EVALUATION: OCTOBER 2017 PUBLICATION: DECEMBER 2019

Dairy producers in Rangpur queue to deliver their milk at the collection centre. Photo credit: Rachel Corner/Oxfam

MAKING MARKET SYSTEMS WORK FOR WOMEN DAIRY FARMERS IN BANGLADESH

A final evaluation of Oxfam's Gendered Enterprise and Markets programme in Bangladesh

DR MANU SAVANI & ALASTAIR STEWART

www.oxfam.org

OXFAM

CONTENTS

ACKNOWLEDGEMENTS

We would like to thank the staff of Oxfam Bangladesh for their support in carrying out this evaluation, with particular thanks to Mutashim Billah, Syeeda Poorna, Shafiqur Rahman and Ashish Kumar Bakshi.

Thanks are also due to Oxfam's partner organizations SEED, SKS and MMS. Oxfam acknowledges the smallholder farmers and wider stakeholders in Bangladesh who generously gave their time for this evaluation.

Fieldwork was conducted by Professor Jahangir Alam and Professor Mohd Kabir at the Bangladesh Agricultural University, ably assisted by Farzana Nizhum, Ismat Begum and their team of enumerators.

Special thanks also to Miranda Morgan, Kanako Yoshikawa and Simone Lombardini for their roles in the data analysis and report revision processes.

EXECUTIVE SUMMARY

The Gendered Enterprise and Markets (GEM) programme has enhanced the incomes and opportunities of poor households in Bangladesh. Women smallholder farmers (SHFs) were supported with a series of training activities that has improved dairy farming practice and promoted higher revenues and farm productivity. Rigorous evaluation techniques combined quasi-experimental analysis of project impacts through a household survey, along with in-depth community-level focus group discussions and stakeholder interviews. The results indicate project households enjoyed greater food consumption and undertook higher investment and spending than comparison households. Novel measures were employed to investigate the quality of income from dairy farming, and the results indicate this livelihood strategy offers predictable and timely income that smallholders value.

The project brought together women from seven districts across Bangladesh to form 84 dairy producer groups. GEM's approach focused on empowerment of women as market actors, through training on rights, assertivenes and leadership skills. This evaluation finds evidence of positive and significant impact on access to market services, participation in household and community decision-making, and intra-household bargaining over care duties. The results point to the beneficial impacts of GEM activities with the potential for spillover effects on local governance and engagement in community decision-making, with women more widely reporting having a greater voice in community planning and organizing.

There is no evidence of unintended negative consequences in the form of gender-based violence, despite the programme aiming to challenge and shift entrenched gender norms and a balance of power within the household. Rather, there is some evidence that the women who participated in GEM activities have lower tolerance for violence against women. Fieldwork was conducted following a period of excessive rainfall and flooding, which caused widespread damage and disruption to livelihoods during 2017. Despite the higher exposure of GEM households to these income shocks, the evidence points to positive project effects, which suggest the use of effective coping strategies to deal with the unexpected loss of assets and support recovery.

The GEM approach also targeted system-level changes to shift the policy and governance environment for smallholder farmers. Successes identified include the creation of a multi-stakeholder forum, and efforts to leverage access to formal finance for smallholders. Advocacy and partnership with Rupali Bank has extended credit to women dairy farmers with early evidence suggesting this has supported farm enterprise. Close working with the private sector helped facilitate the opening of milk chilling centres in two districts. System-wide effects are at an earlier stage of impact, but show promise for generating important advantages for dairy farmers in the future.

The evaluation highlights there is much to learn from the GEM approach. Key points for future programming include:

- Access to markets deserves further attention and resources, and programme design must explicitly apply a gendered lens to build on the achievements made so far. System-wide improvements in transport and connectivity, addressing concerns around safety and security, and redesigning physical markets as spaces equally welcoming to female and male entrepreneurs may go a long way in shifting entrenched gender norms.
- GEM has had a positive impact on women's power in household decision-making, particularly around economic issues, such as farming practice and who works outside the home to earn money. Yet, the evaluation also revealed that key areas of women's lives remain heavily influenced by gendered norms, notably around care work. Women's heavy and unequal responsibility for unpaid care work remains a barrier to them achieving, earning and leading outside the home. Future programmes may want to identify more closely where progress has been made, and what the next frontier is in order to support men and women to disrupt the status quo. Lessons from GEM respondents suggest that community norms and the ability to share care work publicly should be two such areas of focus.
- Despite being more flood-affected, GEM households were able to demonstrate advantages that positively affected their incomes and investments over the previous year. Project activities that promoted disaster-resilient models, such as model farms, fodder cultivation and preservation, and flood shelter may have supported these outcomes. Arguably, these results show that dairy farming can be designed and managed to be flood-resilient. What is less known, although the data is suggestive of this, is how far households deliberately used dairy farming as a resilience strategy. The evaluation finds that GEM households could benefit further from access to information to support preparedness. Future programme design would benefit from research into how borrowing and asset-selling behaviour sits within the household's own perceptions of resilience.
- Access to credit remains an area where more action is needed, and attitudes towards formal banking reflect a wariness among smallholder farmers. However, experience with the Rupali Bank pilot scheme shows strong signs of promise, both in making credit accessible and promoting rapid improvements to livelihoods. Priority next steps are to assess the feasibility of scale-up to meet existing demand, to give further attention to promoting demand by considering how access in *char* areas can be improved, and to address the negative perceptions of farmers towards formal banking.

1. ABOUT GEM

Gendered Enterprise and Markets (GEM) is Oxfam GB's approach to market systems development. The GEM approach facilitates change in market systems and social norms, with the aim of ensuring more sustainable livelihood opportunities for marginalized women and men. The GEM Department for International Development (DFID) AidMatch Programme (June 2014–February 2018) works within the soya, milk and vegetable value chains targeting women smallholder farmers in areas of poverty. Some 63,600 people (10,600 smallholder households) living in Zambia, Tajikistan and Bangladesh were expected to benefit directly from this programme through increases in household income, women having greater influence over key livelihood decisions within their households and communities, and engaging in livelihoods more resilient to shocks, such as natural disasters and market volatility.

The programme works with a combination of private sector, government and civil society actors and supports improved public services, the development of new business models and improvements to linkages between smallholder farmers and small and medium agri-businesses. This engagement is being deepened and strengthened to consolidate gains and scale-up impact.

1.1. GEM IN BANGLADESH

The GEM programme in Bangladesh was implemented under Oxfam Bangladesh's flagship REE-CALL programme (Resilience, through Economic Empowerment, Climate Adaptation, Leadership and Learning). GEM operated in seven districts across Bangladesh, with the project activities implemented by seven local partners. The project aimed to establish 84 producer groups for smallholder dairy farmers, and this was achieved during the first year. Building on these local networks, GEM aimed to deliver a suite of training and support covering assertiveness, rights and leadership skills, agricultural practice and disaster risk management, in order to achieve the three broad objectives.

The theory of action flowchart (below) was constructed with project partners, beneficiaries, Oxfam staff and local stakeholders, including the private sector and local government officials, in September 2017. It represents the intended pathways through which GEM activities were expected to lead to improved outcomes for women's empowerment, household incomes and resilience to natural disasters. The theory of action provides a framework to assess impact both at the outcome level and to understand key steps and intermediary outcomes that determined the overall success of the programme. All seven partner organizations were involved in the workshop to launch the evaluation design. For the purpose of the evaluation, three districts were to be investigated in depth

through the household survey and community-level focus group discussions: Gaibandha district (where SKS were the implementing partner), Rangpur (SEED) and Sirajgonj (MMS). These districts were selected through a participatory approach in consultation with local partners.

The theory of action for the GEM Programme in Bangladesh details the intended pathways from the project activities, outputs, outcomes and intended impacts. The theory of action is aligned to the three impact areas:

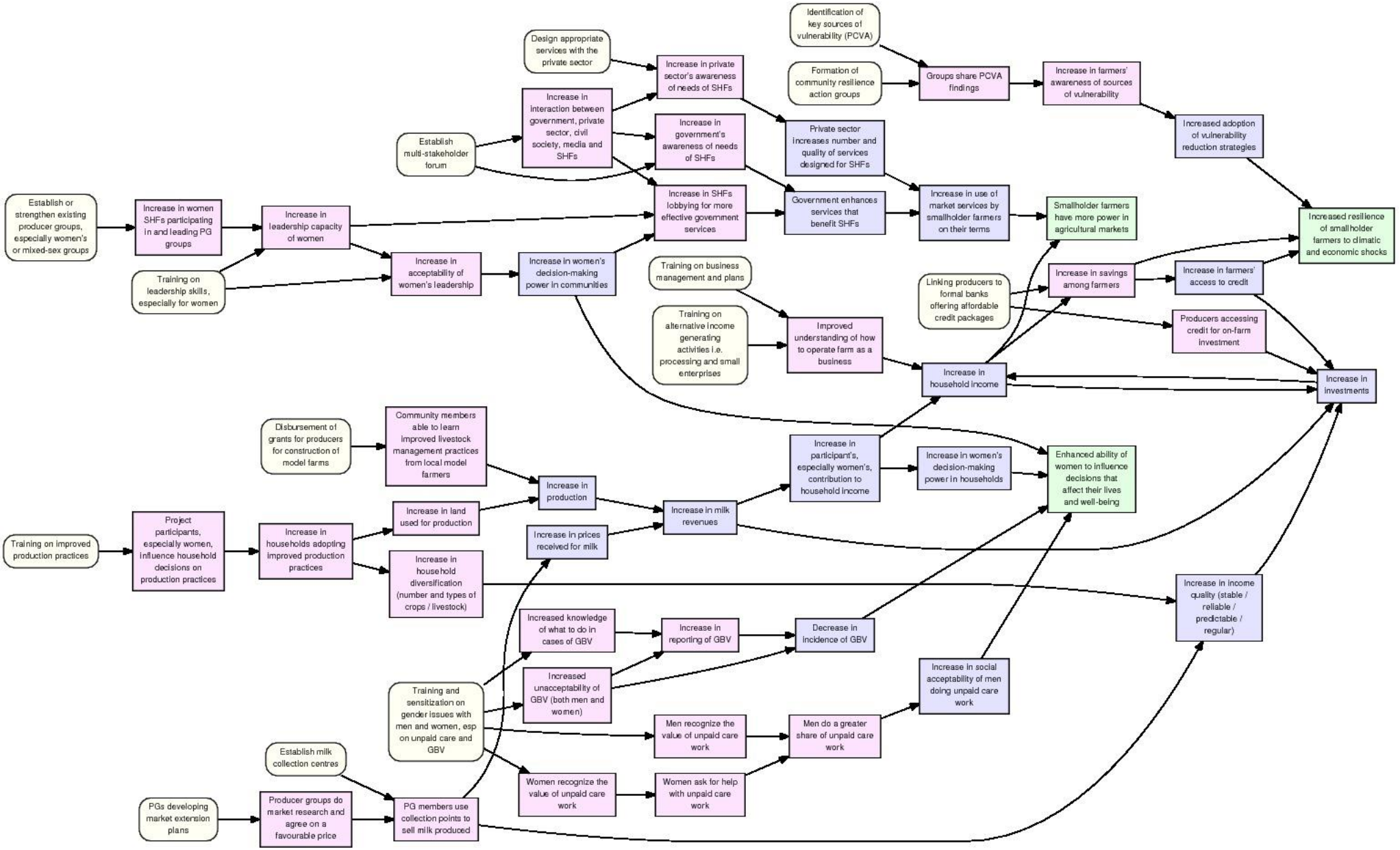
- Enhanced ability of women to influence decisions that affect their lives and well-being: this is achieved through the outcomes of increasing women's decision-making power at household and community level; increasing the social acceptability of men doing unpaid care work and challenging the acceptability of gender-based violence (GBV). The GEM programme in Bangladesh has sought to achieve this through training and sensitization on gender issues with men and women on unpaid care and GBV, training on leadership and assertiveness skills for women producer group (PG) members, and establishing and strengthening producer groups to increase women's participation and leadership opportunities.

