

IDS Bulletin

Transforming Development Knowledge

Volume 53 | Number 1 | February 2022

THEORY-BASED EVALUATION OF INCLUSIVE BUSINESS PROGRAMMES

Issue Editors **Giel Ton and Sietze Vellema**



Notes on Contributors	iii
Introduction: Contribution, Causality, Context, and Contingency when Evaluating Inclusive Business Programmes Giel Ton and Sietze Vellema	1
Systems, Sapiens, and Systemic Change in Markets: The Adopt-Adapt-Expand-Respond Framework Ben Taylor and Jake Lomax	21
Using Theory-Based Evaluation to Evaluate Systemic Change in a Market Systems Programme in Nepal Edward Hedley and Gordon Freer	43
Assessing the Contribution to Market System Change of the Private Enterprise Programme Ethiopia Giel Ton, Ben Taylor and Andrew Koleros	63
The Search for Real-Time Impact Monitoring for Private Sector Support Programmes Fédes van Rijn, Haki Pamuk, Just Dengerink and Giel Ton	87
Monitoring Systemic Change in Inclusive Agribusiness Sietze Vellema, Greetje Schouten and Marijn Faling	103
Assessing Contributions Collaboratively: Using Process Tracing to Capture Crowding In Marijn Faling	123
Understanding Behaviour Change in Theory-Based Evaluation of Market Systems Development Programmes Jodie Thorpe	141
Glossary	165

Using Theory-Based Evaluation to Evaluate Systemic Change in a Market Systems Programme in Nepal

Edward Hedley¹ and Gordon Freer²

Abstract The complexities of markets and market environments are felt in the design and the evaluation of market systems development (MSD) programmes. The authors reflect on a recent evaluation of an MSD programme in Nepal in which they used contribution analysis as a means of navigating these complexities. The planned niceties of the proposal soon departed ways from the reality on the ground, forcing the authors to adopt a more iterative evaluation approach, while ensuring evaluative robustness. This article outlines the iterative process and what the authors have learned regarding the applicability of contribution analysis within a theory-based evaluation, in a dynamic, changing environment.

Keywords theory-based evaluation, market systems development, contribution analysis, dairy value chain, inclusive business, Qualitative Impact Protocol (QulP), Nepal.

1 Introduction

Market systems development (MSD) programmes work with a variety of public and private sector actors to improve the way that markets function for the poor as consumers, producers, or employees. These programmes often work in complex environments where the level of complexity may grow more intricate given circumstantial factors such as a rich donor environment or thin markets or fragile socioeconomic contexts. Evaluating MSD programmes necessitates a methodology that is 'complexity-aware' and enables the evaluators to unpick the role of multiple overlapping drivers of change to uncover and understand the inner workings of programme processes (Chen and Rossi 1980, 1983, 1992) – unpacking the notorious 'black box' of evaluation (Stame 2004). In a recent evaluation of an MSD programme, we used contribution analysis as a means of categorising market forces and contributions within the

© 2022 The Authors. *IDS Bulletin* © Institute of Development Studies | DOI: 10.19088/1968-2022.104



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non Commercial 4.0 International licence (CC BY-NC), which permits use, distribution and reproduction in any medium, provided the original authors and source are credited, any modifications or adaptations are indicated, and the work is not used for commercial purposes.

The *IDS Bulletin* is published by Institute of Development Studies, Library Road, Brighton BN1 9RE, UK
This article is part of *IDS Bulletin* Vol. 53 No. 1 February 2022 'Theory-Based Evaluation of Inclusive Business Programmes'; the Introduction is also recommended reading.

programme 'black box'. In this article, we explain the design and application of our process, as one possible tool to use within theory-based evaluation.

The programme we evaluated worked in a variety of agricultural sectors and we illustrate the approach we took by drawing on examples from our evaluation of the dairy sector. Initially, the evaluation was planned as a two-step process to determine the level of contribution of the programme. However, early in the process, we realised that we needed to adapt this approach. In the article, we explain the original design, our adaptation, and our rationale for taking these steps. We then reflect on the application, noting the insights gained, and using examples that illustrate how the methodology helped us to answer the question 'What interventions worked and why in generating systemic change?'

The article is structured as follows. Sections 2, 3, and 4 provide some context to the evaluation and focus on why we selected a theory-based evaluation design to meet our objectives. Section 5 highlights how we applied it – drawing on a worked example from the dairy sector. It discusses what we learned at each step in the process and how we adjusted our approach along the way. We conclude in Section 6 with our reflections on the process as a whole.

2 Background

In 2020, Itad Ltd³ completed an evaluation of the Samarth-Nepal Market Development Programme (Samarth-NMDP) for the Foreign, Commonwealth & Development Office-Nepal (FCDO-Nepal) (Itad 2019). The Samarth-NMDP programme ran over a six-year period from April 2012 to March 2018 and was the first programme to apply the MSD⁴ approach in Nepal. It worked across several agriculture sectors and tourism, with the aim of making these markets more inclusive of poor people. Itad was commissioned to undertake an endline impact evaluation of the Samarth-NMDP programme, with several major objectives. A significant objective of the evaluation was to determine the extent to which programme interventions had initiated change that might result in systemic change.⁵

3 The programme context and the evaluation challenge

By their nature, MSD programmes present evaluators with challenges. Their *raison d'être* is to include the excluded poor in fair, functioning markets; their ultimate objective is systemic change to overcome this challenge; and their *modus operandi* is to operate at arm's length, encouraging sustainable solutions while avoiding creating distortions within emerging market dynamics. In this way, MSD programmes seek to change the way markets work to benefit the poor by facilitating systemic change. Fundamentally, promoting systemic change derives from a wish to make the benefits of development intervention as inclusive and long-lasting as possible. Programmes achieve that aim by empowering market players to understand and overcome the challenges in market

rules and functions, rather than by offering more traditional direct delivery development assistance. However, the lack of direct involvement, the longer-term goals, and the complex, multifaceted market environment present the evaluator with evidence that is more circumstantial than direct or physical in nature.

