality assured by Patrick Nolen, who has particular expertise in quantitative evaluation methods and experience in providing a similar QA role for a DFID-funded CT programme in a neighbouring East African country.
- **e-Pact-level oversight:** A permanent member of staff (Patrick Ward) from the e-Pact management team will carry out 'certification of the quality process', following the pattern of an International Organization for Standardization system.

Internal QA

Our internal QA process ensures a quality check at each stage of project delivery (start-up, implementation, deliverables, reporting). Integral to this QA process is participation and ownership of the evaluation by various stakeholders (government in particular). We ensure that counterparts are fully incorporated in the review system, enabling the views and priorities to be articulated and integrated into our systems. All draft action plans, strategy documents and reports will be quality assured by the TL and project manager. This will enable appropriate action if the realisation of the outputs is endangered. Set out in Figure 9 is an example of an overview of the review process for the programme, highlighting the various interactions among our programme staff, counterpart institutions, and our quality control and approval processes. During the inception phase this will be developed further in order to identify and agree on counterpart interactions within this.

TABLE 10 EXAMPLE OF REVIEW PROCESS

OUTPUT	PRODUCED BY	REVIEWED BY	APROVED BY
Workplans	Workstream leaders	Team Leader / PILU	Project Manager
Progress reports	Team Leader and Kenya National Project Coordinator	PILU	Project Manager
Key deliverables	Workstream leaders	Team Leader / PILU	Project Manager

External QA

All final deliverables will be quality assured by our external QA expert. In addition, where relevant, project reports and other outputs (such as workshop presentations, briefing notes, etc.) will be reviewed by the senior experts pool. This pool will also be available to provide peer reviewing and QA on a call-down basis for specific intermediary outputs. The TL will access the QA panel through the project manager, who will also draw on OPM's networks of externals to identify any additional expertise needed for a specific QA work as required.

Quality in impact assessments

OPM carries out a wide range of evaluations and programme reviews, including operational, process and impact evaluations, for bilateral and multilateral agencies and for governments. It conducts independent, rigorous and evidence-based evaluations that reflect the Paris Declaration principles of country ownership and harmonisation and

that are in keeping with the OECD-DAC evaluation criteria and draft evaluation quality guidelines. OPM is committed to ensuring that the IEs it carries out are to the highest standard possible. An important component of this is ensuring that we collect good primary data.

Ethical Review Committee (ERC) of OPM

OPM regularly carries out research studies in various parts of the world that collect primary data from human subjects. As a value-driven organisation OPM is always respectful of the rights of the participants of its research projects and has a policy to ensure complete adherence to research ethics. In 2013 the management team of OPM approved the establishment of an internal independent ERC. The overall aim of the committee is to ensure that all OPM research activities are carried out to the highest ethical standard. The ERC has a Chair, a co-chair, a coordinator, an external member from the Ethox Centre of the University of Oxford, three members and a secretary. The committee is currently chaired by Professor Hugh Annett, who is a public health specialist and was a member of the Appeal Committee of the National Institute for Health and Clinical Excellence. A multidisciplinary team including a medical demographer, clinician, public health specialist, sociologist, communication expert and economists are also included in the committee. Studies/surveys that secure funding and involve primary data collection from human participants go through the ERC approval process. The study protocols can go through a full-board review, an expedited review or an exemption, depending on their level of risk to human subjects. After review the committee may agree on one of the four outcomes: accepted, resubmit with minor modifications, resubmit with major modifications, or declined.

Assessing specific risks in Kenya

We have conducted a provisional risk assessment based on our experience on HSNP Phase 1 and our understanding of the ToR. During the design phase we will review these risks and, where appropriate include additional risks. Currently we identify the following risks:

- **Insecurity:** intertribal conflict in Turkana and Marsabit; Al-Shabaab in Mandera and Wajir; bandits; general lawlessness.
- **Logistical challenges:** inaccessible areas; long distances between fieldwork sites; limited or no ability to communicate with survey teams in many fieldwork sites; electricity for charging computer-assisted personal interviewing (CAPI)-fieldwork devices; limited internet connectivity for timely transmission of CAPI-collected survey data.
- **Capacity constraints in HSNP programme management:** problems with coordination and collaboration between implementing agencies; lack of understanding and ownership of evaluation approach and evidence; lack of coordination with implementing agencies to facilitate robust IE design (e.g. staggered programme roll-out).
- **Risk of drought/emergency:** this may disrupt survey fieldwork (for example, households may migrate, and survey participation may be difficult to secure); emergency response may disrupt fieldwork logistics and has the potential to disrupt operation of programme.

DFID's DoC risk assessment rates this project as a high risk (4). The project has a significantly greater than normal risk, with risks of terrorism (rated 5) and violence/crime, civil unrest, security, and transport (rated 4) presenting particular challenges for the project. Our budget has factored in the costs associated with ensuring the security of our project staff (see 0 for summary of the evaluation security plan).

7.3 EVALUATION BUDGET

The table below gives a summary breakdown of the budget provided with the original proposal. Alongside this it presents the costs of the evaluation plan articulated in this inception report and the difference between the two.

TABLE 11 BUDGET COMPARISON BETWEEN PROPOSAL AND CURRENT EVALUATION PLAN

BUDGET ITEM	PROPOSAL COST	UPDATED COST	DIFFERENCE
Inception phase			
Fees	£59,480	£59,480	£0
Reimbursables	£22,465	£22,479	£14
Total	*£81,945*	£81,959	£14
Implementation phase			
Fees	£703,060	£795,910	£92,850
Quantitative IE survey	£418,113	£472,011	£53,898
Qualitative fieldwork	£117,000	£165,112	£48,112
Operational monitoring survey	£154,440	£159,343	£4,903
Process and capacity review	£1,536	£1,536	£0
Reimbursables	£147,730	£151,964	£4,234
Total	*£1,541,878*	*£1,745,876*	*£203,998*
TOTAL	£1,623,822	£1,827,836	£204,014

Notes: Totals may not sum exactly due to rounding to the nearest £.

The additional cost of the current evaluation plan compared to the plan articulated in the original proposal is £204,014. Almost exactly half of this cost is accounted for by increases to the primary data collection stemming from revisions to the IE design.

In terms of the quantitative survey, the increase in costs derives from the increase in sample size and the combining of the two rounds of panel survey into a single, larger cross-sectional survey. This is required in order for the evaluation to provide estimates of impact across different population groups as requested by DFID (i.e. heterogeneity analysis; see Section 0 above).

By way of an alternative comparison, we can compare the current cost of the quantitative survey –which is designed so as to meet the objectives of the LEWIE (see Section 0), the IE (Section 0) and the micro-simulation analyses to be carried out under Workstream 4 (Section 0)– with what it might be if we sacrificed the IE component and just structured the sample in order to deliver the LEWIE and micro-simulation analysis. For these purposes we would require around 3,000 households for a single, cross-sectional survey. This sample size is driven by the minimum sample required for LEWIE design, which aims to break down the multiplier across the three cluster types that build the model of the local economy (town, village and remote cluster), across four income groups, and as well between beneficiaries and non-beneficiaries. Such a sample size would provide a representative sample of the whole population and would also thereby satisfy the needs of the micro-simulations.