- Smallholder farmers (SHFs) have more power in agricultural markets: this is achieved by increasing the use of market services (information, input supply, extension) by smallholder farmers and increasing household income and investments. The GEM programme sought to achieve this through linking smallholders and the private sector, establishing a multi-stakeholder forum and designing appropriate services with the private sector, training SHFs on improved production practices, agriculture as a business and value addition, and establishing milk collection centres.

- Increased resilience of smallholder farmers to climatic and economic shocks: this is achieved through the outcomes of increasing farmers' access to credit and savings including through a pilot programme with a commercial bank, and training and support for dairy farmers around livestock health, including strategies that would boost resilience in the event of natural disasters.

The theory of action diagram (Figure 1) visualizes the pathways from activities to outcomes and intended impacts.

Figure 1: GEM Bangladesh theory of action diagram



2. EVALUATION APPROACH

Why evaluate?

The purpose of the final evaluation was threefold:

1. To provide an opportunity for Oxfam, partners and a range of programme stakeholders to reflect on and learn from what and how change has happened because of the GEM programme.
2. To provide an opportunity for Oxfam GB and partners to learn more about select elements of GEM's theory of change within and across GEM countries.
3. To meet the requirement for accountability to DFID and programme stakeholders to measure the agreed impact and outcome indicators.

How to evaluate?

To achieve these goals, an evaluation approach was chosen that combined rigorous observation and analysis of changes at and across multiple levels (from individual to systems level) with maximizing opportunities for learning for a range of stakeholders, including the people we work with ('project participants'[1]), local market systems actors, programme partners, Oxfam and DFID. Based loosely on the Participatory Impact Assessment and Learning Approach (PIALA)[2], the evaluation aimed to combine rigour, inclusiveness and feasibility to assess the GEM programme and its contributions to change in several areas. The elements drawn from PIALA that differentiate this evaluation from others include:

- Stakeholders at all levels contributed to framing and designing the evaluation, and positioned the GEM programme theory of action within their larger theories of change.
- Mixed-methods data collection, analysis and triangulation to understand if the programme's desired changes occurred at multiple levels as well as how progress is enabled or challenged (see Box 1).

In addition, a process review was conducted by an external consultant to better understand the factors that enabled or constrained effective programme implementation and sustainability with a view to using this learning to improve future programming.

Limitations of the above methods and the fieldwork are outlined in Annex 5.

[1] Also referred to as 'beneficiaries'.

[2] Heinemann, E., Van Hemelrijck, A. and Guijt, I. (2017) 'Getting the most out of impact evaluation for learning, reporting and influence: Insights from piloting a Participatory Impact Assessment and Learning Approach (PIALA) with IFAD.' https://maintenance.ifad.org/documents/10180/91825437-ffa9-4451-a7ef-9dbd5efbeba0

Box 1: Complementary mixed methods

Both quantitative and qualitative methods were used to observe *if change occurred* in key outcome areas and make plausible claims of contribution to the observed results, as well as to understand *how change happens* over time.

1. Household survey and quasi-experimental analysis: Interviews were conducted with 747 women, of which 300 were directly supported by the project in the dairy value chain (referred to as the intervention or project group) and 447 were identified in villages were GEM project partners did not conduct any activities (the comparison group). For more information on the survey method and sampling strategy refer to Annex 1.
2. Focus group discussions (FGDs) in communities: FGDs were conducted in four communities supported by the programme. Participatory FGD exercises were designed and conducted for different groups in each community: (i) village leaders; (ii) female PG members; (iii) male PG members or family members of PG members (mostly husbands); and (iv) female PG members and non-members. For more information refer to Annex 2.
3. Key informant interviews (KIIs) with market systems actors: During the design phase theory of change workshop, an outcome harvesting approach was used to identify potential systems-level outcomes. KIIs were then conducted with a range of key stakeholders to verify these outcomes. For more information on the OH approach and list of interviewees refer to Annex 3.

What to evaluate?

Like the GEM programme, the evaluation aimed to provide a holistic view of change and shed light on the multiple levels (from individual to systemic) and interconnected dimensions (economic, social, personal and political) of change that enable and sustain women's increased participation and power over time.[3]

The evaluation was designed to investigate if and how the GEM programme might have contributed to its intended outcomes – not only in the lives of individual women smallholder farmers targeted by the programme (by enhancing their livelihoods or changing the way they influence decisions in their households, for example), but also in changes in their communities and the larger market system. It also sought to capture any potential unintended outcomes of the programme, especially those that are well known to manifest as negative outcomes in women's economic empowerment programmes (such as increased gender-based violence or amount of work).

The specific evaluation questions guiding the evaluation were primarily informed by the four impact indicators of the GEM programme log-frame. Table 1 shows the source of the main evaluation questions with links to the question numbers in the findings below.

3 See Oxfam's fundamental principles for achieving sustainable, positive change for Women's Economic Empowerment, pp.23–24 in Kidder, T. et al. (2017) *Oxfam's Conceptual Framework on Women's Economic Empowerment* (https://policy-practice.oxfam.org.uk/publications/oxfams-conceptual-framework-on-womens-economic-empowerment-620269)

Table 1: Evaluation questions			
	Global impact indicators	**Locally identified questions**	
Changing market systems	1.1 Number and type of changes to local and national policies and practices related to economic opportunities of smallholder farmers, especially women	What has been the policy influence of the GEM-sponsored multi-stakeholder forum?	Q1
		Has GEM increased access to market services and credit?	Q2, Q3
Enhancing smallholder power in markets	2.1 Average annual income of beneficiary smallholder farmer from dairy value chains	Has GEM improved productivity and incomes?	Q4, Q5
	2.2 Perceptions of beneficiary smallholder farmer on the quality (timeliness, predictability, stability, purpose) of their income	Do women smallholders receive high quality income from dairy farming?	Q6
Increasing women's economic leadership	3.1a Perceptions of women in targeted communities on their ability to engage in decision-making processes at household level, especially economic and care work responsibilities	Has GEM empowered women to engage in decision making in their households and communities?	Q7, Q9
		How has the heavy responsibility for unpaid care work changed for women?	Q8
	3.1b Perceptions of women in targeted communities on their ability to engage in decision-making processes at community level	Can women's economic empowerment (unintentionally) lead to increased rates of gender-based violence?	Q10
Resilience		Has GEM supported smallholder farmers to cope with extreme weather events and natural disasters?	Q11

3. RESULTS

The results below present a combined analysis of the quantitative and qualitative data to address the key evaluation questions. It should be noted that all survey respondents were women. This approach enabled greater statistical power in answering the questions of the programme's impact on women's incomes and empowerment.

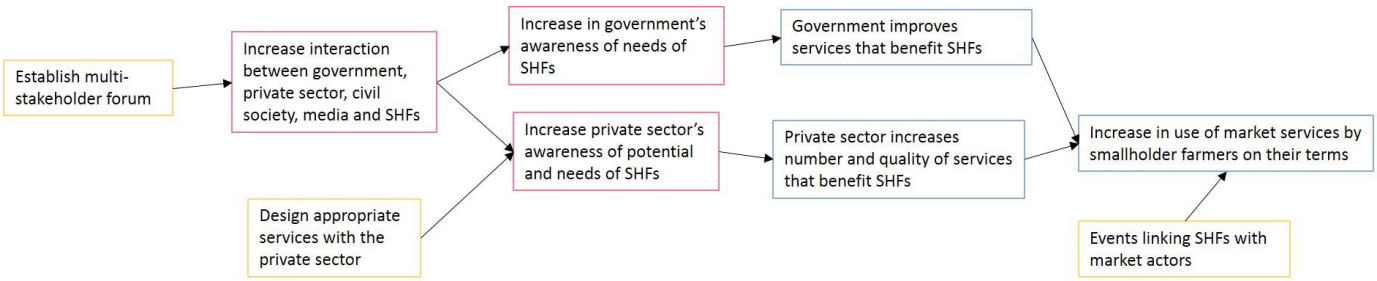
When reporting the results from the quasi-experimental impact evaluation analysis, the results refer to the average difference between women living in communities where the project was implemented (the 'intervention group') and the matched women in communities where the project was not implemented (the 'comparison group'). As women in the comparison group represent an estimate of what would have happened to the project participants in absence of the project, the difference between the two groups represents an estimate of the project impact. In general, the results are reported as 'significant' if they have a p-value of less than 0.05, but results with higher p-values (between 0.05 and 0.1) will also be noted as 'weak evidence'; the lower the p-value the more confident we feel that the measured estimate reflects the true impact. Results with a p-value of more than 10 per cent are not considered to be statistically significant.