The Samarth-NMDP programme was designed from the outset to be an MSD programme. In the first three years of implementation, however, in an experimental attempt to achieve scale more rapidly and to gain buy-in for the unfamiliar MSD approach from key government stakeholders, the programme partnered with established non-governmental organisations (NGOs) that were active in the programme's sectors of focus and unfamiliar with the MSD approach. Unintentionally, this resulted in a more direct delivery model being practised and limited traction for the MSD approach. After a 'midcourse correction', the programme opted for a more facilitative market systems approach. Market development programmes generally require time to catalyse systemic change; with the programme implementing a fully MSD approach for only two or three years, we anticipated that any observed changes in the market system would likely be embryonic. This made our evaluation that much more difficult. Systemic change is generally difficult to pinpoint, and now we were obliged to look instead for early indicators of this change.

Events external to the programme also created a challenging environment. The programme worked in a congested donor environment in which multiple government and other donor programmes occupied similar geographic locations. These were working in the same sectors, often with overlapping aims and objectives and sometimes in direct tension with the programme's market systems approach. The programme also worked against a backdrop of rapid socioeconomic change (such as rural-to-urban migration), a changing and fluid political landscape, and encountered a series of serious shocks, including the 2015 earthquake and the 2015 Indian Economic Blockade. Finally, while some of the interventions focused on sectors with significant market activity, in other areas, the programme operated in a classic 'thin market' context characterised by sparse and underperforming market-supporting functions – especially market failures in agricultural input markets and post-production services, poor physical infrastructure, and weaknesses in the policy and regulatory environment.

In short, given this context, we acknowledged at the outset that it would be difficult to collect evidence of systemic change and that we would have to seek indirect evidence that would give a plausible indication that systemic change would manifest itself in the future. Even where we found proof, the programme's contribution to change might not be obvious, given the multiple overlapping drivers of change, and the long and indirect pathways to impact.

In response to these challenges, we developed an overarching theory-based approach to the evaluation framed by contribution analysis. We selected contribution analysis precisely for its ability to deal with complexity, for its capacity to weigh up the relative importance of different factors, and its iterative and exploratory nature. Within this contribution analysis framework, we developed different evaluation modules based on a mix of different methods, including household surveys and key informant interviews.

4 What were we looking for?

The holy grail of MSD programmes is a systemic transformation that makes a market beneficial for poor producers and consumers. However, despite decades of implementation, a clear definition of systemic, transformative change remains elusive.

Given that Samarth-NMDP (hereafter known as 'Samarth') had already used the Adopt-Adapt-Expand-Respond (AAER) framework (see Taylor and Lomax, this *IDS Bulletin*) in its reporting on results achieved by the programme in the area of systems change, we opted to use this framework as a primary lens to identify evidence of systemic change. We purposively selected five programme sectors (dairy, vegetables, pigs, ginger, and tourism)⁶ for study, based primarily on the programme's reporting of achievement against the AAER framework and the sectors' reach across the programme portfolio, considering their importance in terms of numbers in the target population (smallholder producers). This selection was made to balance the twin evaluation aims of accountability and learning. The dairy sector was selected, for example, because of its importance within the Samarth portfolio and its importance to the Nepali economy (contributing an estimated 8 per cent to Nepal's gross domestic product, with more than 3.5 million households engaged in the sector, of which 500,000 are producers and sellers of milk). Within each sector, we then identified three or four interventions for in-depth study based on similar criteria.

Figure 1 Examples of systems change within the AAER framework



Source Authors' own, adapted from Nippard, Hitchins and Elliott (2014).

Box 1 Top-down and bottom-up lenses

The **top-down phase** was designed to develop our understanding of the market contexts in which the programme had intervened, to deepen our understanding of the mechanisms through which the programme sought to influence these market systems, and to assemble existing evidence of change in the market system. This phase was designed to cover the first four steps of the typical contribution analysis cycle, drawing largely on programme data and secondary sources, but with additional primary research among key informants. This helped us to refine the research agenda for further data collection.

The **bottom-up phase** was designed to collect evidence of change in the market system introduced by individual interventions 'on the ground'. It harnessed a range of research methods appropriate to the different types of programme participants: among poor producers we used household-level quantitative surveys and participatory Qualitative Impact Protocol (QulP) interviews (see Box 2); among market-level participants (business, associations, and government agencies) we used semi-structured qualitative interviews following a snowball sample).