The total cost of such a single-round survey of 3,000 households would be roughly £242,000. Compare this to the cost of the current proposed single-round survey of approx. 6,500 households (approx. £472,000), and we see the survey cost of producing some kind of quantitative estimate of impact is approx. £230,000.

The other additional costs to the current evaluation plan result from a combination of the fees and reimbursable expenses required to deliver the additional analyses requested by DFID during the inception phase (including the study on the emergency payments, the special study on the ASP, and the review of the HSNP registration instrument), as well as the additional analysis required around the revised IE approach (namely the heterogeneity analysis).

References / bibliography

FAO (2013) *The economic impacts of the CT-OVC programme in Kenya*.

Fitzgibbon, C. (2014) *HSNP Phase II Registration and Targeting Lessons Learned and Recommendations*, July.

Njagi, R. Bori (2014) *Sublocation HSNP Registration Completion Report*, NDMA, November.

OPM (2013) *Kenya Hunger Safety Net Programme Monitoring and Evaluation Component Analysis for HSNP Phase II business case*, April.

Taylor and Filipski (2014). *Beyond Experiments in Development Economics Local Economy-wide Impact Evaluation,* Oxford: OUP.

White and Ellison (2006). *Wellbeing, Livelihoods and Resources in Social Practice*, WeD Working Paper 23, Economic and Social Research Council, October.

Glossary

Instrumental Variable (IV) – IV regression is an estimation technique that is employed to address bias caused by omitted variables (unobserved variables that cannot be included in the regression), simultaneous causality (explanatory variable affects the outcome, but the outcome also affects the explanatory variable), and bias derived from measurement errors.

Local Economy-Wide Impact Evaluation (LEWIE) – LEWIE is a methodology designed to understand the full impact of CTs on local economies, including on the production activities of both beneficiary and non-beneficiary groups of observations, how these effects change when programmes are scaled up to larger regions, and why these effects happen.

Monte Carlo Method – A Monte Carlo method is a simulation technique used to understand the impact of risk and uncertainty in financial, project management and forecasting models. The Monte Carlo simulation performs risk analysis by building models of possible results through substituting a probability distribution for a range of possible outcomes in any choice of action.

Propensity Score Matching (PSM) – PSM is a statistical matching technique used as a quasi-experimental design in the statistical analysis of observation data. PSM attempts to estimate the effect of a treatment (i.e. policy, programme or other intervention) by accounting for the observable characteristics of the unit observation that predict receiving the treatment.

Regression Discontinuity (RD) – RD is used as a quasi-experimental design that attempts to estimate the effect of a treatment (i.e. policy, programme or other intervention) when the treatment is determined by a cut-off point on an observed assignment variable. Observations just below the cut-off are seen as good comparisons to observations just above the cut-off.

Annex A Quantitative IE technical annex

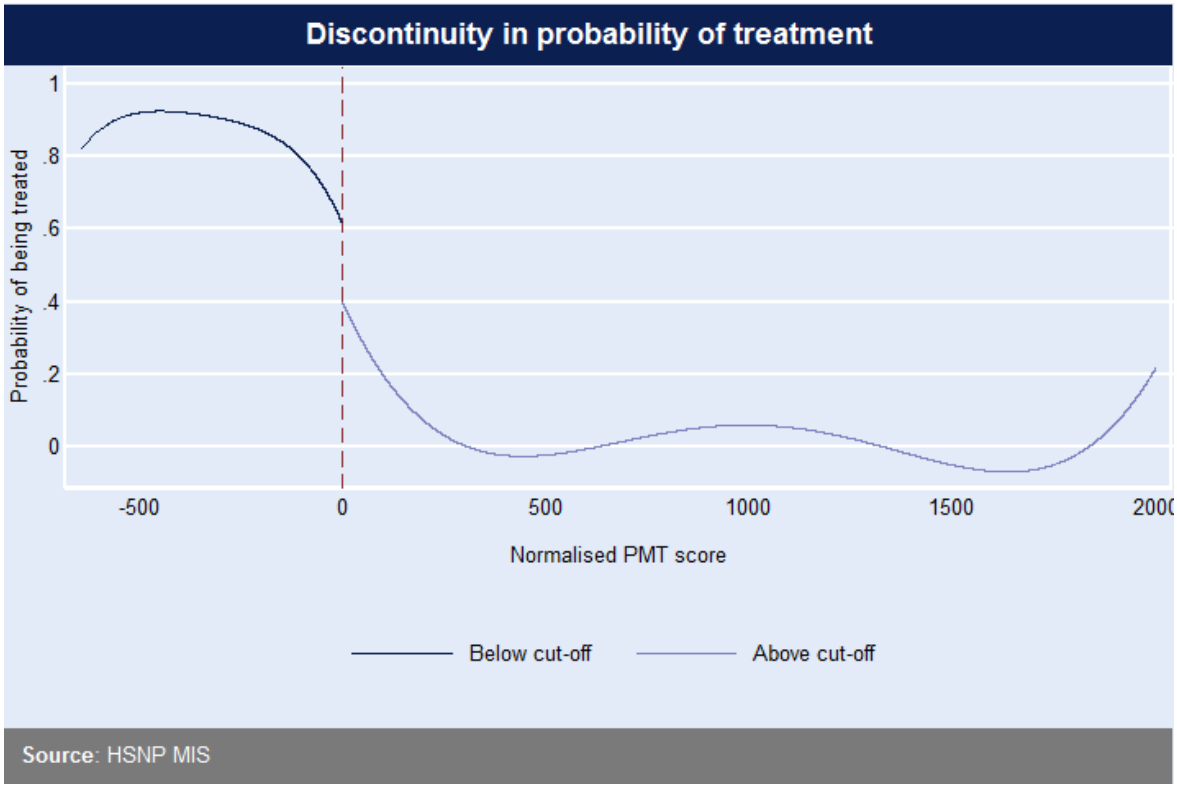
A.1 RRD DESIGN

A.1.1 Internal validity tests

Three validity tests were performed to determine whether the key assumptions underpinning the RD design are satisfied at the county level and the design is therefore internally valid:

- Test of the continuity of the density of the assignment variable: This test produces a positive result as there does not appear to be a discontinuity in the density of the assignment variable around the PMT cut-off. This means that there is no sign of manipulation of the PMT eligibility score by treatment households, proving that counties are in fact the level at which the PMT score threshold is set.
- Test of the discontinuity in the probability of treatment at the cut-off: We find a statistically significant but relatively small (0.2185) discontinuity in the probability of treatment, based on the normalised PMT score threshold set at the county level. This small discontinuity is driven by the proportion of households with PMT scores above the county PMT cut-off that become HSNP beneficiaries and vice versa, which is due to the combined targeting at the village level. Figure A.1.1 presents the discontinuity around the normalised county-level PMT cut-off. Apart from showing a relatively small discontinuity, the graph also demonstrates that the probability of receiving the HSNP transfer is different from zero (0%) for some observations above the normalised cut-off (cross-over households), and is different from one (100%) for some observations with PMT scores below the cut-off (no-show households).