3.1. INFLUENCING MARKET SYSTEMS

The project sought to deliver changes to the number and type of local and national policies and practices related to the economic opportunities of smallholder farmers, especially women. The GEM theory of action assumes the creation of the multi-stakeholder forum will lead to increased interaction and therefore understanding among the dairy value-chain actors, resulting in improved awareness of the needs of smallholder farmers. This in turn should lead to the development and roll out of improved market services (information, input supply, extension) that smallholder farmers have access to and use, making them more informed and productive market players. The findings in this section relate to this pathway of change and integrate findings around the role of the multi-stakeholder forum in policy influencing.

The evaluation relied on key informant interviews to gather evidence of the nature and extent of change GEM had brought about in the broader market system. Three key findings emerged from in-depth interviews with partners, government officials, private sector actors and dairy farmers. These results indicate positive effects from GEM activities, particularly in relation to increasing farmers' access to services and to credit, and influencing policy through the multi-stakeholder forum.

Figure 2: Snapshot of 'Influencing Market Systems' pathway of theory of action



Question 1: What has been the policy influence of the GEM-sponsored multi-stakeholder forum?

A positive outcome of the project was the establishment of eight multi-stakeholder forums, including one at national level, which successfully enabled representation and advocacy of smallholder dairy producers' interests. Achievements resulting from the multi-stakeholder forum mentioned by members included improved provision of services, especially in hard to reach areas, access to credit, and influence over the National Dairy Development Policy. The forum had identified four priority areas for the policy: access to finance; access to extension services; inclusive sourcing and responsible business; and care work. Discussions on the National Dairy Development Policy were collated from all provincial forums and relayed to the appropriate government agencies.

Despite optimism from some stakeholders, notably Oxfam Bangladesh, that the policy had the scope to improve women's participation in the market and relieve a number of market failings, other stakeholders were less hopeful. The Department of Livestock Services office in Dhaka highlighted a number of gaps in the existing policy. While it may have considered many aspects of gender imbalance and the role of women as dairy producers, often these considerations were not perceived as realistic or the priority issues for women. For example, strategic issues around the role of middlemen in the market and how this affects women's engagement, as well as the accessibility of the formal milk market, were feared to have 'not been addressed' in the draft policy. Reports from theory of change workshop participants, who also represented a diversity of stakeholder voices, echoed these strategic issues left out of the National Dairy Development Policy as key constraints to women farmers. The draft policy is still waiting for final approval at the time of the evaluation. Of interest is the evidence that voices within the government are willing to critique their own approach, and this bodes well as an opportunity for the multi-stakeholder forum to lobby for improved policy and government leadership.

Question 2: Did smallholders gain increased access to market services?

GEM activities served to increase the number of stakeholders providing essential input services, including feed, medicines, vaccinations and

artificial insemination (AI). Household survey data shows statistically significant effects from the project on access to veterinary and AI services and other inputs (see Table 2).

Focus group discussions reflect some of the themes raised in interviews. When asked about livestock services, women belonging to producer groups were both familiar with a range of different services – including training, vaccination, AI, doctors, local service providers trained by GEM partners, and the district livestock officer – and reported that all their producer group members were able to access these services when required. They appreciated the facilitation efforts and information from Oxfam's partner organizations. Village leaders also had a sense that producer groups were 'well informed' of new technologies, through training on farming practices and financial support, and believed that women farmers outside of the GEM producer groups were 'inspired' to replicate such practices in order to improve their own lives, suggesting the potential for spillover effects of GEM interventions on the wider community.

Table 2: Project impact on access to market services

Proportion of respondents answering 'yes' when asked if they:	GEM (n=296)	Comparison (n=442)	Project effect (% points)	Statistically significant
Accessed veterinary or artificial insemination services in the past year	73%	19%	+ 54 (0.04)	Yes
Purchased inputs for dairy farming in the last 12 months	92%	63%	+ 29 (0.04)	Yes
Received inputs for free in the last 12 months	29%	12%	+ 17 (0.04)	Yes

Notes: For each hypothesis test of group means, the difference between project and comparison groups was statistically significant with p<0.01.

Data gathered through in-depth interviews highlighted the roll out of government vaccination services and training of female service providers at the local level. Both have seen high demand and use. Feed and fodder banks, mostly run by women, are directly linked to private dealers and wholesalers. This business model has the scope to be promoted and replicated more widely as a means of connecting remote communities to mainstream markets. Government has provided AI services, and private AI technicians have been trained in harder-to-reach areas. The data indicate that companies such as ACI, Mamus Medical Hall, Seed Vaner, Masum Traders and Sagar Traders have expanded their operations in the evaluated districts because of increased demand from smallholders.

Despite these examples of positive change, the river islands (*char* villages) remain constrained in their access to such services due to weak and inadequate infrastructure and the logistical challenges of reaching markets and larger settlements by boat. These constraints were exacerbated during 2017 because of excessive rainfall and flooding.

While communities broadly praised the improved access to services, producer group members from *char* villages were often reliant on project activities for services. For example, they highlighted the difficulty of receiving timely visits by doctors, and the challenges of reaching the district livestock offices because of the distance and travel time from these villages to the mainland. Despite project activities to support local service providers who are based in the communities, evidence from GEM participants suggests there remains unmet demand, particularly in more remote areas.

Another valuable development was the establishment of a milk chilling centre in Rangpur district. However, the evidence suggests it has not yet been used extensively by GEM participants (seeming to reach only 48 producer group members). It is possible the centre has greater reach among the wider dairy sector, but data were not available to corroborate this. The finding suggests that while there are promising signs of impact, especially for some producers, it may be too early to see widespread adoption. It also highlights that time is needed for policy and market changes to become embedded in practice and to produce more visible impacts.

Question 3: Did smallholders gain increased access to credit?

Evidence suggests there has been some increased access to credit for smallholder dairy producers through a pilot credit scheme with Bangladesh Bank. Reports from Oxfam Bangladesh indicate 100 loans were distributed from Rupali Bank Ltd and National Bank Ltd. Rupali Bank reports that 30 loans were provided in Rangpur with another 30 planned. Stakeholder interviews indicate that demand for the pilot credit scheme is outstripping the funds made available by Bangladesh Bank. While there is potential for wide-scale reach, the pilot is at an early stage and repayments have not yet started.

The role of GEM in this process of improving access to credit is relatively small scale in Rangpur, but potentially important in terms of bringing in the involvement of government-owned banks, such as Rupali Bank, and private sector actors, such as National Bank (in Kurigram district, which was not directly evaluated in this exercise). Through GEM, Oxfam has promoted the formation of local networks and groups, providing skills in financial literacy, and connecting them to private sector actors. Rupali Bank indicated this was a helpful intervention for them, as it opens up a new market of eligible and reliable clients. At this level, the influence of GEM activities on system-wide processes has been positive, and future impact could be greater still.

The impact at household level is less clear at this stage. No significant differences were detected between the project and comparison groups on amounts borrowed in the past year or the proportion of households who borrowed at all. The survey uncovers, however, an interesting contrast between past borrowing experience and outlook. When asked if they would be able to borrow 40,000 taka for a hypothetical business

opportunity, 57 percent of GEM households reported they would, against 40 percent in the comparison households. The data appear to indicate higher confidence in access to credit among the project households. One explanation for why this was not translated into differences in actual borrowing might lie in the circumstances of 2017, notably extreme flooding, which may have hampered new business investment and appetite for expansive ventures.

Among those who did report taking out loans over the past 12 months, only 19 (4 percent) of the comparison group and 2 (1 percent) from the project group relied on commercial credit. Respondents from both project and comparison groups were more likely to mention microcredit institutions and informal borrowing as sources of credit than commercial banks.

These findings are reflected at the community level, where mixed feelings were expressed towards formal bank loans. Focus group discussions highlighted a degree of wariness of formal banks, especially among those who had not taken a loan in the past year, both in terms of the likelihood of being accepted for a loan ('we never get loans from the formal bank due to living in the *char* area') and navigating the process of applying ('we don't want to. [It's a] very complex procedure to get the loan… [we have a] lack of collateral, procedure of taking a loan at the bank is time consuming, and sometimes we need to give a bribe.'). Such pessimism contrasted with group discussions in one community where women producer group members had recently taken a loan from Rupali Bank for livestock production. They reflected on the experience: 'The loan was so useful for us. We have now increased our farm and our income. But the number of us is very few. All of us want to get a loan from Rupali Bank'.

'…we don't want to. [It's a] very complex procedure to get the loan… [we have a] lack of collateral, procedure of taking a loan at the bank is time consuming, and sometimes we need to give a bribe.'