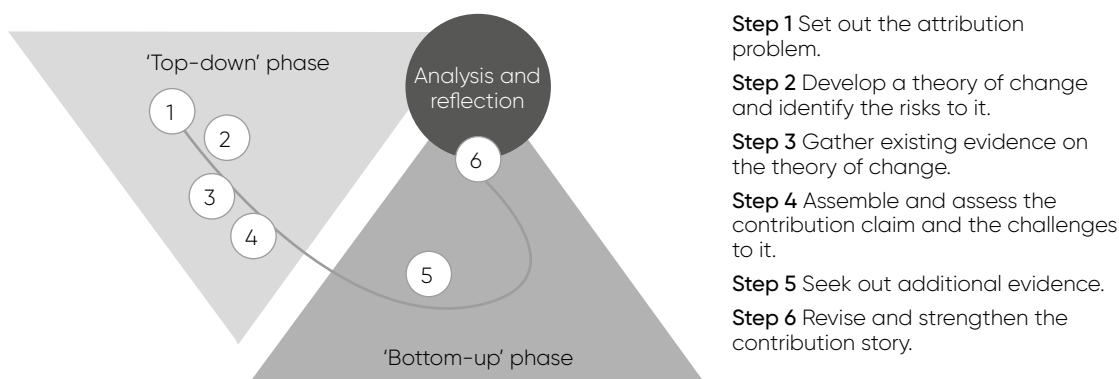
Source Authors' own.

In terms of what would constitute evidence of systemic change against the AAER framework, and taking our constraints into consideration, we were looking for evidence within the AAER quadrants which would suggest a subtle change of role or 'function shift' of some of the market actors (Fowler and Lomax 2021). We were also looking for evidence of replication of the interventions by other market actors ('crowding in') and for similar shifts in thinking and action by market actors in other areas of the same value chain. Figure 1 highlights the type of change that would constitute elements of systemic change against the AAER quadrants, with examples drawn from the dairy sector for illustration.

5 How did we identify early signs of systemic change?

Our approach built on one that Itad first developed for the Growth and Employment in States (GEMS) project in Nigeria to capture change at different levels of a system and then to attribute this change where possible to MSD interventions (Ruffer 2012). During the inception phase of the Samarth evaluation, we structured our approach into two broad phases, '**top-down**' and '**bottom-up**'. We mapped these research phases onto the standard phases of contribution analysis as illustrated in Figure 2.

Figure 2 Steps in contribution analysis illustrating our top-down and bottom-up phases



Source Authors' own, adapted from programme documentation.

Structuring our research in this way served a similar purpose to the 'helicopter' and 'intervention' lenses recently described by others in their *Pragmatic Approach to Assessing System Change* (Posthumus et al. 2020).

At the end of the process, we put the two lenses together through a process of reflection and analysis to understand not only **if** interventions (or combinations of interventions) had influenced change in the system (and the role played by the programme), but also **why** this change had occurred, and if this change was likely to be sustained and scaled, even if evidence for this was nascent.

By structuring our research into these top-down and bottom-up phases we expected that, collectively, our varied methods would build a good picture of how the market system(s) were operating and why certain interventions within these systems appeared more likely to generate lasting change than others.

5.1 Top-down phase

5.1.1 Defining and understanding the system under study

During the top-down phase of research, we first developed results chains for our selected sectors and interventions (in consultation with programme staff and drawing on programme documentation and reporting). These built on the programme's existing results chains developed as part of the Donor Committee on Enterprise Development (DCED) audited results measurement system. However, it added further detail in order to capture not only the programme's vision for system change in each sector, the role played by different interventions in supporting this vision and key assumptions and risks, but also external factors, including the perceived role and influence of other actors. These complexity-aware results chains (Britt 2013) therefore took into account what were thought to be relevant factors in the broader environment, including other market players and drivers of change, and

captured our initial understanding of the parameters and dynamics of the market systems that the programme sought to influence.

Our dairy results chain covered Samarth's work to improve incomes for poor households through improved market access and improvements in productivity and milk quality by adopting good manufacturing practices (GMP) and improving animal husbandry. We captured the roles and responsibilities of the most significant actors who were either engaged directly in these interventions or had a relevant role in the wider milk production and marketing system. As such, it covered smallholder producers and milk aggregators (including milk cooperatives and private regional-level milk processors), national-level milk processors of milk and cheese products, national-level dairy associations, and the various government agencies responsible for setting milk prices and defining and enforcing milk quality standards. While recognising and appreciating the 'soft' boundaries of market systems, this range of actors acted as a *de facto* delineation of the dairy market system, for the purposes of our evaluation.

At the end of this stage, we developed an initial set of sector-specific research questions, which translated our evaluation questions into a set of more tailored research questions for each sector based on our sector and intervention results chains.⁷ These guided the next stages of our research.

5.1.2 Top-down research

Our top-down research phase consisted of a review of secondary documents (in particular, Government of Nepal and donor reports) and a series of high-level interviews and workshops with sector key informants to get a 'birds-eye view' of the sector. These key informants were drawn from the public, private, and non-governmental sectors.⁸

The purpose of this phase was to deepen our understanding of system dynamics in our selected sectors and to focus on topics such as recent socioeconomic history and key development trends, as well as contemporary political economy and incentive structures for key actors. As part of this research, we interviewed our key informants about changes and developments in the sectors and the relationship between these changes and the programme's activities.

In the dairy sector, the top-down phase of research helped us to identify the key opportunities and challenges faced by the sector as a whole and in doing so to understand the aims, objectives, and reported achievements of programme interventions in this context. It also enabled us to develop a more informed and nuanced research agenda for the bottom-up phase of research to follow.