Figure A.1.1 Probability of the discontinuity of treatment



- Test of the continuity of observable covariates at the cut-off: this test determines whether or not the condition that observable covariates must be a continuous function of the assignment score at the eligibility threshold is satisfied. There appears to be a significant number of discontinuities in observables part from a few exceptions, the discontinuities are found across variables associated with household demographics, assets and food consumption. Table A.1.1 below contains the list of variables showing these discontinuities at the county-level cut-off.

TABLE A.1.1 CONTINUITY OF OBSERVABLES AT THE COUNTY LEVEL WITH PMT CUT-OFF

DEFINITION OF VARIABLE	NAME OF VARIABLE	COEFFICIENT	P-VALUE
Household Size	*hhsize*	0.435492***	0
No. of Children under 5	*children_under5*	-0.00313	0.826413
No. of Elderly over 60	*elderly_over60*	0.014335**	0.020409
DR	*dependratio*	-0.07647***	0.000137
No. of Children in School	*attendschool*	0.309921***	0
No. of People Working	*working*	-0.03777*	0.050316
Female Household Head	*FemaleHHead*	-0.01924***	0.006141
No. of Female Members	*femalehhmembers*	0.249661***	0
Proportion of Female Members in Household	*prop_femalemembers*	-0.00622**	0.045275
Proportion of Children in Household	*prop_children*	-0.02849***	0
Proportion of People Working in Household	*prop_working*	-0.03829***	0
Proportion of Elderly in Household	*prop_elderly*	0.01066***	0
Wall Is Made of Grass	*wall_grass*	0.003545	0.575924
Wall Is Made of Iron	*wall_iron*	0.033436***	0
Wall Is Made of Mud	*wall_mud*	-0.04472***	0
Roof Is Made of Grass	*roof_grass*	-0.03839***	0
Roof Is Made of Asbestos	*roof_asbestos*	0.039349***	0
Floor Is Made of Earth	*floor_earth*	-0.04508***	0
Floor Is Made of Other Materials	*floor_other*	0.013085**	0.024113
Floor Is Made of Cement	*floor_cement*	0.013088***	0
No. of Rooms	*no_main_rooms*	0.032635***	0.000239
Land Cultivated in Acres	*land_cultivated_1yr_acres*	0.029001**	0.034232
The Household Owns:			
Sofa	*sofa_set_owned*	0.005569***	0.002903
Bed	*bed_owned*	0.06988***	0

Refrigerator	*refrigerator_owned*	0.000394	0.601999
Charcoal Jiko (Stove)	*charcoal_jiko_owned*	-0.00114	0.666298
Kerosene Stove	*kerosene_stove_owned*	-0.00036	0.735619
Electric Iron	*electric_iron_owned*	-0.00054	0.361308
Charcoal Iron	*charcoal_iron_owned*	-0.0013	0.236849
Paraffin Lamp	*paraffin_lamp_owned*	0.022703***	0
Frying Pan	*frying_pan_owned*	0.041537***	0
Mattress	*mattress_owned*	0.044486***	0
Towels	*towels_owned*	-0.00467***	0.00366
Mosquito Net	*mosquito_net_owned*	-0.09149***	0
Mobile Phone	*cellphone_owned*	0.002638	0.602951
Animal Cart	*animal_cart_owned*	0.00482*	0.05293
Motorcycle	*motorcycle_owned*	-0.00073	0.454321
Bicycle	*bicycle_owned*	-0.00401***	0.002081
Radio	*radio_owned*	-0.00489*	0.079968
TTV	*tv_owned*	-0.00083	0.500263
Zebu Cattle	*zebu_cattle_owned*	-0.23937***	0.00085
Exotic Cattle	*exotic_cattle_owned*	-0.03261	0.321355
Shoats (Young Pigs)	*shoats_owned*	-5.23395***	0
Camels	*camels_owned*	0.077819	0.257258
Donkeys	*donkeys_owned*	-0.19786***	0
Nets	*nets_owned*	0.008243	0.136674
Lines	*lines_owned*	0.007275	0.119523
Hooks	*hooks_owned*	0.009344**	0.030858

In the last week, the household ate[32]:

[32] The number of households is considerably reduced for the food consumption variables, as there are 208,313 missing observations, which correspond to 57.41% of the total number of observations in the MIS dataset used for the analysis.

A.2 PSM

In light of the issues related to the quality of the data at our disposal and the methodological concerns affecting the proposed county-level RD design, we believe that it is advisable to envisage the possible design and development of an alternative evaluation approach based on propensity score analysis. The PSM approach relies on a matching procedure that attempts to avoid the bias that emerges when beneficiary households differ from non-beneficiary households in a series of defining characteristics that could affect the outcome of interest and thus confound any estimate of the programme impact. The PSM approach attempts to find control units that are as similar as possible to the treatment units (i.e. households receiving the HSNP transfer in our case), so that any measurable difference in the outcome indicators can be confidently attributed to the intervention.

The proposed PSM represents a back-up approach that will be based on a propensity score model constructed at the household level, which would therefore allow us to avoid the analytical level of the village that is affected by data quality issues. In particular, the PSM approach relies on a first-stage PSM that estimates each observation's probability of receiving the intervention and is meant to achieve a balance across treatments' and controls' characteristics at baseline. A comparison can then be made between treatment and control households with similar 'propensity scores'. Under the assumption that the true likelihood of receiving the intervention is adequately captured by the chosen observable characteristics, this method overcomes the selection bias in an attempt to isolate the effects attributable to the HSNP intervention. A necessary condition for PSM to be valid is the existence of a sufficient overlap in the observable characteristics of treatment and control units. A provisional analysis of the MIS data was undertaken to establish whether this condition is likely to be met for our beneficiary and non-beneficiary households.

This preliminary PSM analysis was carried out in two stages. First, the propensity score for each household was calculated (the probability of exposed to the intervention) on the basis of a probit regression, with treatment status as the dependent variable and the matching variables as covariates. Secondly, each treatment household was matched with one control group household that has the most similar propensity score to its own (control households are not allowed to be replaced). Since 27% of the MIS sample was assigned to the intervention, there were sufficient control group households to ensure that all were matched. For the purposes of this preliminary PSM feasibility investigation we adopted the simple rule that treatment observations are dropped if their propensity scores are above the maximum propensity score observed in the control group, or lower than the minimum.