~ Female PG member

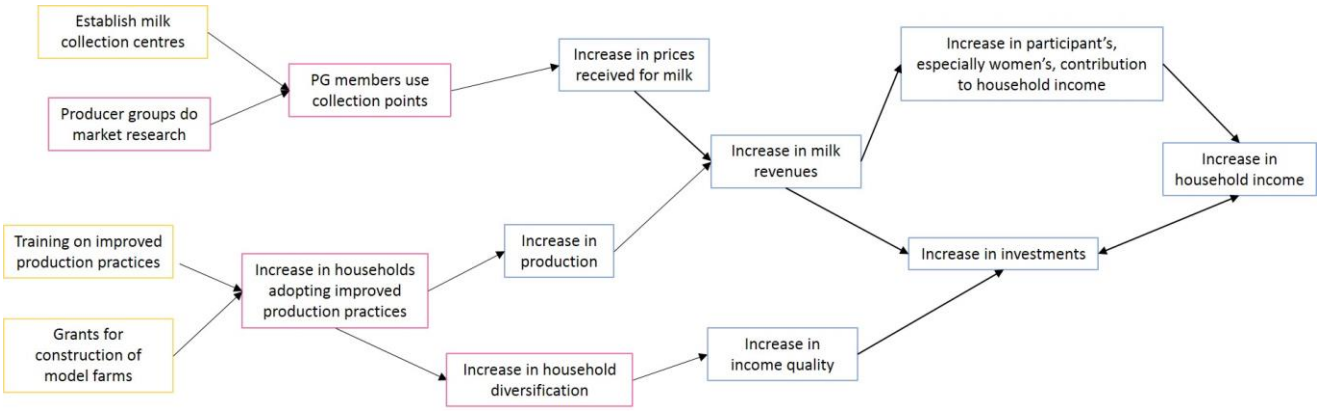
3.2. ENHANCING SMALLHOLDER POWER IN MARKETS

The theory of action set out a number of pathways by which GEM interventions would lead to increased smallholder incomes and power in markets. For example, various training activities were expected to lead to take up of improved farming practices, which would deliver higher milk production and revenues. The establishment of milk collection centres was intended to make it easier for women smallholders to connect to local markets and secure a more reliable source of income.

The evaluation relied on household survey data and quasi-experimental impact evaluation methods to assess the nature and extent of change GEM had brought about in incomes for women smallholders. Three key findings emerge. The project supported higher incomes from dairy farming, linked to wider take up of good farming practices. GEM households reported higher overall food consumption and income. Further, income from dairy farming was reported to be of high quality,

judged by its timeliness and predictability.

Figure 3: Snapshot of 'Smallholder Power in Markets' pathway of theory of action

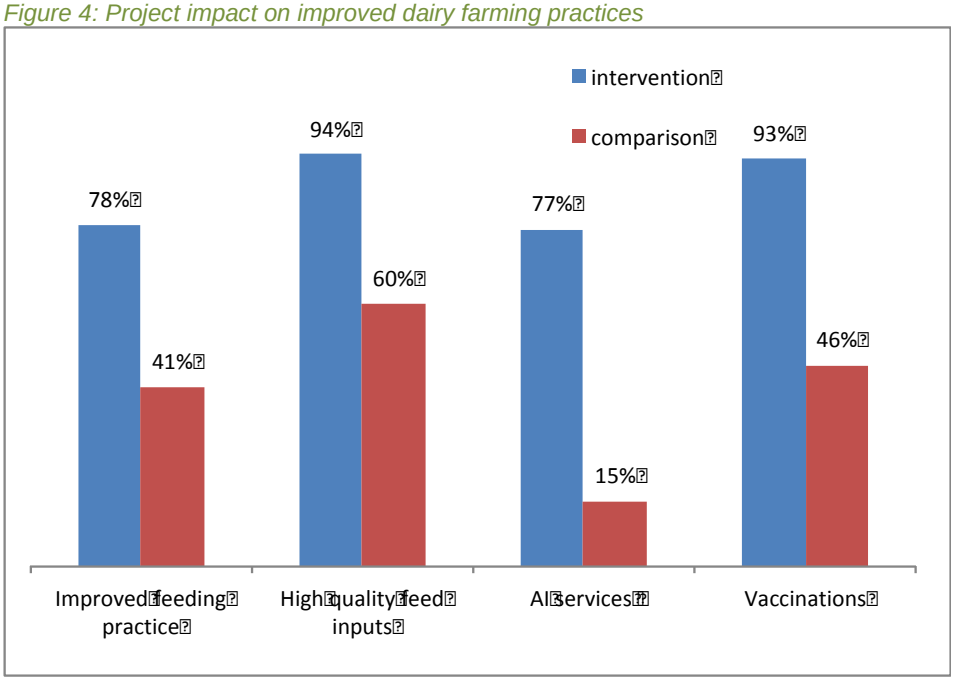


Question 4: Has GEM improved productivity and income from dairy farming?

Households who received GEM support reported significantly higher milk production and sales, and ultimately higher income from the dairy value chain: approximately 24,300 taka more per year. To put this finding into context, the dairy income differential is one-third of average household spending over the year across all project households surveyed. The headline message is that the GEM project had positive impacts on dairy income, and accounted for a substantial proportion of household income in 2017.

The survey asked households about their farming practices. Households who received training and support through GEM project activities were more likely to have implemented good practice, as reflected in Figure 4. For example, GEM households were more likely to have vaccinated their cow in the past year (93 percent against 46 percent in the comparison group), used recommended feeding practices such as concentrated feed (94 percent against 60 percent), used separate feeding mangers (78 percent against 41 percent), and employed AI services over the past three years (77 percent against 15 percent). Each of these findings represent positive project impact on farming practice, which the theory of action links to productivity and income in the dairy value chain.

Figure 4: Project impact on improved dairy farming practices



Notes: Improved feeding practice was measured as a positive response to 'ever used a separate feed and water manager in the last 12 months'. High quality feed inputs were measured as a positive response to 'ever given your livestock concentrated feed in the last 12 months'. AI services were measured as a positive response to 'ever used artificial insemination over the past 3 years'. Vaccinations were measured as a positive response to 'vaccinated your cow in the past 12 months'. Hypothesis tests on group means are statistically significant with p<0.01 for each indicator. See annex for further details on number of observations in each analysis.

In line with the evidence that the project helped improve farming practices, the household survey further identifies that GEM households recorded higher dairy output and enjoyed greater revenues from their produce. As Table 3 demonstrates, the project led to more successful dairy farming in terms of output, sales, revenues and, ultimately, income.

Table 3: Project impact on dairy production, sales, revenues and income

	GEM	Comparison	Project effect	Statistically significant?
Dairy production (litres, in the past week)	19.5	7.3	**+ 12.2** (2.34)	Yes
Quantity of milk sold (litres, in the past week)	16.3	3.7	**+ 12.4** (2.35)	Yes
Value of milk sold (taka, in the past week)	645	177	**+ 468** (106.95)	Yes
Revenues per milking cow (litres, in the past week)	469	138	**+ 331** (72.74)	Yes
Dairy income (taka per year)	33,547	9,190	**+ 24,357** (5,736.43)	Yes

Notes: For each hypothesis test comparing group means, the difference between project and comparison groups was statistically significant with p<0.01. See annex for further details on number of observations in each analysis.

Question 5: Do women smallholders supported by the GEM programme report greater household income?

Women from GEM households reported significantly higher incomes than comparison households when measured by two important indicators: food consumption and annual household spending. The difference in reported food consumption over the previous week and expenditure on household needs over the past year were statistically significant between project and comparison groups. Spending by project households was also significantly higher in forms of investment, such as education, farming equipment and off-farm activities, such as small enterprise (see Table 4).

Some indicators provide a more mixed impression of impact, however. The project had no effect on changing household wealth. The nature of assets held varied across the two groups of households, for example comparison households were more likely to own assets such as a smartphone (42 percent compared to 31 percent in project households) and a TV (23 percent compared to 12 percent in project households). On the other hand, project households were more likely to own solar panels (41 percent compared to 23 percent in the comparison group). The evidence suggests that while incomes had improved, this has not yet delivered a change in overall wealth status. The outcomes may reflect the challenges of acquiring assets during a year of extreme flooding. An alternative interpretation is the relatively short timeframe of the evaluation may not provide enough time for wealth accumulation to be accurately measured, and it is therefore a less meaningful indicator of impact in the short term than household spending.

Table 4: Project impact on incomes, investment and wealth

	GEM (n=296)	**Compari-son (n=442)**	**Project ef-fect**	**Statistically significant?**
Food consumption (Taka per day per capita over past week)	102	91	**+ 11** (3.53)	Yes
Food consumption (Log)	4.54	4.44	**+ 0.1** (0.04)	Yes
Annual expenditure (Taka over past 12 months)	73,574	65,987	**+ 7,506** (6,648.20)	No
Annual expenditure (Log)	10.76	10.54	**+ 0.21** (0.09)	Yes
Investment (Taka over past 12 months)	27,524	22,847	**+ 4,677** (3,430.34)	No
Investment (Log)	9.38	8.88	**+ 0.5** (0.17)	Yes
Change in wealth (Index)	0.01	0.05	**- 0.04** (0.04)	No

Notes: indicators are statistically significant when variables are expressed in log-arithmic terms. Real values are presented here to give a sense of magnitude of incomes in the sample.

Another puzzle in the findings is that while GEM households reported spending more over the previous year, expenditure over the previous month showed no significant difference with comparison households. The disparity between these different indicators might reflect the categories of expenditure being different across monthly and annual household spending, and possible measurement issues around recall. The data might also reflect seasonal patterns or recent changes in levels of spending and ownership of items in response to household shocks. If monthly expenditure is a relatively restricted snapshot of a particular period of time, then the annual figures are arguably a better reflection of differences between the project and comparison groups.

The impact monitoring framework also asks for results disaggregated by the sex of the household head. Across the household survey as a whole 90 percent have a male head of household (87 percent in the project group and 91 percent in the comparison group). Considering female-headed households restricts the sample to 77 observations, results are therefore indicative only. On this basis, evidence is suggestive that female-headed households benefited through GEM, and report higher dairy income and annual household expenditure (both findings are statistically significant). Weaker impacts are found on monthly spending, and no differences are detected between project and comparison households on food consumption or investment (see Annex 1 for more detail).[4]

The overall message from the household survey data is that GEM improved incomes overall, which translated into higher annual spending and food consumption. However, project households do not spend more at all times of the year, and their income advantages have not delivered faster change in wealth.