Our initial top-down research focused on validating programme information, filling evidence gaps, and providing us with a foundation from which to assess programme contribution to change. This confirmed that the dairy sector had experienced rapid growth in annual demand for milk and milk products (estimated at 8 per cent p.a.), which had outstripped supply from domestic sources and offered opportunities for producers to increase production. It also confirmed some of the key constraints in the sector related to production, value chain coordination, and the enabling environment. These constraints ranged from supply-side constraints such as inconsistent knowledge and application of GMP to regulatory constraints such as poorly enforced national quality standards. These constraints combined to tighten profit margins for milk producers despite the increasing national demand for the product.⁹

5.1.3 Learning from the top-down assessment of systemic change A strong and evaluable sector-level results chain or theory of change (ToC) that is 'complexity-aware' is critical to the research process.

We initially aimed to use the programme's overarching ToC and its sector results chains. However, these did not systematically and explicitly capture the expected system-level changes in all sectors. Nor did the overarching ToC articulate the relevant system constraints, assumptions, and risks which might have a bearing on these changes. The ToC diagram and results chains needed to be much more granular. Therefore, the research team spent more time than anticipated in creating 'nested results chains' in each sector in which groups of results chains for selected interventions fed into one 'complexity-aware' results chain for each evaluated sector. This proved to be a useful and necessary investment. It enabled us to define the systems we sought to study, understand their wider dynamics and strengthen our understanding of the intended role of the programme in driving change. These results chains provided the backbone for our subsequent research agenda and the platform against which we later combined our top-down and bottom-up lenses during our reflection and analysis phase.

A top-down preparation and research phase is particularly useful to gain a broader perspective on the system(s) under study and to help the evaluators get 'up to speed'.

Our top-down research helped us to quickly understand the dynamics of the systems under study. We found the combination of programme reporting, secondary document review, and interviews and workshops with sector key informants to be a particularly effective and efficient way of building this broad but detailed picture. The analysis of systems change is highly context-specific and requires in-depth knowledge of the sectors under study and the roles, responsibilities, and incentive structures of key actors. The top-down phase helped us to build a more detailed picture of the key assumptions and risks present in programme delivery

across the sectors and alternative explanations for change, given the wider dynamics present in the sectors.

A careful selection of programme areas for evaluation is vital.

Evaluators seeking signs of systemic change may need to delve more deeply into fewer programme areas, rather than adopting a wider but shallower approach. Systemic change itself takes time, and a deeper study is more likely to uncover these early signs. The inherent risk of a shallower approach is that the evaluated programme areas do not have sufficient depth to be able to reveal any signs of systemic change. This makes the initial selection process vital for evaluation success. To maximise the likelihood for success and opportunities for learning from this evaluation, we prioritised sectors in which the programme had been working for longer periods of time and had reported stronger early signals of systemic change. As a result, we decided not to include a number of more recent programme sectors (for example, the fish sector). Only where programme sectors were important from an economic and portfolio perspective – but where we determined that these had not been designed according to MSD principles, such as tourism – did we adopt a lighter-touch approach.

Evaluators need to be careful and deliberate in defining and delimiting the boundaries of the system(s) under study. Beyond selection of programme areas for study, a key finding for the team was that evaluators also need to be careful in how they define the systems selected for study and, within a finite resource envelope, not to overcommit in the research phase. Many potential avenues of research may open up during the top-down phase (and later during the bottom-up phase); the evaluator needs to make explicit choices as to where to focus their energy and resources. For example, in each sector the team made numerous decisions concerning which other potential drivers of change they should investigate and to what depth. In the dairy sector, for example, the team determined that Danida (Danish International Development Agency) support to develop guidelines for improved milk production was directly related to the aims and objectives of Samarth's interventions and was worthy of further investigation. By contrast, the team determined that the new relationships that Samarth had facilitated among market players in the sector had not been significantly impacted by another programme, the United States Agency for International Development's (USAID) large-scale Agricultural Growth Programme for livestock.

5.2 Bottom-up phase

5.2.1 Research strategy

Our bottom-up research phase involved a mix of different research methods designed to meet the varied aims of our evaluation.

Box 2 The Qualitative Impact Protocol (QulP)

The Qualitative Impact Protocol (QulP) draws on contribution analysis. QulP's approach places project beneficiaries' voices at the centre of the evaluation, enabling them to share and feed back their experiences in an open, credible, and respectful way. QulP gathers evidence of a project's impact through narrative causal statements collected directly from intended project beneficiaries. Respondents are asked to talk about the main changes in their lives over a pre-defined recall period. They are prompted to share what they perceive to be the main drivers of these changes, and to whom or what they attribute any change – which may well be from multiple sources. In some applications of QulP, the researchers do not know for which project the analysis is being done, limiting bias in deciding on contribution claims.

Source Authors' own based on the method developed by Bath Social and Development Research (2021).

We conducted **quantitative surveys** of programme participants (treatment groups) and compared the results with groups of non-participants with similar characteristics (comparison groups). Our aim was to determine with confidence whether programme interventions had resulted in producers adopting the new practices, whether these changes had become embedded in the market and were sustained, and whether they were continuing to produce benefits for poor households in terms of increased productivity and income (all core aims of the MSD approach). Data from these surveys also supported our understanding of the resilience and gender equity of the 'adapt' phase of market system change.