The outcome of this preliminary investigation indicates that there is in fact a case to apply PSM techniques in this setting, as the imbalance between treatment and control groups is considerably reduced after the two groups of households are matched. There is however one specific consideration: the consumption module of the MIS data was only available for less than half of the respondents. As consumption data are associated with socioeconomic status of households, the inability to make use of this information for over half of the sample was a shortcoming of our preliminary analysis. In view of this issue, we explored two options. The first was to estimate a matching model over the full population covered by the MIS data, and not include any consumption information in the matching procedure. This ensures that the model is most representative of the population. The second option was to estimate the model over a restricted sample containing only the households for which consumption data were collected. The two tables containing the results for the matched and unmatched samples obtained with these two options are presented below.

A.2.1 PSM for the entire MIS dataset

Our first PSM model included variables that could be collected for the full sample contained in the MIS data, to ensure maximum representativeness. The matching variables selected did not include consumption data, but instead a range of other variables expected to have an influence on the probability of receiving the HSNP transfer. The variables were purposefully selected to capture a range of dimensions of household socioeconomic status. Accordingly, a combination

of household-level and (constructed) individual-level variables were represented, including dummy variables for each of the counties that make up the sample (with one omitted category).[33]

The PSM can be thought of as having been effective if the matched treatment and control group households have similar average observable characteristics. As explained above, this indicates that the selection bias has been effectively removed, and that a comparison between the two groups will reliably isolate the impact of the intervention (assuming that unobservable differences do not persist). The results of our preliminary PSM balance diagnostic are reported in Table A.2.1 below. In all cases, the difference in mean outcomes between treatment and control households in the unmatched sample is highly significant. The results on Rubin's B,[34] which is an overall indicator of the similarity of means across the set of matching variables, show a tangible decrease in overall imbalance after matching. However, some significant differences remain across individual variables in the matched sample, though this is thought to be partly driven by the very large number of observations included in the model.

A.2.2 PSM for the MIS dataset with food consumption

We repeated the matching procedure outlined above on a subset of the MIS population that was covered by the consumption module of the questionnaire (43.84% of the observations in the MIS dataset – 157,230 households). As discussed, the inclusion of consumption variables in the PSM routine is expected to increase the relevance of the matching algorithm, though this is achieved at the cost of the representativeness of the matched sample. The benefits gained in terms of improving the strength of the matching routine are shown in Table A.2.1 below. Compared to the PSM conducted over the full sample, there are fewer significant differences between means in the treatment and control groups of the matched sample. Furthermore, the test regarding the overall measure of balance shows a more marked improvement[35] than in the case of the overall MIS dataset.

[33] A routine was adopted to ensure that the choice of variables did not suffer from multicollinearity issues. Collinearity of the selection means that two or more variables are highly correlated, which can distort the significance of the results.

[34] According to accepted standards, a value of below 25 is acceptable for PSM. In this case the value of 10.8 suggests that this set of matching variables could allow for a feasible PSM technique to be performed on this dataset.

[35] The value of Rubin's B has fallen to 6.35.

TABLE A.2.1 BALANCE OF OBSERVABLE CHARACTERISTICS ACHIEVED BY PSM OVER THE FULL SAMPLE

VARIABLE	TREATMENT MEAN UNMATCHED	CONTROL MEAN UNMATCHED	T MEAN MATCHED	C MEAN MATCHED	OBS.	% BIAS UNMATCHED	% BIAS MATCHED	% CHANGE	P-VALUE UNMATCHED	P-VALUE MATCHED
Household has female head	0.52***	0.43	0.52	0.52	3586 49	18.7	0.4	98	0	0.407
Proportion of household that is female	0.5***	0.48	0.5	0.5	3586 49	9	-0.6	93.5	0	0.187
Household has at least one child aged between 5 and 15	0.89***	0.74	0.89**	0.88	3586 49	37	0.9	97.5	0	0.015
Proportion of children aged 5 to 15 attending school	0.43***	0.37	0.43***	0.41	3586 49	14.4	3.5	75.7	0	0
Household has no bed	0.71***	0.51	0.71***	0.71	3586 49	42.9	1.5	96.5	0	0
Walls made from grass	0.33***	0.27	0.33**	0.34	3586 49	12.9	-1	92.3	0	0.031

Household owns mosquito net	0.2***	0.45	0.2**	0.21	3586 49	-53.7	-0.8	98.5	0	0.045
Household owns frying pan	0.11***	0.33	0.11*	0.11	3586 49	-54.6	0.7	98.8	0	0.057
Mean households with walls made from grass in village	0.3***	0.29	0.3***	0.31	3586 49	4.2	-2.6	38.3	0	0
Mean households with female household head in village	0.46***	0.45	0.46	0.47	3586 49	9	-0.6	93.7	0	0.205
Mandera	0.21***	0.22	0.21***	0.19	3586 49	-3.1	4.4	-40.9	0	0
Turkana	0.39***	0.36	0.39***	0.41	3586 49	6.4	-5.8	8.8	0	0
Wajir	0.21***	0.29	0.21***	0.22	3586 49	-18.1	-3.1	82.8	0	0

Notes: Rubin's B before matching: 91.19. Rubin's B after matching: 10.79. Asterisks denote level of significance: ***=99%; **=95%, *=90%.

TABLE A.2.2 BALANCE OF OBSERVABLE CHARACTERISTICS ACHIEVED BY PSM OVER THE CONSUMPTION DATA SAMPLE

VARIABLE	T MEAN UNMATCHED	C MEAN UNMATCHED	T MEAN MATCHED	C MEAN MATCHED	OBS.	% BIAS UNMATCHED	% BIAS MATCHED	% CHANGE	P-VALUE UNMATCHED	P-VALUE MATCHED
Household has female head	0.57***	0.5	0.57	0.57	10561 5	14.8	0.5	96.5	0	0.502
Proportion of household that is female	0.51***	0.49	0.51	0.51	10561 5	11	0.1	99.1	0	0.893
Proportion of household members older than 5 who were working in last 7 days	0.28***	0.34	0.28***	0.29	10561 5	-22.4	-3.1	86.3	0	0
Ratio between dependents (children under 15 and over 64) and working-age people	1.98***	1.66	1.98	1.98	10561 5	22.2	0.3	98.8	0	0.74
Household has at least one child aged 5–15*, proportion of children aged 5–15 attending school	0.44***	0.39	0.44***	0.45	10561 5	11.5	-2.3	80.2	0	0.004
Household has at least one child aged 5–15	0.87***	0.74	0.87	0.86	10561 5	32.9	0.9	97.2	0	0.169

Meat or fish eaten in last 7 days	0.44***	0.58	0.44	0.43	10 56 15	-29.2	1.2	95.9	0	0.122
Wheat or beans eaten in last 7 days	0.87***	0.92	0.87	0.87	10 56 15	-16.5	0.1	99.5	0	0.925
Vegetables or fruit eaten in last 7 days	0.17***	0.33	0.17	0.17	10 56 15	-37.2	-0.2	99.5	0	0.78
Marsabit County	0.36***	0.26	0.36*	0.36	10 56 15	22.8	1.4	93.7	0	0.08
Turkana County	0.58***	0.71	0.58***	0.59	10 56 15	-27.4	-2.1	92.3	0	0.01
Mandera County	0.06***	0.03	0.06*	0.06	10 56 15	11.8	1.5	87	0	0.077

Notes: Rubin's B before matching: 64.65. Rubin's B after matching: 6.35. Asterisks denote level of significance: ***=99%; **=95%, *=90%.