Question 6: Do women smallholders receive high quality income from dairy farming?

To explore the 'quality of income' from the dairy value chain, timeliness and predictability of income were probed further in the survey. Women from GEM households reported higher quality income from dairy farming when assessed by stability and frequency of income; but there were no differences with the comparison group when quality of income was assessed by predictability, timeliness or sufficiency. Across all households surveyed, incomes were not sufficient to prevent selling assets to cope with household needs during 2017. These three broad findings are expanded below (see annex for further detail).

Stability was measured by the number of months in a year that at least one cow was providing milk. GEM households reported five months of milk production over the year, compared to four months reported by comparison households, and the difference was statistically significant.

Frequency of income was measured by how often households received

[4] The sample for statistical analysis falls to 77 when only female-headed household are investigated and the matching methods are applied, which is too small to allow definitive conclusions to be drawn.

their income from dairy, including daily, weekly, monthly, seasonally and annually. More frequently received income was expected to allow households more predictability of their income, and facilitate their household budgeting and planning. The data provide mixed results against this indicator of quality of income. On the one hand, a significantly higher proportion of GEM households reported receiving their dairy income on a daily or weekly basis over the year. On the other hand, however, across both project and comparison groups, the proportion of smallholders who reported this frequency of income was low: only 1 percent of comparison households and only 7 percent of GEM households.

Village leaders highlighted the importance of 'a regular income source' for income and social mobility.[5] The findings from the household survey indicate that despite the improvements from GEM project activities, there is considerable scope for further improving the stability and frequency of income from dairy farming.

A second finding is that GEM households were not significantly different from comparison households in terms of predictability of income or timeliness. In both sets of households, 43 percent reported that the amount of income received was as expected, and over 80 percent in both groups received income when they expected it. Predictability has scope for improvement in both comparison and project households.

The findings on timeliness are positive overall, in that the majority of households had a reasonable degree of certainty around the timing of income. Focus group discussions corroborate this finding. Dairy producers in *char* areas indicated that selling to a milk collection centre was an important development in raising the predictability of income where there was no market and a boat ride was required to transport dairy produce: 'We sell to the milk collection centre; the local market is very far from here…before we could hardly sell the milk. We had no buyer.' In mainland communities, the milk collection centre was less important if there were local markets nearby, or competitive middlemen who could offer a higher rate and visit women at their homes. Producers were in broad agreement that dairy represented a reliable source of income, with payments delivered on time for the 7 to 8 months of the year they were able to generate income from milk.

'We sell to the milk collection centre; the local market is very far from here…before we could hardly sell the milk. We had no buyer.'

~ Female PG member

Future research could delve further into how programme design could focus on these aspects of high quality income streams, and investigate the channels through which more predictable and frequent income streams can affect enterprise planning, savings behaviour, and coping strategies for unexpected income shocks. Qualitative evidence supports the idea that certainty around dairy income is highly valued by smallholders, and further work could investigate how this certainty affects household planning and foresightedness, and the wider potential benefits for resilience and enterprise.

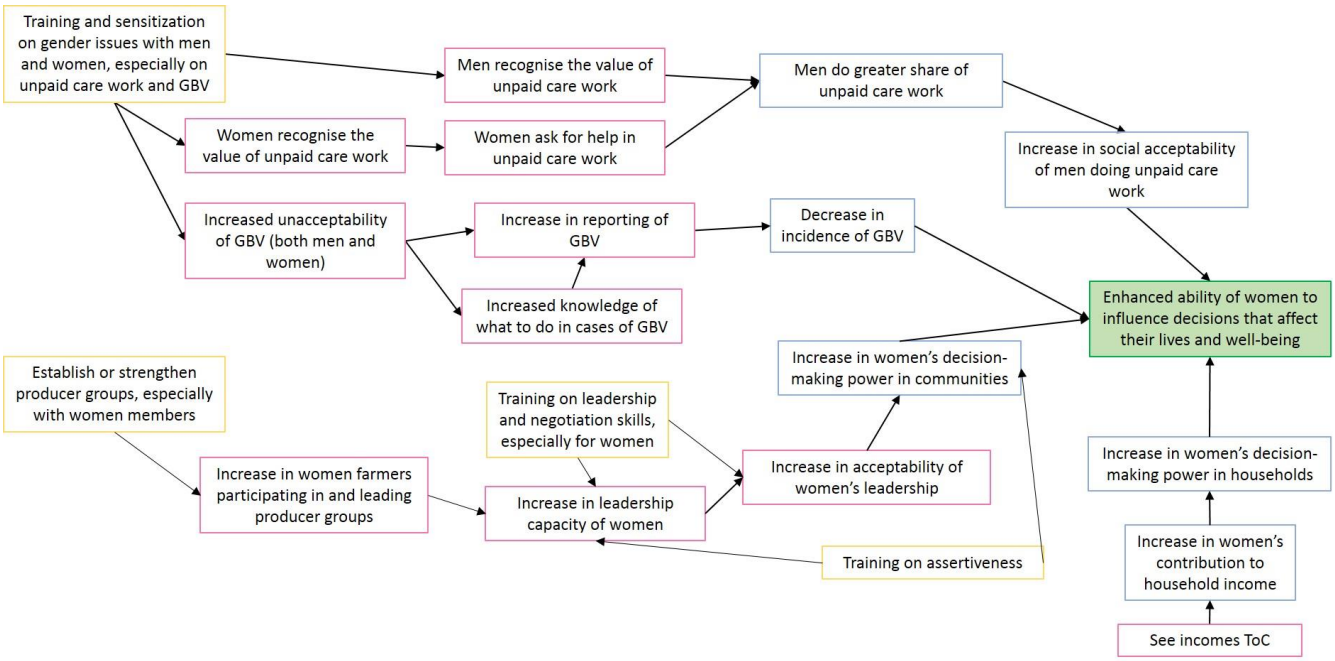
[5] From focus group discussion with village leaders, October 2017, using a 'ladder of life' research tool to investigate social mobility and community perceptions of poverty.

Finally, income sufficiency was generally low across all households surveyed. A sizeable proportion of households in both groups reported having to sell assets in the past 12 months to cover household expenses, with a slightly higher proportion in the project group (46 percent compared to 42 percent; the difference is not statistically significant). These results might reflect the fact that both intervention and comparison groups remain classified as poor households unable to claim sufficiency of income. The sufficiency indicators are consistent with the relatively low savings rates reported in the survey: half of all households reported having no savings at the time of the survey, and across the other half of households the average amount of savings was approximately 16,500 taka. Comparing across groups, project households had significantly lower savings than comparison households. Further, in a year where Bangladesh experienced extreme flooding, it was to be expected that many households responded to weather shocks through coping strategies such as drawing down on savings. The impact of the floods is discussed in Section 3.4.

3.3. INCREASING WOMEN'S ECONOMIC LEADERSHIP

The GEM approach aimed to empower women to take a stronger role in decision making within the household. Project activities, such as the use of rapid care analysis (RCA) tools and training, were designed to improve assertiveness and foster leadership. The evaluation sought to assess how well these activities had translated into an increase in women's participation in household decision-making. Results show that the project had a number of positive impacts on increasing women's participation in decision making, both in important day-to-day household matters and community-level decision-making. The evidence also suggests women were increasingly able to challenge the status quo by actively asking for men to share the responsibility for care work, and the project effected change through increased support from men in care work.

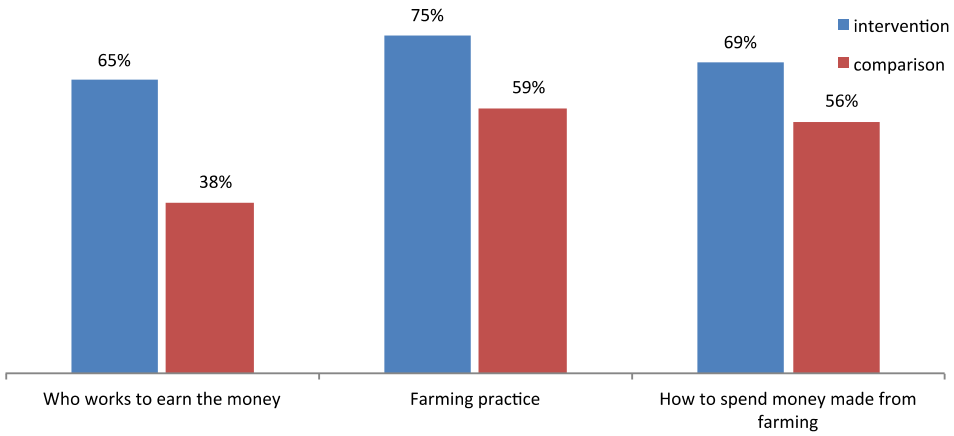
Figure 5: Snapshot of 'Women's Economic Leadership' pathway of theory of action



Question 7: Has GEM empowered women to engage in economic decision-making in their households?

Project impacts are identified for women's participation in decisions relating to 'who works to earn money for the family', 'farming practices' and 'how to spend money made from farming' (see Figure 6). In each case, women who were involved in GEM report a higher degree of participation within household decision-making in these important areas, and the difference was statistically significant.

Figure 6: Project impact on participation in household decision-making



Notes: Participation is recorded if respondent reported being involved in household decision-making on various issues. Hypothesis tests on group means are statistically significant, with p<0.01 for each of the indicators in this figure. See Annex 4 for further details on number of observations in each analysis and full set of decision-making issues.

The evaluation also sought to differentiate between women's participation in a household decision and, for those respondents who reported they did not take the decision alone, her perceived ability to influence the outcome of that discussion. Tellingly, perceived influence remains higher in the intervention group than the comparison group. However, the size and strength of project effects is lower when it comes to perceived *influence* over a decision area than participation in the discussion itself. For example, 75 percent of women in the project group felt able to participate in decisions about farming practice. Among those who take the final decision jointly with others in the household, 69 percent believed they could influence that decision. This figure compares favourably with the 60 percent of the comparison group who felt they could influence the decision, but the project effect is both smaller (9 percentage points, compared to 16 percentage points for participation) and weaker in terms of statistical significance ($p<0.10$, compared to $p<0.01$ for the participation outcomes). What the data indicate is that there remains progress to be made in empowering women and shifting intra-household dynamics to allow women to be more influential in economic matters.