In the dairy sector, we undertook two rounds of data collection with 500 producer households, which were consistent with, and incorporated baseline and endline data collected through Samarth's own monitoring system. The evaluation team aimed to add further rigour by providing additional resources to increase sample sizes beyond those used normally by the programme; we added comparison groups and a further round of data collection after the programme had ended to better assess the sustainability of programme interventions.

Paired with these quantitative surveys, we commissioned **Qualitative Impact Protocol (QulP) studies** (see Box 2 for further detail on the method) to understand the issues faced by dairy farmers in greater depth and to identify the most important challenges, obstacles, and drivers of change. We proposed to triangulate this information with other evidence on the

contribution played by programme interventions that we had collected in both the top-down and bottom-up research phases. The design of the QuIP studies also enabled us to explore these issues in a participatory way from the perspective of smallholder dairy farmers on the ground, especially women and marginalised groups. In the dairy sector, our QuIP study comprised 24 interviews with individual respondents and four focus group discussions in two locations – one in a programme district and one in a matched control location.

Finally, since we aimed to understand the broader impact of interventions beyond the direct sphere of programme influence, we deployed a **semi-structured survey that asked qualitative information of market actors** following a 'snowballing sample' technique. These interviews started with market actors who were directly engaged in our selected interventions and had been identified with the support of programme staff, and then expanded outwards from there. We used insights gained through the initial interviews and the team's own market intelligence to identify additional actors who were thought to have reacted to, and potentially replicated, new practices as a result of the programme interventions. In the dairy sector, this included interviews with local milk producers, aggregators and traders, regional and national-level processors, producers' cooperatives, and local government agencies.

We opted for this semi-structured survey of market actors based on the expectation that evidence of systemic change would be limited and would need to be carefully identified and 'unpicked'. We were also faced with resource constraints and out-of-date and/or incomplete lists of market actors operating in programme districts and neighbouring locations. Unfortunately, there had been very limited record-keeping on the part of Samarth in terms of 'adjacent' actors that they had engaged with tangentially or indirectly during programme implementation. This made identifying credible sample frames difficult. In this context, we determined that quantitative surveys of a sample of market actors would not be feasible or cost-effective. This was one of the reasons why a snowballing sample technique was more feasible than a sampling strategy, which attempted to draw a representative sample of market actors from a pre-determined sample frame.

During our qualitative surveys, our researchers played the role of 'detective'; they tracked down systemic change 'leads', often based on partial or contradictory information, until they were satisfied that they had developed a fairly accurate picture of the change that had occurred and what factors were driving it. In situations with potentially conflicting data, asking probing questions of all the respondents and carefully corroborating this evidence was important, especially where the evidence was incomplete and potentially based on interviewees' incomplete recall over a time horizon of several years.

This combination of methods helped us to build up a picture of change in the market system 'on the ground' and to understand the programme's role in sparking this change, both within its sphere of direct influence and more broadly. While the qualitative interviews of market actors were undoubtedly the most important method in identifying and following up specific evidence of changes in the system beyond the direct influence of the programme, each method had a role to play in developing our understanding of the market and the programme's place within it. This is discussed in Section 5.3 on analysis and synthesis.

5.2.2 Learning from the bottom-up perspective of systemic change

The qualitative interviews with market actors placed particular demands on the team. We learned that researchers need to be knowledgeable about the MSD approach and be highly versed in the aims and objectives of the interventions and sectors they are studying. Experience in conducting qualitative research was particularly important so that researchers could effectively step into the role of 'detective' to identify, follow up on, and substantiate 'clues' as to the role of the programme, while at the same time knowing when to stop. In practical terms, we found that there needed to be consistency in the research process, with researchers being brought in as core members of the team for both data collection and analysis. Where we hired in researchers for a few interviews, the quality of the evidence suffered.

The reality of our research was 'messier' than anticipated and our bottom-up research did not always follow on neatly from the top-down phase as initially envisaged. From a practical perspective, a number of external factors complicated our research timings. These included agricultural cycles and intervention close-out activities which in some instances dictated that our bottom-up research needed to commence more rapidly than anticipated and overlapped with the conclusion of our top-down research phase. In some cases, our understanding of system dynamics was not as strong as we would have preferred and we missed opportunities to ask salient questions in our surveys.

We also found that our understanding of the system and its boundaries continued to evolve throughout our research. In practice, we continually updated our intervention results chains and research questions as subsequent research activities deepened our understanding of the role of key actors and important aspects of context; in some cases, we identified additional strategic top-down activities (additional research and interviews with sector key informants, for example) as a result. In sum, our research was much more iterative and irregular than we first envisaged.

Nepal's fluid political situation arising from the ongoing process of decentralisation proved to be challenging in a practical sense (although it did hold broader opportunities). It proved

Figure 3 The evaluation team plot evidence for dairy sector interventions on a white board



Source Authors' own.

challenging during our bottom-up research phase to identify and track down key respondents in local government who may have moved position following their central role with the programme. This resulted in numerous dead ends, with little value. However, it also presented an opportunity to contact a range of respondents who we may not have ordinarily identified and who were willing to provide their own thoughts and commentary on the sector situation and the role of different interventions in their geographic area of responsibility. This allowed us a far broader view of the practical realities on the ground, as well as a more comprehensive, unfiltered view of the regulatory environment governing the sector from the bottom up.