However, Tables A.2.3, A.2.4 and A.2.5 give an indication of the extent of the loss of representativeness caused by restricting the sample to the sub-population for which consumption information is available. Each table shows mean values for different variables, in the overall sample and disaggregated according to whether consumption information is available. Column five reports the p-values associated with the difference between the means in the consumption and non-consumption sub-groups. The tables show that the differences are always highly significant. This reflects in part the very large number of observations in the MIS data, so it is also instructive to consider the numeric difference in the means between the two groups. Reviewing the tables reveals that there are some variables where the means appear to be quite different; these are highlighted in bold. Overall, it appears that restricting the sample to contain only households included in the consumption module does carry a cost of representativeness.

TABLE A.2.3 MEAN DIFFERENCES ACROSS INDIVIDUAL-LEVEL CHARACTERISTICS

INDICATOR	OVERALL	CONS. DATA	NO CONS. DATA	P-VALUE	TOT N
Female Household Head	0.453	.46***	0.448	0	358649
Average Household Age	19.299	**19.788*** **	**18.918**	0	358649
DR	1.815	1.75***	1.866	0	358649
Proportion of Household Members Working	0.252	.276***	0.233	0	358649
Proportion of Household Female Members	0.487	.488***	0.485	0	358649
Presence of School-Aged Children (5–15)	0.784	.77***	0.794	0	358649
Proportion of Children Attending School	0.385	**.407*** **	**0.368**	0	358649
Wives in Settlement	0.016	.021***	0.011	0	358649

Notes: Overall F-test; F-stat: 811.1322; P-value: 0

TABLE A.2.4 MEAN DIFFERENCES ACROSS HOUSEHOLD-LEVEL CHARACTERISTICS

INDICATOR	OVERALL	CONS. DATA	NO CONS. DATA	P-VALUE	TOT N
Acres of Land Cultivated	0.065	.068*	0.063	0.067	358649
Wall Is Made of Grass	0.29	**.66*** **	**0.001**	0	358649
Roof Is Made of Grass	0.507	**.616*** **	**0.422**	0	358649
Household Has No Bed	0.566	**.664*** **	**0.489**	0	358649
Household Has No Frying Pan	0.73	**.748*** **	**0.717**	0	358649
Household Has No Mosquito Net	0.621	**.593*** **	**0.644**	0	358649
Household Owns:					
Sofa	0.024	.029***	0.021	0	358649
Bed	0.434	**.336*** **	**0.511**	0	358649
Refrigerator	0.003	.004***	0.003	0	358649
Charcoal Jiko (Stove)	0.045	.063***	0.031	0	358649
Kerosene Stove	0.008	.009***	0.007	0	358649
Electric Iron	0.002	.002***	0.003	0	358649
Charcoal Iron	0.007	.009***	0.005	0	358649
Paraffin Lamp	0.122	**.144*** **	**0.105**	0	358649
Frying Pan	0.27	**.252*** **	**0.283**	0	358649
Mattress	0.335	**.296*** **	**0.366**	0	358649
Towels	0.019	.022***	0.016	0	358649
Mosquito Net	0.379	**.407*** **	**0.356**	0	358649
Mobile Phone	0.184	**.189*** **	**0.179**	0	358649
Animal Cart	0.034	.016***	0.048	0	358649
Motorcycle	0.005	.005***	0.006	0	358649

Bicycle	0.009	.012***	0.007	0	358649
Radio	0.048	.05***	0.047	0.001	358649
TV	0.012	.014***	0.01	0	358649
Zebu Cattle	0.812	**1.127*** **	**0.566**	0	358649
Exotic Cattle	0.215	.261***	0.179	0	358649
Shoats (Young Pigs)	14.814	**15.136*** **	**14.562**	0	358649
Camels	1.237	1.021***	1.405	0	358649
Donkeys	0.357	**.366*** **	**0.35**	0.001	358649
Boats	0.173	**.395*** **	**0**	0	358649
Nets	0.101	**.23*** **	**0**	0	358649
Lines	0.092	**.209*** **	**0**	0	358649
Hooks	0.089	**.203*** **	**0**	0	358649

Notes: Overall F-test; F-stat: 12916.6129; P-value: 0

TABLE A.2.5 MEAN DIFFERENCES ACROSS CLUSTER (VILLAGE)-LEVEL CHARACTERISTICS

INDICATOR	OVERALL	FOOD DATA	NO FOOD DATA	P-VALUE	TOT N
Female Household Head	0.453	.461***	0.447	0	358649
Average Household Age	19.299	**19.798*** **	**18.91**	0	358649
DR	1.815	1.75***	1.866	0	358649
Proportion of Female Household Members	0.487	.489***	0.485	0	358649
Presence of School-Aged Children (5–15)	0.784	.77***	0.794	0	358649
Proportion of Children Attending School	0.385	**.407*** **	**0.368**	0	358649
Wall Is Made of Grass	0.29	**.648*** **	**0.011**	0	358649
Roof Is Made of Grass	0.507	**.612*** **	**0.425**	0	358649

Household Owns Bed	0.434	.336***	0.511	0	358649
Household Owns Paraffin Lamp	0.122	.144***	0.105	0	358649
Household Owns Frying Pan	0.27	.252***	0.283	0	358649
Household Owns Mosquito Net	0.379	.407***	0.357	0	358649
Household Owns Mobile Phone	0.184	.189***	0.179	0	358649
Household Owns Donkeys	0.357	.367***	0.35	0	358649

Notes: Overall F-test; F-stat: 74670.1305; P-value: 0

Although the loss of representativeness caused by restricting the sample to the sub-population with consumption data is a concern, this analysis seems to indicate that balanced matched samples of HSNP beneficiary and non-beneficiary households can be achieved, if the appropriate types of selection variables are included in the first stage of the matching procedures.

A.2.3 A back-up approach

Given the considerations made in the section above, we believe that a PSM method could be used as a back-up approach based on a combination of MIS data and additional data collected in the first round of the evaluation. However, when the first round of data collection will be undertaken, the HSNP CT will have already produced some effects on beneficiary households' and individuals' general welfare. Therefore we will not be able to use variables that are influenced by the programme in the PSM selection model as the variables that need to be included in the first-stage propensity score selection model are supposed to influence programme participation but cannot be influenced by the effects of participating in the programme (i.e. receiving the HSNP transfer in our case). This additional information would therefore be derived from so-called 'static' variables that do not change over time and are not influenced by the programme intervention. We believe that the wealth of information collected with the multipurpose survey that we will be used in the first round of data collection will allow us to construct the set of additional variables on household and individual characteristics that have not been affected by the HSNP and can therefore inform our matching algorithm.