Qualitative data broadly confirms this finding. Women PG members were asked to discuss a vignette about a married couple making a decision on how to spend their money. In the scenario, the wife wants to pay for artificial insemination to get an improved calf, while the husband wants to buy a bicycle. When asked how easy or difficult it would be for the wife to go ahead with AI without her husband's support, many female respondents suggested it would be easy: 'We know how to motivate our husbands…we can manage the situation by explaining the priority of our investments.' However, there was some recognition that 'you cannot decide alone…there will be no harmony in the home'. Some women were less optimistic, suggesting that there could be some difficulties in the wife acting without her husband's support, and referring to a need for 'permission' and recognition that ultimately husbands are the head of the family.

'We know how to motivate our husbands…we can manage the situation by explaining the priority of our investments.'

~ Female PG member

When faced with the same scenario in a separate group discussion, husbands of PG members in one community indicated that it would be easy for the wife to go ahead with her spending decision, explaining 'we have full confidence of our wives' decisions'. Male participants believed that couples need to work together and that the husband could not 'take the decision alone'. The men referred to the benefits of AI and the GEM partner organizations, indicating they were well familiar with GEM activities and the content of training provided. It is plausible that this group of men have also been affected by the project and their focus group discussion reflects a willingness to empower their wives in farming and business. What is less clear is how widespread these attitudes are. There is also the potential for social desirability bias to affect the content of responses given to the research team.

Other discussions, both at the theory of change workshop and in communities, highlighted the issues around women taking their milk to sell at markets. This task is generally perceived to pose a risk to women's security, linked to unreliable transport arrangements over long distances and challenging terrain (for example the river islands or *char* settlements), and markets outside the village largely lacking social areas and toilet facilities for women. Village leaders spoke of the ongoing 'social bindings' that meant women could not move easily, or get a job easily, or be flexible to work in any place or environment. When asked why some women remain trapped below the poverty line, village leaders mentioned that women 'need a safe environment', but were pessimistic about being able to provide this. Husbands of women who participated in GEM projects mentioned the negative effects on the reputation of a man whose wife goes to the market: 'Society will not agree. Because why do women go to the market for selling milk since their husbands are present in the world. Why will men stay at home to do housework?' [everyone laughs]. This excerpt from a focus group discussion suggests the concept of women taking milk to the market is akin almost to a complete role reversal of duties, with the response from other men in the group suggesting it is a laughable idea.

'Society will not agree. Because why do women go to the market for selling milk since their husbands are present in the world. Why will men stay at home to do housework?' [everyone laughs].

~ Husband of GEM producer

No significant project effects are found on either participation or influence in decisions relating to care work and family planning. These findings arguably reflect the entrenched social norms on care work and reproductive rights. The majority of women who do not take decisions on family planning by themselves, believe they can influence decisions on family planning either to some extent or to a large extent. In the comparison group, 75 percent of women reported having this influence. Although this proportion was higher at 81 percent of women in the project group, the difference was not statistically significant. Reproductive rights were not an outcome directly targeted by the GEM programme, but these data were collected in order to better understand the potential spillover effects of economic empowerment on other important areas of decision making in women's lives. An emerging message from this evidence is the sizeable minority of women who feel they cannot influence family planning decisions, and this is clearly an area for future efforts linking economic empowerment with personal empowerment.

The overall conclusion is that while the project has had a beneficial impact on participation in economic decision-making, longstanding norms and a traditional, gendered division of labour around household duties, farming and business practices remain. It is beyond the scope of any one project to fully address and overturn gender imbalances. The GEM project has contributed to the necessarily long-term process of empowering women through direct training and support of women, and by framing and facilitating debates on fundamental issues of gender equality. Such debates include, for example, the evolving and dual roles of women within the family and as entrepreneurs. GEM has also provided and encouraged demand for complementary services and investment essential to create safe and equal opportunities for women farmers so that they might assume leadership roles in the household and local markets.

Question 8: How have women's responsibilities for unpaid care work changed?

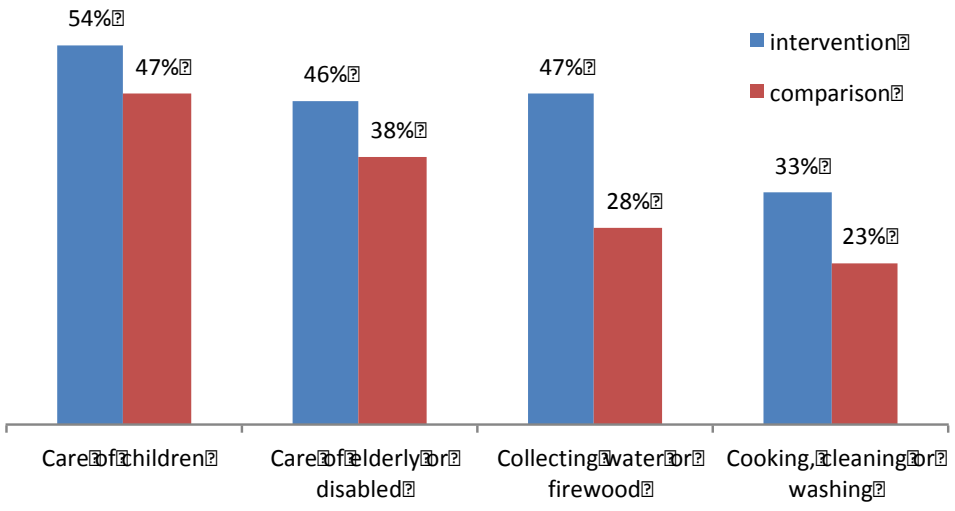
As mentioned in Question 7, women in the project group do not report much of a difference in household decision-making on unpaid care work. Yet, the actual sharing of care work at home shows marked differences between the GEM and comparison households. Men are reported to give more support to women in the project group across a range of care duties, including taking care of children, taking care of elderly or disabled family members, collecting water and firewood, and cooking, cleaning or washing (see Table 5). These results are arguably linked to the finding that women in the intervention group were significantly more likely to ask for help with such household duties (Figure 7). To better assess attitudes, household survey respondents were asked whether they agreed with the statement: '*If the wife is working outside the home, the husband should help her with household chores*'. Among GEM respondents, 95 percent of women agreed or strongly agreed with the statement, compared to 82 percent of women in comparison households, and the difference was statistically significant. This too indicates a shifting of individual beliefs about the appropriate sharing of household duties as opportunities and responsibilities outside the home expand.

Table 5: Project impact on men's contribution to care work

Proportion of women who reported men's support had increased for:	**GEM (n=296)**	**Comparison (n=442)**	**Project effect (% points)**	**Statistically significant?**
Care of children	54%	32%	**+ 22** (0.04)	Yes
Care of elderly or disabled	48%	28%	**+ 20** (0.04)	Yes
Collecting water and firewood	31%	12%	**+ 19** (0.04)	Yes
Cooking, cleaning or washing	27%	10%	**+ 17** (0.04)	Yes

Notes: Hypothesis tests on group means are statistically significant, with p<0.01 for each indicator.

Figure 7: Project impact on women asking for men's help with care work



Notes: Indicators measure proportion of women in project and comparison groups who reported asking for husbands or sons to support with care work activities in past month. See Annex 4 for details of sample size on each indicator. Hypothesis tests comparing group means varied in their degree of statistical significance with p<0.1 for care of children, p<0.05 for indicators on care of the elderly and cooking, cleaning and washing; and p<0.01 for collecting water or firewood.

Community norms also play a role, with GEM participants more likely to report that a majority of men in their communities believed it acceptable to be involved in care work: 21 percent of women in GEM households thought their communities would accept men cooking, cleaning or washing, compared to 9 percent of women in comparison households. Similarly, 28 percent of women in project households thought their communities would accept men collecting firewood or water, but only 12 percent of women in comparison households felt the same. In both cases the difference between the quasi-experimental groups is statistically significant. There was less of a contrast in perceived acceptability for care of children or the elderly, indicating a distinction between care of family members and household chores. In the case of childcare and care of the elderly, no significant difference was found between project and comparison groups, but both reported higher levels of acceptability for men to be involved (approximately 51 percent in the GEM households and 43 percent in comparison households). All care work is not the same, and GEM communities appear to have made progress in tackling norms around some duties that have traditionally been the preserve of women. The views of husbands of GEM producers were sought in focus group discussions, and these indicate some shifting of attitudes also: 'Ten years ago, women did all the housework, men never did household activities and they thought it will decrease their dignity…now the idea is changing. We help them in rearing livestock and sometimes collect water and fuel.'

'Ten years ago, women did all the housework, men never did household activities and they thought it will decrease their dignity…now the idea is changing. We help them in rearing livestock and sometimes collect water and fuel.'

~ Husband of GEM producer

While some households appear to have changed their approach to the sharing of care work, this is by no means universal. In line with the household survey findings, there are still men who will not contribute to housework; they were described by some (male) focus group respondents as the 'powerful men, busy men'. Focus groups reflected diverse perspectives even within a small group of men whose wives took part in GEM training. Respondents in one community believed men 'sometimes help the women by taking care of their children and livestock rearing to a small extent'. But when asked if most men in the village took care of household duties such as washing and caring, the response was different: '…very few of us. We don't have to do housework. Very rarely we have to help them. They manage all of their work'. Family structure was also important. In a nuclear family, men were described as being more at ease with housework, 'but in a joint family men never like to do housework, they are afraid of other family members... they are scared of society'.