5.3 Analysis and synthesis: putting the 'top' and 'bottom' together

Given the challenging context for this evaluation, the final phase of analysing and synthesising the data was of central importance to building the programme's contribution story. Throughout the data-gathering process we had identified strands or threads of evidence that pointed to early indications of systemic change, which included hints and nuances in interviews and glimpses into the manner in which 'business was being done differently'. The phase of synthesis and analysis aimed to gather these hints, nuances, and threads of evidence, triangulate them, and weave them into a coherent evidence-based contribution story.

To do this, we reverted to the results chains we had developed during the top-down phase. We engaged the whole evaluation team in a participatory exercise to visually map the evidence we had collected through our research phases onto the result chains (see Figure 3). This mapping process was an iterative and inclusive

process, held over a few days, allowing all team members to record their thoughts and interpretations of data and evidence. All team members were encouraged to challenge the data, the strength of evidence, and even the positioning of the data within the visual map. All of this added value in testing the strength of evidence for early signs of systemic change, where this change seemed to be happening, and the contribution played by the programme.

Throughout the mapping process, we colour-coded the evidence by source and strength to aid the triangulation process and paid particular attention to evidencing the key assumptions made by the programme. Merging the wider, more inclusive top-down evidence with the more fine-grained evidence collected bottom-up allowed us an opportunity to critically examine our own 'data picture'. We could then take a step back to scrutinise this picture and to understand the selected interventions' influence within the system as a whole. This allowed us to consider other explanations of change in light of what we knew about the context in which they were operating.

Synthesising the data from both the top-down and bottom-up approaches in a visual manner assisted us in understanding the dynamics and external factors within the market system. The process was particularly useful in helping us to contextualise the changes introduced by the programme and to make an assessment as to the extent to which changes within the market system were likely to be sustained and/or scaled. This process sometimes led to a satisfactory understanding of the programme's contribution to change; on other occasions it raised further questions which led to additional data collection, either through further top-down document review and interviews, or additional interviews from the bottom up.

In the dairy sector, evidence from our various research methods helped us to identify a clear difference in the depth and sustainability of market system change brought on by first- and second-phase interventions. For example, although interventions in the first phase had successfully encouraged smallholder producers and the cooperatives and processors that purchase their milk to adopt changed practices, leading to increased productivity and milk quality, our quantitative household surveys revealed that adherence to these practice changes had begun to erode over time and the expected impact on household income had not emerged.

During the synthesis we had hoped that our QulP studies would provide a source of data triangulation to understand the programme's contribution in bringing about these changes, but this proved difficult given the challenges encountered in isolating the precise role of programme interventions within the QulP data. This resulted from a number of factors including the nature of

Box 3 A new business model in the dairy sector

In a second-phase intervention in the dairy sector, Samarth provided support to national-level processors (in the form of brokering new linkages with suppliers, providing technical advice and some financial resources) to enable them to form new working relationships and adopt new business models with regional suppliers of fresh milk and semi-processed dairy products. The national-level processors were encouraged to provide technical support to these regional processors to introduce new production practices, with the expected benefit of improved quality and consistency in the supply of raw milk to these national-level processors. In turn, the regional suppliers were encouraged to work with their smallholder producers of milk to support them to adopt new practices to improve raw milk quality, underpinned by the incentive of increased prices.

Source Authors' own.

the QuIP design, especially the 'blindfolding' of researchers, the indirect nature of the MSD approach (working at arm's length through the market), and the complexity of the research setting (in which multiple interventions were operating, often with similar objectives).¹⁰ Nevertheless, the QuIP data (alongside evidence from our qualitative survey of market actors and data from the household surveys) helped us to better understand the contextual factors that undermined the durability of the changes introduced into the market system by first-phase interventions and that limited their scope.¹¹

A key factor to emerge through the synthesis from both top-down and bottom-up evidence was the impact of increasing production costs. Programme participants identified increasing costs as a primary factor in eroding incomes from milk over time. The costs had not been offset by the expected 'price premium' from the production of higher-quality milk. This resulted in reduced willingness on the part of producers to sustain new practices. This, in turn, highlighted the importance and impact of a key contextual factor identified during the top-down research: out-of-date and poorly enforced national quality standards for milk which the first-phase interventions in the dairy sector had been unsuccessful in addressing. This undermined incentives for milk aggregators and processors to offer higher prices on a sustained basis for higher-quality milk.

The picture to emerge through the evidence synthesis for second-phase interventions in the dairy sector was quite different. There were early signs that a number of these interventions were gaining

traction in the market system and were contributing to changes that held the prospect of sustainability and scale. In particular, evidence from our qualitative survey of market actors revealed that a second-phase intervention that had facilitated the introduction of new linkages and working practices between large national-level dairy processors and regional-level suppliers of processed milk and cheese products (see Box 3) had resulted in both of these actors adopting changes to their practices to the benefit of smallholder producers. Those changes were sustained beyond the end of the financial support offered by the programme.

In this intervention, these new working relationships seemed to be proving to be durable because of the way that incentives had been aligned along the value chain: our qualitative survey of market actors revealed that the regional processors were continuing to implement improved practices. They were also continuing to work with smallholder producers after programme support had been withdrawn in response to the prospect of continued access to new, large, and higher-value urban markets. For their part, national-level processors were also continuing to support these suppliers in order to secure access to additional supplies of higher-quality and more consistent milk products.¹² In doing so, they had overcome the twin constraints of increasing production costs and weak official quality frameworks.