In any case, once the sample of households is selected for the analysis, further sensitivity checks will be performed on the resulting sub-sets of households with and without food consumption data to determine whether the distribution of their characteristics still shows significant differences. It is in fact possible that the final sample emerging from our multi-level cluster sampling strategy based around the RD county-level PMT threshold could present a different distribution of household characteristics, as compared to the entire MIS dataset. This will be reassessed with a distribution analysis focused on treatment and control households as well as households with or without food consumption data.

Annex B HSNP evaluation security plan executive summary

Having conducted the first phase of the HSNP successfully, OPM bid for and successfully won the contract to implement the second phase of HSNP on behalf of DFID and the GoK. This will entail the implementation of a bespoke communications and learning workstream, which will work closely with the PILU and NDMA to ensure a solid understanding and ownership of the programme's evaluation approach and findings.

The team structure is designed to respond to the particular needs of the HSPN programme. It comprises a mix of leading international expertise and Kenyan knowledge and experience. This is embodied in the pairing of an international TL and full-time KNPC, both based in Nairobi for the duration of the project. All of the field research will be conducted by leading a Kenyan field research company, RGA, with continuous and extensive involvement from relevant international team members under each of the four proposed workstreams.

While OPM does not have a DoC for the field research team (RGA), the programme will employ a core evaluation team of 12 experts, supported by a pool of senior experts, all of whom will come under OPM's DoC. In order to deliver the project successfully, in addition to the TL and KNPC based in Nairobi, other members of the core evaluation team and senior experts will be required to work and travel to Kenya and potentially into a number of high and extreme risk areas throughout the country. While it is likely that travel will initially be contained within Nairobi and some of the larger towns/cities, travel into more remote and potentially riskier areas may be required as the project progresses. Security in each of these locations must be managed to a standard that meets the requirements of the OPM DoC statement and DFID's DoC requirements as set out in the HSNP ToR.

Kenya has seen an increased risk of terrorism and kidnapping over the last few years from extremists linked to the Somali Jihadist militant group, Al-Shabaab, who have pledged allegiance to Al Qaeda. In addition to the highly publicised Westgate Mall attack in Nairobi in September 2013, there have been several other significant attacks in Kenya in the past year, which have included attacks mainly in Mandera County on the north-eastern border with Somalia, but also in Wajir, Lamu and Tana River Counties, and Mombasa on the coast. The Foreign Office recommends against all but essential travel to many of these locations, especially within 60 km of the border with Somalia. At the time of writing, at least six attacks have occurred in the past week in Mandera and Wajir.

In addition to the terrorist threat, there is also a threat of banditry throughout the country, cattle-rustling (that has led to violence) and intertribal conflict in some northern areas.

While not confined to Nairobi, there is a significant threat in the capital from carjackings, street crime (including mugging, sexual assault, armed robbery), credit card fraud, kidnap, terrorism, civil unrest and police corruption. Throughout the country, road traffic accidents are commonplace through poor road and driving conditions, poor driving, and the poor state of vehicles.

The project manager and TL will require support in assessing and managing their risks, as well as providing training and security briefings to programme team members. An analytical security product is required to give managers a forward-looking snapshot of the current risks that might affect their teams on a regular basis.

The security plan for HSNP will be a layered structure, with a series of mitigating elements to manage the risk to the team, any assets and the programme itself. No single measure will defeat risk alone, and the strategy has been designed to give a large measure of resilience to the programme, and ensure all programme employees and consultants have maximum awareness of their risks, and have access to the means with which to mitigate them.

The following layers form the security structure of the project:

Security management: Given that there are only two project members based permanently in Nairobi, and the rest of the team will visit Nairobi and further afield on an ad hoc basis, and in limited numbers, there is no requirement for a permanent security manager to be appointed to the programme. The Spearfish senior consultant (SSC) will provide remote security oversight for the programme and remain in regular contact with the OPM HSNP project manager. He will provide information, 24/7 response, and requisite training and advice where necessary.

Field security: The location of personnel will be managed by the use of satellite and mobile phones throughout the team where possible, with advice on vehicle selection and field standard operating procedures.

Security information: There will be country-specific products providing forward-looking statements on a periodic basis, as well as spot reports as incidents happen, overseen and tailored by the SSC.

Security training: A module including first aid for all staff and security awareness training has already been provided to core members of the programme team. Further training can be provided on an ad hoc basis when deemed necessary and by mutual agreement.

This plan is structured in 3 parts: Part 1 – security policy and objectives, Part 2 – standard operating procedures and Part 3 – contingency plans. Day-to-day responsibility for the implementations and upkeep of this plan will rest with the SSC.

The full security plan is available on request.

Annex C Evaluation ToR extract – purpose and objectives as well as scope of work

Department for International Development

UKaid from the British people

PO 6711

Evaluation Terms of Reference

Title:	Kenya Hunger Safety Nets Programme – 2 and Arid Lands Support Programme

Introduction

1. The Government of Kenya, with support from DFID Kenya and AusAID, has recently started the second phase of the Hunger Safety Nets Programme (HSNP) which focused on people living in extreme poverty in 4 counties in Northern Kenya. The first phase 2007-2013 has provided 69,000 households (496,800 people) with predictable electronic cash transfers worth approximately £13 per household per month. Independent evaluation has shown that HSNP works effectively as a safety net, particularly for the very poorest. It has helped families to be more food secure, hold onto their assets during shocks, and spend more on health.

2. HSNP 2 started in July 2013 and will run till March 2017. It will continue to focus on the same 4 counties as in HSNP phase 1. DFID plans to provide £85.59 million and this time the Government of Kenya (GoK) will also contribute funding to these transfers as part of the National Safety Net Programme (NSNP) – by 2017 it is envisaged that 49% of total programme costs will be met by the GoK. Under HSNP 2, all households in the 4 counties have been registered for bank accounts and the HSNP will together provide up to 100,000 of these households (targeting is using a combination of a proxy means test derived from the registration exercise with a Community Based Wealth Ranking with an adjustment done using the Commission of Revenue Allocation formula that incorporates the resource sharing that is currently being used by the Government of Kenya to allocate resources to the devolved county governments) with cash transfers worth approximately £17-£19 a month. The money will be provided directly into recipients' bank accounts.

3. At present there is no intention to vary the level of the cash transfer across households; however this is an issue of interest and so might change before 2017.

4. As well as the predictable cash transfers, HSNP 2 also has provision for additional Government transfers to be triggered in the event of a drought or other shock, which may be targeted at those in receipt of the monthly transfers and/ or other beneficiaries meeting requisite criteria.

5. These ToR also cover evaluation of the DFID Arid Lands Support Programme (ASP) which complements the HSNP by focusing on longer-term resilience to drought and shocks- including supporting community adaptation strategies, establishing livestock insurance, encouraging livelihoods diversification and using the HSNP mechanism to provide emergency cash transfers in times of drought. No separate evaluation of the Arid Lands Support Programme is envisaged.