Overall, data from surveys and focus groups concurs in depicting communities where attitudes are shifting within the household, behind closed doors. Actions are starting to follow in the same direction, with women asking for and receiving more help at home. However, there is

much to be done to make such changes acceptable to wider community members. Arguably GEM activities have supported promising steps towards a fairer distribution of care work at home. The data highlight the important social and household effects that can arise from leadership training and the promotion of opportunities for women in dairy enterprise and markets.

Question 9: Has GEM empowered women to engage in decision making in their communities?

Being part of the GEM project had an effect on wider participation in community groups. As well as being part of a producer group, the majority of women benefiting from GEM were also members of a savings group (68 percent) and/or some other community based organization (72 percent). Overall participation in groups outside the household was significantly greater than for women in the comparison group.

Once in these groups, women from GEM households were likely to be moderately involved in group decision making, with 75 percent of women reporting they were involved 'to some extent' in producer groups, 81 percent in savings groups and 79 percent in other community-based organizations. Across the different community groups, approximately 7 percent reported they were not involved at all. This being the first exercise to collect such data, it is not clear whether the share of women reporting lower engagement within the groups is normal, low or higher than other projects and groups (data was not available for the comparison group to draw conclusions on project impact on this indicator).

Table 6: Project impact on women's leadership and community decision making

Proportion of women who 'agree' or 'strongly agree' with the statement:	**GEM (n=296)**	**Comparison (n=442)**	**Project effect % points**	**Statistically significant?**
I feel I have a number of good qualities	91%	66%	**+ 25** (0.04)	Yes
I continue to work on hard tasks even when others oppose me	64%	39%	**+ 25** (0.05)	Yes
Even when my farm is doing well I keep my eyes open in case I find a way to improve it	84%	75%	**+ 9** (0.04)	Yes
I feel comfortable speaking out at a meeting of men and women to help decide on infrastructure (roads, water supplies, well) to be built in my community	66%	39%	**+ 27** (0.27)	Yes

Notes: Hypothesis tests on group means are statistically significant, with p<0.01 for each indicator.

Alongside higher participation in community and producer groups, women from GEM households report significantly higher self-esteem and confidence (see Table 6). These findings relate not only to confidence in their own problem solving skills and resilience, but also to their confidence in speaking out at meetings involving men and women relating to community investments and spending. Community level focus groups provided some examples of how women participated to influence wider issues. In one case, a group of people including women and men approached local leaders to discuss flooding and river erosion linked to the construction of a bridge by the government. Some women felt that they were part of community decisions on difficult and sensitive issues too, such as violence against women, child marriage, dowry and divorce: 'Ten years ago, women were not involved in any community decisions but now they are participating. They have learned many things from [GEM] training for the improvement of our life'; 'we can now talk about our rights, we earn money, our decisions and opinions get value and priority in many cases'.[6]

Such participation is not limited to women who participated in GEM. In another area, women who belonged to a village where GEM operated, but were not producer group members themselves, highlighted their efforts to discuss new road-making plans with men in the village. Elsewhere, women reported grouping together to consult a local leader about building a bridge in their community. These anecdotes further support the idea of positive spillover effects on community norms and the voice of women in community decisions; both through direct effects on the women who participated in GEM activities, as well as indirect effects for other women by creating space and making it acceptable for them to participate in local planning.

Question 10: Can women's economic empowerment (unintentionally) lead to increased rates of gender-based violence?

The evaluation recognized the potential for unintended negative consequences from activities that shift power dynamics within the community and household. Gender-based violence was one such consequence that was important to examine, particularly as the GEM project design actively sought to empower women and encourage them to assert their own preferences both within and outside the household. On the other hand, such training may also have had the opposite effect of mitigating against the risk of violence against women in the household, and it was not clear from the theory of change which of these effects might be predominant.

The survey asked respondents if they were aware of any abusive behaviour or violent acts being committed against someone they knew over the past 12 months. The questions were deliberately indirect to take

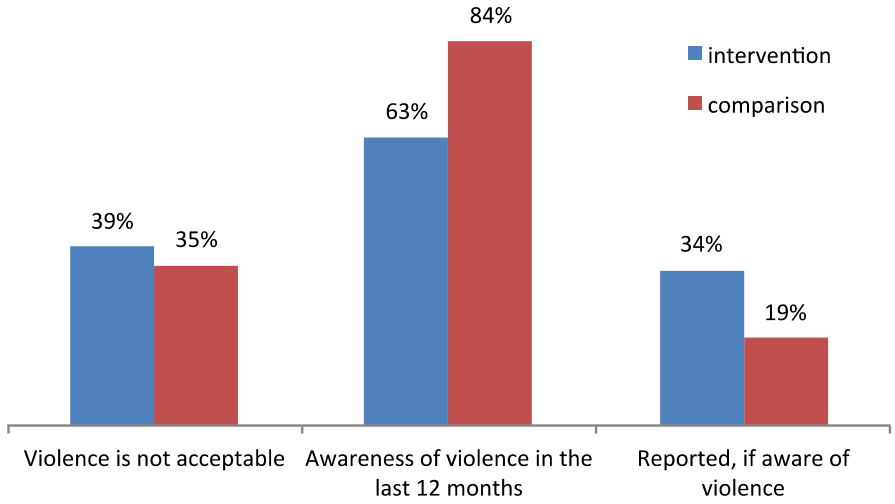
[6] Women from Taluk Shabaj and Paschim Deluabari producer groups, October 2017.

a context-sensitive approach in light of local norms and customs.[7] The results therefore do not distinguish between violence committed in the community and those experienced by the respondent herself, but the expected trade-off was to elicit more responses from participants who might otherwise be reluctant to share their views at all, or who might be afraid to give honest responses.

The data presents a series of important findings (see Figure 8). Firstly, levels of violence are high, with 63 percent of women in the project group reporting awareness of any abuse (physical or psychological), rising to 84 percent in the comparison group. Secondly, the difference between the groups is statistically significant and suggestive that women and communities who have participated in GEM training are less likely to experience violence.

Thirdly, and in line with estimated prevalence of violence, a minority in both groups reported such violence was *not* acceptable: only 39 percent in the project group and 35 percent in the comparison group (a difference that was not statistically significant). Estimates of acceptability and awareness of violence evidently point to the need for continued support and resources to tackle high rates of gender-based violence. The evaluation focused on female respondents, but future research and programme design may want to consider the attitudes and awareness among male members of the community.

Figure 8: Project impacts on gender-based violence



Finally, the evaluation suggests that women in the GEM households were more likely to report violence if they were aware of, with 34 percent indicating they had reported to someone relative to only 19 percent in the comparison group. The confidence to take the issue to an authority figure outside the home or community is a key indicator of empowerment, and is a promising sign of existing norms continuing to be challenged in the future through transparency and women's capacity to act.

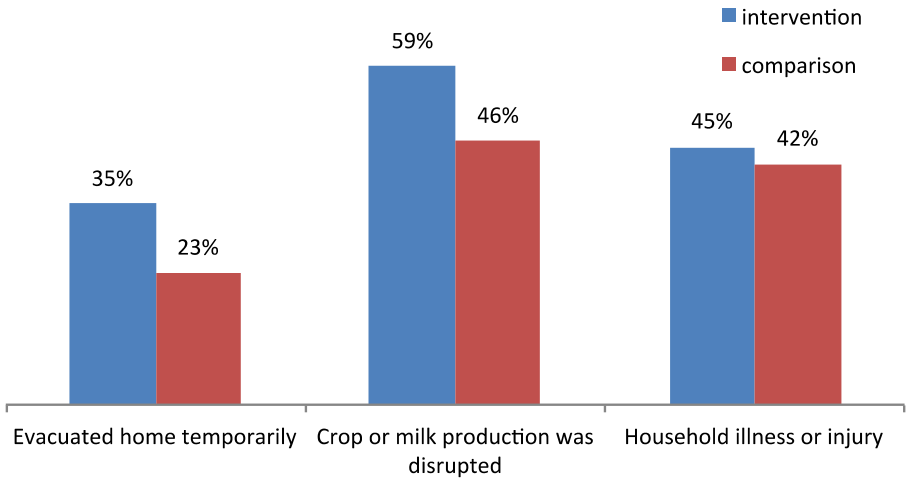
[7] Questions around gender-based violence were only asked if privacy was assured during that part of the interview to preserve confidentiality for all respondents, and at any time participants could ask to move the survey on to the next topic. The sample size for this part of the household survey is lower than the other questions, with 388 responses on questions relating to the acceptability of violence in the community, 381 responses on awareness of violent acts over the past 12 months, and 278 responses on whether violence was reported and to whom. Approximately 50 percent of survey participants engaged with this set of questions overall, with 37 percent providing a response on the final question around reporting violence if they were aware of it.

3.4. RESILIENCE

Question 11: Were households supported by the GEM programme more resilient to natural disasters?

Bangladesh experienced significant flooding during 2017, with some two-thirds of the country under water in August. While some level of flooding is an annual monsoon phenomenon, some areas experienced levels of flood not seen since 1988.[8] In the context of the evaluation this had two implications. Firstly, it played a role in the sampling strategy, with some areas still too flood-affected to visit during September 2017 when fieldwork was underway. However, the evaluation made efforts to ensure flood-affected households were represented in the household survey. Full details on the sampling strategy are set out in the annex. Secondly, the impact of flooding was expected to affect the measured outcomes and impacts during the evaluation, potentially making it more difficult to detect progress if household assets and livelihoods were affected by natural weather events in the months immediately prior to data collection. Nevertheless, the timing of the evaluation also allowed an opportunity to examine how prepared GEM households were to cope with the floods, and to reflect on the disaster risk and resilience of dairy farming in flood-prone areas.

Figure 9: Flood exposure and effects in 2017



Survey data shows that project households were widely affected by heavy rainfall and subsequent flooding. Respondents were asked about the extent to which they were affected, with 44 percent reporting they were 'severely' affected by floods in 2017 and 35 percent reporting they temporarily evacuated their homes. Clear differences emerged between the comparison and intervention groups as reflected in Figure 9. On two important dimensions, GEM households were more affected by flooding

[8] https://www.oxfam.org.uk/what-we-do/emergency-response/south-asia-floods

than the comparison households: the share of households who were temporarily evacuated and the share of households reporting disruption to crop and milk production.[9] These differences would have made it more difficult to find evidence of project impact, since comparison households may have reported better outcomes simply as a result of being less flood-affected during the year of the evaluation. There are also indicators that show little difference between the two groups, for example the share of households reporting illness or injury to a family member due to the floods.