There was also evidence from the qualitative surveys of market actors, albeit nascent, that this intervention had gained traction in the market and held the potential for scale. The national-level processors engaged by Samarth were seeking to adapt the model (for instance, by offering finance to support smallholder producers to adopt practice changes to improve the quality of milk produced) and to expand their relationships to other regional processors. Regional processors meanwhile were taking further steps to adapt their production processes to the demands of urban markets and offer new processed and semi-processed milk products. In addition, other national-level dairy factories were seeking to crowd in with similar versions of the model described in Box 3.

It should be noted that Samarth was not the only programme to support some of these market actors – other programmes had offered equipment to at least one of the regional milk processors, for example. One of the key questions debated during the synthesis, therefore, was the degree to which Samarth could claim to have contributed to these changes. Based on evidence from multiple interviews with market actors and set against evidence collected through the top-down phase, the team concluded that Samarth had indeed made a very significant contribution, given that it could claim credit for initiating and supporting the development of the new working relationships between national and regional milk processors – which was the critical factor in explaining the intervention's success.

6 Concluding reflections on the process as a whole

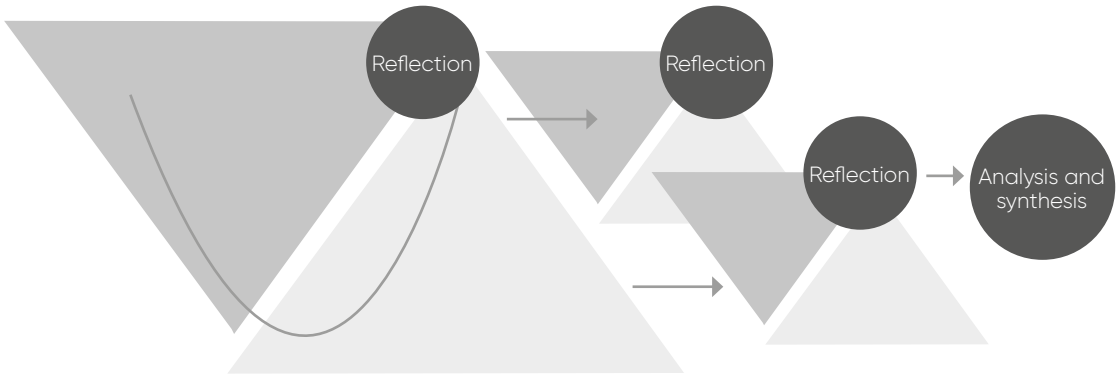
Overall, we found that our contribution analysis-based approach, framed by top-down and bottom-up research lenses, assisted us in identifying early signs of systemic change. Contribution analysis provided an appropriate analytical framework to navigate a complex environment (Mayne 2011), characterised by long and indirect results chains and multiple overlapping interventions and actors. The 'thinness' of the market and the impact of two huge, unanticipated external events (the trade embargo and an earthquake) would have significantly impacted our ability to identify a plausible 'business as usual' counterfactual through the results of the quantitative survey only, given its focus on inferring impact through with-without measurements in treatment and comparison groups only (Stern *et al.* 2012).

Our top-down and bottom-up research phases provided complementary inputs to help us determine if and how the systems we were studying had changed, the likelihood that these changes would be sustained and scaled in the market, and the programme's role in driving this. Faced with the initial programme ToC and results chains which were not 'systems-aware' and only providing a very loose definition of the market system, the top-down phase enabled us to develop more detailed impact logics for our chosen interventions; to define more clearly the boundaries of the systems in which they were working; and to deepen our understanding of these systems' characteristics (in essence, the contextual backdrop against which the programme was seeking to effect change). This work helped us to develop a detailed research agenda for our bottom-up phase.

Through our bottom-up phase, we progressively collected evidence of change in the market system against our interventions' results chains, drawing on a range of methods. We found that an exploratory, flexible qualitative survey of market actors, based on a snowball sampling approach, was an appropriate technique to identify and follow up on emergent evidence of systemic change, especially where it lay beyond the immediate scope of programme influence. However, this approach required skill on the part of researchers, who needed to know enough about the context and the intervention to develop relevant questions and have the ability to know where to probe further and where not.

We had initially planned that our contribution analysis framework would follow what might be referred to as a 'linear spiral' of ongoing evaluative activity, with our top-down and bottom-up lenses forming discrete phases of research, followed by a period of reflection and analysis. However, one of our key findings was that the reality of our evaluation implementation was much more iterative and 'messier' in reality, as illustrated in Figure 4. For example, we found that our understanding of the market systems in which our interventions operated continued to evolve during

Figure 4 'Messy' reality of applying our top-down and bottom-up research lenses



Source Authors' own.

the bottom-up data collection phase – as we learned more about context and the roles and responsibilities of key actors, for instance – and this necessitated further top-down research activities and periods of reflection.

Nevertheless, we found that our top-down and bottom-up framing added considerable value, especially during the analysis and synthesis phase. By combining a wide-angle view of the system as a whole with a narrower, more detailed perspective on change from the bottom up, we were able to take a step back to understand the selected interventions' influence within the system as a whole. This took into account context and other influences and explanations of change, which ultimately enabled us to identify those interventions that were demonstrating early signs of sustaining and scaling change in the market system.