6. In understanding the requirements of these ToR, it is useful to be aware that the GoK's National Safety Net Programme (NSNP) brings together all the social protection programmes in Kenya under one umbrella. The World Bank's Programme for Results (P4R) programme is supporting development of the NSNP and some of the indicators that will trigger payments to Government under this P4R programme rely on data from the HSNP programme and its evaluation.

7. Within the Government of Kenya, the National Drought Management Agency (NDMA) is responsible for leading on HSNP 2. A Programme Implementation and Learning Unit (PILU) is being created to manage and monitor HNSP 2, provide oversight of a rights and grievances mechanism for the programme and also oversee the independent evaluation covered by these ToRs. The PILU will report to the NDMA and will contain a mixture of NDMA civil servants and external counterparts with technical expertise that will be imparted to NDMA over the lifetime of the PILU.

8. DFID is separately tendering the contract for the evaluation of HSNP 2 as we want this to be independent of the PILU and ensure we get the best possible evaluation provider; however the intention is that the PILU will take on responsibility for managing the evaluation component after contracting. It will not be possible for the same contractor to be awarded both the main PILU contract and the evaluation work specified in these ToRs. . Current ToRs for the PILU are available at Annex 1 for reference purposes

Purpose and Objectives

9. **Purpose:** Independent evaluation of the 2013-2017 HSNP 2 programme is required to provide evidence on programme performance of use to: all programme implementers, DFID and Government of Kenya with regard to future decision making and accountability for funding, and the wider community interested in cash transfers nationally and internationally.

10. **Objectives:** Some broad objectives for the evaluation are:

    a. To determine the extent of the HSNP 2 programme impacts across a range of dimensions including at the household and community level;

    b. To determine strengths and weaknesses of operational elements of the programme and allow improvements to be made during the programme's lifetime.

c. To facilitate DFID Kenya and GoK decision making, in particular, to inform future targeting and the size of a sustainable caseload for the GoK which maintains an effective safety net function.

d. To help create a platform for wider Evaluation of work in the Arid and Semi Arid Lands (ASALs) of Kenya where the HSNP is operating. The data collected will be made available to all interested parties.

11. Service Recipients: NDMA and DFID will be the primary recipients of this work.

Approach to the evaluation

12. It is envisaged that the evaluation will look at both impact (through quantitative and qualitative methods) and process questions. For some questions a case study approach may be appropriate and for some issues the evaluation should take a theory-based approach and consider the underpinning theory of change and how that holds in practice. In particular, this could be useful around the interface of the HSNP and Arid Lands Support Programme. The Theory of Change for HSNP 2 is shown in Annex 2 and for the ASP in Annex 3. While we have a set of questions for the evaluation to address and a clear idea of how to approach some of the questions, the service provider will need to determine the evaluation approach for some questions and there may also be scope to add to or amend the evaluation questions over the lifetime of the evaluation.

13. A quantitative impact evaluation was undertaken under HSNP 1 which tracked beneficiaries and non-beneficiaries from 2009 to 2012 (OPM/IDS HSNP Quantitative Impact Evaluation report May 2012; OPM HSNP Quantitative Impact Evaluation report June 2013). Evaluation of HSNP 1 also considered operational elements of the programme (OPM Operational Monitoring Report, June 2013. Evaluation reports are published on the UK Government's website, HSNP's website http://www.hsnp.or.ke/, OPM's own website and in the World Bank micro data portal - HSNP1- IE report. While we feel it is necessary under HSNP 2 to continue with some longitudinal tracking of beneficiaries and non-beneficiaries in order to learn more about the impact of the transfers on households and help improve future targeting, we are not able simply to extend the impact evaluation that has been undertaken in HSNP 1. This is because some non-beneficiaries from the earlier study are now scheduled to enter the programme and receive payments and some beneficiaries will not be eligible for payments under the new targeting approach – only 28% of phase 1 beneficiaries are included in phase 2 of the programme. Thus a different approach is envisaged under HSNP 2, although it may be that there is more that can be done drawing on the phase 1 evaluation dataset, and potentially using it as a sampling frame for the future for some questions. This dataset has been publicly lodged with the appropriate documentation at the World Bank's micro data archive if bidders wish to access it.

14. The proposal for the evaluation of impact of HSNP 2 is to draw heavily on the registration and re-registration data for the programme, enhanced with an end-line; re-registration is not the responsibility of the evaluators but the final end line data collection on a sub-sample will be. Thus there is not an immediate need to create a comprehensive baseline position for this major part of the evaluation, but

rather to do comprehensive analysis of the existing data and plan for studies to respond on the additional questions which cannot be addressed in this way. There may still be a need for a baseline on some specific issues that are not sufficiently captured by the registration questionnaire as discussed later in these ToRs. The registration instrument is shown at Annex 4.

15. While there is currently an agreed commitment and accompanying budget for re-registration to take place at the end of the second year of operation of the programme, if for any reason this decision was changed, there would be implications for the design of the evaluation which would need to be addressed. This is not considered a likely issue and so it is proposed that it is only considered further should the need arise, but we do need the evaluation team to have the skills to tackle this if necessary.

16. The evaluation team will rapidly need to understand the monitoring system of HSNP 2 which may be of use for evaluation purposes and be aware of IFMIS (the Integrated Financial Management System) used by GoK. There is also a Perception Survey of beneficiaries being undertaken (PIBHS) – where the lead is the Social Protection Secretariat. Liaison will be required with other agencies over plans for further data collection, sharing information and trying to develop a common evaluation platform to maximise the use of the evidence being generated while maintaining individual confidentiality. It is currently anticipated that the Government of Kenya's national Integrated Household Budget Survey (KIHBS) will take place in 2015 – with fieldwork running over a year – and a reasonable sample size in the arid lands, however data from this will probably not be available until late 2016.

Evaluation Questions

17. The business case for HSNP 2 identified some key evaluation questions and listed them as 'first' and 'second' order. This list has now been refined and added to by looking also at the evaluation requirements of the NSNP, the needs of the Arid Lands Support Programme and recommendations from the annual review of DFID Kenya's current Social Protection Programme. A few of the second order questions have been dropped as it was felt the evaluation could not really answer them.

18. The questions of interest are grouped according to the DAC evaluation criteria below. It is possible that more will emerge over time. The list covers evaluation issues relevant to HSNP and the Arid Lands Support Programme (ASP); questions particularly focused on ASP are marked with an asterisk. While some questions will be possible to address under any circumstances, there are others where our ability to address them will be contingent on what happens in practice – for example, around using the HSNP mechanism to trigger payments, and whether different transfer values are delivered (either in a triggered response, in a deliberate test, or with another funder joining the programme and offering a different size of transfer).