In this case, we applied contribution analysis to identify embryonic traces of systemic change and to weigh up evidence from multiple perspectives as to whether the programme had contributed to the existence of these changes. Trying to define and identify systemic change in any MSD programme is challenging. However, we may take refuge in Justice Stewart's¹³ wisdom that often we may 'know what it is when we see it'.

Notes

- 1 Edward Hedley, Principal Consultant and Evaluation Director, Itad, UK.
- 2 Gordon Freer, Evaluation Team Leader, Department of International Relations, University of the Witwatersrand, South Africa.
- 3 Edward Hedley is a Principal Consultant in Itad's Inclusive Growth and Climate Change Practice and was Project Director for this evaluation. Gordon Freer is an independent evaluator and was Team Leader for this evaluation.

- 4 At the time of programme inception, the MSD approach was widely known as Making Markets Work for the Poor (M4P).
- 5 While there is no clear definition on what constituted systemic change, four perspectives of different types of systemic change are presented in Jenal (2019). See also Nippard *et al.* (2014).
- 6 Out of a total of 11 programme sectors which included pigs, dairy, fish, feed, vegetables, ginger, mechanisation, crop protection inputs, agriculture reconstruction, tourism, and media.
- 7 These questions covered the relevance, effectiveness, impact, and sustainability of the selected interventions, including questions around the extent to which they had produced systemic change, either singularly or collectively.
- 8 These key informants were selected based on their anticipated knowledge of the sector, including the work of Samarth and other programmes. These key informants included individuals who had a direct relationship with Samarth and those who did not in order to get an external perspective on the work of the programme. This latter group included representatives of other donor organisations, other government departments, and other 'apex' firms not engaged directly by Samarth.
- 9 A comprehensive review of dairy sector constraints is provided in Pant *et al.* (2017).
- 10 The role of the QulP studies in our evaluation and some of these challenges encountered are discussed further in an upcoming Centre for Development Impact Practice Paper co-authored with Bath SDR.
- 11 Factors included increasing production costs and poorly enforced milk quality standards which served to undermine incentives to improve the quality of production.
- 12 Depending on the timing of intervention close and our research, evidence indicated that these actors were still implementing improved practices between 12 and 18 months after the end of programme support.
- 13 *Jacobellis v Ohio* 378 US 184 (1964).

References

- Bath Social and Development Research (2021) (accessed 8 October 2021)
- Britt, H. (2013) *Complexity-Aware Monitoring*, Discussion Note, Washington DC: United States Agency for International Development
- Chen, H-T. and Rossi, P.H. (eds) (1992) *Using Theory to Improve Program and Policy Evaluations*, New York NY: Greenwood
- Chen, H-T. and Rossi, P.H. (1983) 'Evaluating with Sense: The Theory-Driven Approach', *Evaluation Review* 7.3: 283–302
- Chen, H-T. and Rossi, P.H. (1980) '**The Multi-Goal, Theory-Driven Approach to Evaluation: A Model Linking Basic and Applied Social Science**', *Social Forces* 59.1: 106–22 (accessed 23 August 2021)
- Fowler, B. and Lomax, J. (2021) 'Measuring what Matters: Monitoring and Results Measurement', in J. Ledgerwood (ed.),

- Making Market Systems Work for the Poor: Experience Inspired by Alan Gibson*, Rugby: Practical Action Publishing
- Itad (2019) **Samarth-NMDP Programme Final Evaluation Report**, Oxford: Itad and e-Pact (accessed 17 January 2022)
- Jenal, M. (2019) *Measuring Systemic Change in Market Systems Development – A Stock Taking*, Washington DC: United States Agency for International Development
- Mayne, J. (2011) 'Contribution Analysis: Addressing Cause and Effect', in K. Forss, M. Marra and R. Schwartz (eds), *Evaluating the Complex: Attribution, Contribution, and Beyond*, Brunswick NJ: Transaction Publishers
- Nippard, D.; Hitchins, R. and Elliott, D. (2014) *Adopt-Adapt-Expand-Respond: A Framework for Managing and Measuring Systemic Change Processes*, Briefing Paper, Durham: Springfield Centre
- Pant, S.R.; Acharya, M.; Sah, A.K. and Pandet, G. (2017) 'Present Status, Problems and Strategies for the Development of Cattle Farming in Nepal', Proceedings of the 10th National Workshop on Livestock and Fisheries Research in Nepal, Lumle, Kaski, Nepal, 5–7 March
- Posthumus, H.; Shah, R.; Miehlbradt, A. and Kessler, A. (2020) **A Pragmatic Approach to Assessing System Change**, Boekel: Hans Posthumus Consultancy, Springfield Centre, Miehlbradt Consulting, and Donor Committee for Enterprise Development (accessed 23 August 2021)
- Ruffer, T. (2012) **GEMS Results Measurement Handbook**, Hove: Itad (accessed 20 September 2021)
- Stame, N. (2004) 'Theory-Based Evaluation and Types of Complexity', *Evaluation* 10.1: 58–76
- Stern, E. et al. (2012) **Broadening the Range of Designs and Methods for Impact Evaluations: Report of a Study Commissioned by the Department for International Development**, Working Paper 38, London: Department for International Development (accessed 23 August 2021)