Impact

a. What are the overall effects of the cash transfers in terms of: consumption[36], asset retention, well-being (proxy means test as measure), food diversity, health, education, self-esteem[37], comparing beneficiaries and non-beneficiaries?

b. For which sub-groups are effects most pronounced? (taking account of poverty status, household size, family composition, geographic location, livelihood base, gender and disability)

c. (How) Do cash transfers impact on women's control of cash within their (often polygamous) households and their wider empowerment?

d. How do the effects of predictable transfers compare with those of short term transfers triggered in response to acute shocks? How do the larger one off transfers some households will receive due to later than anticipated start of the programme impact on those households?

e. Does the combination of cash transfers and wider livelihoods activities open up new livelihoods opportunities/ income generating activities for poor households? How? (need to underpin this with a broader understanding of the nature of livelihoods in these areas – no. and type of activities and how diversify when additional funds available and during times of shock?)*

f. What kind of multiplier effects are found in local economies?[38]

g. Is there evidence of the programme having an impact on community relations – both within and between communities?

h. To what extent does NDMA policy dialogue contribute to changes in the budget allocations of GoK in respect of the NSNP? (this is also a sustainability issue)

Effectiveness

i. How well does the system based on use of HSNP registration data work as a basis for targeting the transfers and selecting households for a triggered drought response payment?

j. How well does the rights and grievances process work?

k. Does the new payment platform and expansion of financial services provide benefits for beneficiaries and non-beneficiaries?

l. Do the HSNP management structures and processes work well and provide value for money?

---

[36] NSNP indicators include 'net change in beneficiary household monthly per adult equivalent consumption expenditure' and 'net change in beneficiary household dietary diversity score'

[37] There is an NSNP indicator looking at change in self-esteem of decision makers

[38] Some interesting evaluation work around cash transfer programmes is now looking at the wider multiplier effects in local economies. One reference for more on this is: http://www.fao.org/economic/ptop/home/methodology/en/. 'From Protection to Production' is a research initiative jointly funded by FAO, DFID and UNICEF.

m. How effective is the work with NGOs to understand how communities respond to shocks under ASP? Does this get reflected in county plans?*

Efficiency

n. Are triggered payments more/ less cost effective than longer term predictable cash transfer payments?

o. What are the costs and benefits of HSNP 2? Can we undertake CBA using DALYs, rate of return on education, changes in welfare or assets?

p. Do the gains associated with a package of complementary activities to support livelihoods and cash transfers represent value for money compared with just providing the cash transfers?*

Sustainability

q. Is NDMA developing the capacity to deliver HSNP in the future?

r. Can we learn anything about the transfer value for future transfers? Could a lower amount have a similar impact for some Households? Could more have sufficient impact to justify the extra cost?

s. Do the reliable cash transfers build peoples' resilience to climate variability?

Relevance

t. In areas where the HSNP is linked to other activities, is the combination of a targeted cash transfer, mechanism for additional triggered cash transfers and additional work on livelihoods an appropriate response to the situation in the 4 Counties?*

u. Where possible to compare, how does the cost/ impact of HSNP/ ASP compare with other interventions (e.g food aid) used to support households in the four counties?

19. With regard to addressing the **impact** questions, the HSNP 2 registration data gathered already for every household in Turkana, Marsabit, Mandera and Wajir to enable targeting of the cash transfer and provide data to allow the setting up of bank accounts for all households, is a great resource. NDMA will be managing access to the HSNP 2 registration dataset and full access will be available to meet the evaluation needs. The intention is also to re-register all households in 2 years' time. We thus envisage this registration/ re-registration data being used to address a number of impact evaluation questions in bullet (a) and (b) above with the evaluation contractor expected to design and undertake an appropriate end line survey to complement this.

20. Part of the NSNP is to create a 'single registry' for social protection programmes in Kenya and thus it should also be possible to compare households in the registration dataset who receive different/ multiple transfers that are recorded on the registry. It is also possible that over the life time of this programme more interventions in the 4 counties will use the HSNP registration data and provide their own data about who received programmes and with what benefits, enabling more complicated analysis which captures the reality of programming beyond the NSNP and ASP.

21. While the registration data has complete population coverage, the registration instrument does not contain as much detail as might be expected in a survey on a smaller scale – for example it has little on consumption, thus there will be some limits to the questions that can be addressed and we expect to undertake a number of smaller scale qualitative and quantitative studies, which can go into more detail in specific areas. Such supplementary studies will be able to use the registration data as a sampling frame and might include work on education, consumption patterns, livelihood diversification, climate resilience, gender impacts and effects on self-esteem.

22. It is likely that the evaluators will be able to get a few additional questions added to the re-registration questionnaire which will be of use for rounds 2 and 3. The registration data includes geo-references (not necessarily by household, but certainly by community) and so there will be interesting possibilities to link to this, information on access to schools and health facilities and other data which may be of interest in analysis. Some of the impact questions of interest go beyond the beneficiaries of transfers, to consider effects on local economies as we would like to learn more about any multiplier / spill over effects generated by the regular, or targeted/ triggered cash transfers. A study focused on this is covered within these ToRs.

23. In summary, it is currently anticipated that questions **a, b** and **d** will primarily be addressed through the analysis of registration/ re-registration and end line survey data –with additional work likely to be required to look at education attendance, self-esteem and consumption patterns in more detail. Special studies are also likely to be required drawing samples from the registration data to address questions **c, e** and **s**. Question **f** would require a separate study with appropriate methodology. A process evaluation would be required to address questions **g and i-m, q** and perhaps **h** with findings usefully emerging during programme implementation to enable improvements to be made. Questions **n-p,** would require analysis combining evaluation data on impact and effectiveness questions with cost information from monitoring work. Questions **t** and **u** would require some benchmarking against other options and wider assessment of relevant work in the ASALs. Questions r, t and u are probably going to be most difficult to address fully, and are contingent on what happens in practice, but we are keen to learn any lessons we can in this area. The evaluation team will be required to monitor the various aspects of risk within each of the evaluation questions/and possible outcomes.

Scope of work

24. At the end of the 3 month Inception phase, there will be a contract break point to review Inception Outputs. Progress to the Implementation Phase will be subject to the satisfactory performance of the service provider, delivery of Inception outputs and the continuing needs of the Programme. There will be a further break point at the conclusion of the annual review of the programme in September 2015, to allow an on-going assessment of the successful operation of the evaluation work and progress against objectives.

25. The main implementation phase will run from December 2014 to March 2017. The evaluation team will be responsible for evaluation design, new fieldwork and all associated logistics and for analysis and reporting in order to address the questions identified above.

26. In undertaking this evaluation work, DFID's ethical principles for Evaluation and Research should be upheld (see Annex 5). Quality assurance to ensure a high standard of design and reporting is expected within the Evaluation team, but will also be available at certain stages through DFID's external quality assurance service (SEQAS) which provides comments on evaluation terms of reference, design papers and reports. DFID country teams are required to use this service and DFID will facilitate its use. DFID expects all evaluations to adhere to DAC quality standards.

27. During implementation it may be that some sharing of resources or facilities will be possible between this team, the PILU and NDMA to achieve cost efficiencies (e.g. on the costs of security advice or office space if required by field workers in the 4 counties).

28. Please note that DFID expects all data collected under this programme will be made available in anonymised format, to other users, within 12 months of the last data collection period in a data repository.