The data also raises some interesting questions around the differences between comparison and intervention households. For example, although comparison group households were less likely to be evacuated, those that were reported being away from their homes for over 16 days, on average, compared to 9 days for project households. The available data from the household survey does not fully explain what might account for these different experiences, and whether it reflects stronger coping strategies among GEM households or some other explanation. Similarly, project households reported raising, on average, 50,000 taka from sales of assets in order to raise cash, compared to 40,000 Taka among comparison households. The difference is not statistically significant, but may reflect some improved resilience or preparedness through savings and an asset buffer.

The effects on households from flooding and the strategies used to deal with them were probed through focus group discussions. Communities confirmed widespread effects: 'all are damaged by floods, and to a great extent [for example] loss of crop grasses for cows, feed cost increases, house and livestock shelter damage'. When delving into how households prepared for the floods, qualitative data highlighted actions such as putting aside inputs for cows, including fodder and dry food, saline and medicine. Respondents from *char* areas mentioned they had held community meetings and encouraged households to save money, anticipating emergency situations. In line with the theory of action, these preparedness strategies might reflect the influence of capacity building interventions provided by GEM, such as model farms and training activities on fodder cultivation and preservation. Many respondents felt they had taken action to prepare for the flooding, although the height of the rising waters meant that damage and suffering was inevitable.

Despite the differences in exposure and effects of flooding, project households reported very similar outcomes to the comparison group. For example, sufficiency of income was broadly the same across both groups, with 46 percent of project households and 42 percent of comparison households having to sell assets to cover household expenses due to the floods. The positive project impacts on income, dairy revenues and investments reported in earlier sections are arguably conservative estimates given the more adverse effects from flooding on project households, making those results all the more remarkable.

[9] This was unavoidable, as sampling rules had for ethical and practical reasons set limits on the extent that the survey teams would approach flood-affected households who were not part of the GEM network.

The evaluation highlights that project households were no more or less prepared for potential flooding than the comparison group in terms of advance warning about weather patterns. This perhaps suggests there was scope to increase outreach and demand for weather information among project households. However, anecdotal feedback during the evaluation design highlighted that some strategies were employed to mitigate the effect of the floods, including moving cows and assets to higher ground as soon as water levels began to rise. This, along with higher vaccination rates, might help explain why 3 percent of the intervention group reported losing cows due to the floods compared to 5 percent in the comparison group. This finding is not statistically significant but may be indicative that wider practices to improve dairy health may also have had spillover benefits during times of crisis, such as the floods, mitigating against catastrophic dairy losses.

4. LEARNING AND RECOMMENDATIONS

The theory of change workshop during September 2017 raised three priority themes for this evaluation: changes to income, improvements in farm practices, and access to markets and services. Underpinning each of these was resilience to natural disasters, a non-ignorable and stark reality of the field context. Key learning points emerged on each of these themes:

- The project had a positive and significant effect on increasing household incomes, measured through food consumption and annual household expenditure. Indicators assessing quality of income showed women had more stable and frequent income from dairy than comparison households, and the timeliness of dairy income was high.
- Clear improvements can be observed in farm practice, which have arguably played an important role in improving milk production and revenues. High rates of vaccination for cows among project households may have played a key role in protecting valuable assets and livelihoods in the face of extreme weather events. Training delivered through the GEM project has been highly effective in these ways.
- Access to market services has been promoted effectively. GEM supported women as dairy farmers by increasing access to agricultural services, such as artificial insemination and vaccinations. GEM also supported women as entrepreneurs, with improved access to credit and participation in savings groups.
- Further, attitudes among women who have participated in GEM activities reflect strong evidence of empowerment, confidence, self-esteem and decision-making power relative to the comparison group. These features are all key requirements for women to take up economic opportunities.
- While there is much to be lauded in these outcomes, improving access to markets deserves further attention and resources when considered through the lens of women's perceived ability to influence, lead and participate equally in business transactions. Notably, the issue of access to markets for selling produce remains a thorny issue, with considerable stigma surrounding the job of taking milk and other produce to markets. To promote meaningful and swift change, investments in assertiveness and negotiation skills and other activities focused on women's attitudes must be complemented by system-wide improvements in transport and connectivity, safety and security, and redesigning physical markets as spaces equally welcoming to female and male entrepreneurs.
- Results show considerable benefits from GEM activities for women's power in household decision-making, particularly around economic issues, such as farming practice and who works outside the home to

earn money. Yet, the evaluation also revealed that key areas of women's lives remain heavily influenced by gendered norms, notably around care work. While there is a belief that helping with unpaid care work is more acceptable to a majority of men in GEM communities, and men in GEM households are increasingly helping with care activities, the heavy and unequal responsibility for unpaid care work remains a barrier to women achieving, earning, and leading outside the home. Future programmes may want to identify more closely where progress has been made, and what the next frontier is in order to support men and women to disrupt the status quo. Lessons from GEM respondents suggests that community norms and the ability to share care work publicly should be one such area of focus.

- Despite being more flood-affected, GEM households were able to demonstrate advantages in their incomes and investments over the past year, and this is implicit evidence of greater resilience in the face of extreme weather shocks and disruption to family life and farming. Arguably, the results are to some extent a validation that dairy farming can be designed and managed to be flood-resilient, and the GEM dairy business model is worth replicating in wider communities.
- However, indicators suggest that GEM households could benefit further from access to information to support preparedness. Future programme design would benefit from research into how borrowing and asset-selling behaviour sits within the household's own perceptions of resilience. What is less known, although the data is suggestive of this, is how far households used dairy farming and income as a resilience strategy.
- The multi-stakeholder forum has yielded some positive outcomes. Looking ahead, the forum's priorities are to select a dynamic leader from the key stakeholders, register the association under social welfare ministry/department, engage more private sector and government actors, develop an action plan with tangible milestones, and create a fund for its future activities. These practical steps will help to turn the forum from a loosely formed structure into an organization that is fully functional, self-sustainable, and properly equipped to tackle the challenges of the dairy sector.
- Access to credit remains an area where more action is needed, and attitudes towards formal banking reflect wariness among smallholder farmers. However, experience with the Rupali Bank pilot scheme shows strong signs of promise. GEM has leveraged smallholder access to formal credit through effective advocacy. Importantly, the next steps are to assess the feasibility of scale-up to meet existing demand, to give further attention to promoting demand by considering how access in *char* areas can be improved, and to address the negative perceptions of farmers towards formal banking.

This publication was written by Dr Manu Savani and Alastair Stewart. The authors acknowledge the assistance of Miranda Morgan, Kanako Yoshikawa and Simone Lombardini in its production.

For further information on the issues raised in this publication please email astewart1@oxfam.org.uk or mmorgan1@oxfam.org.uk.

Published by Oxfam GB.

ISBN: 978-1-78748-536-5

DOI: 10.21201/2019.5365

Oxfam GB, Oxfam House, John Smith Drive, Cowley, Oxford, OX4 2JY, UK.

OXFAM

Oxfam is an international confederation of 20 organizations networked together in more than 90 countries, as part of a global movement for change, to build a future free from the injustice of poverty. Please write to any of the agencies for further information, or visit www.oxfam.org.

www.oxfam.org

ANNEXES FOR GEM BANGLADESH FINAL EVALUATION REPORT

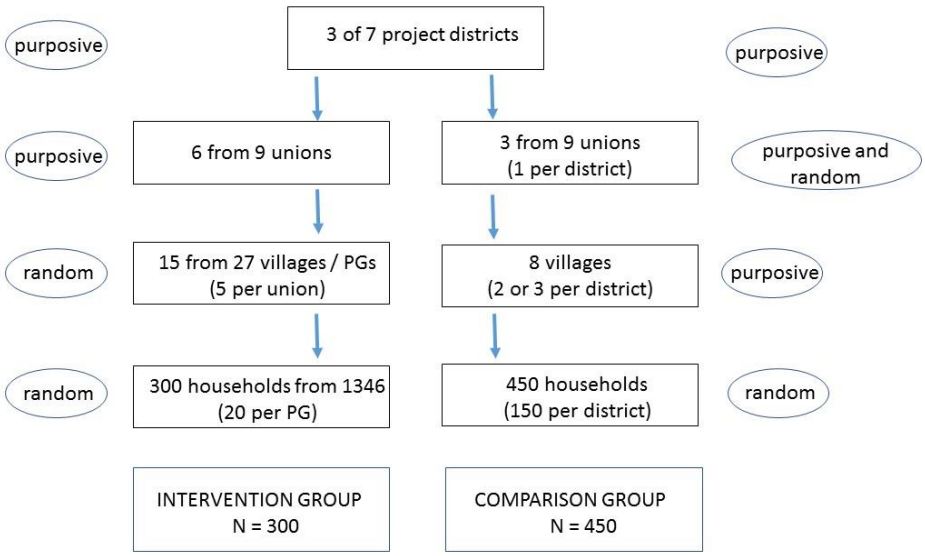
Contents

ANNEX 1: HOUSEHOLD SURVEY AND QUASI-EX-PERIMENTAL ANALYSIS

Sampling

Sampling was undertaken using a two-stage stratified random sampling method. Stratification was undertaken through purposive discussion of various criteria to establish broad fieldwork sites (at sub-district level) in three districts. Households were randomly selected from a clearly established sample frame. The sampling strategy took careful account of local conditions, notably the effects of heavy flooding over previous months, the potential for further weather events as it was the monsoon season still, and the expectation that a team of 10 female enumerators would be in the field and the particular security and logistical needs the fieldwork team would need to be mindful of. An early decision was made to interview female participants only in the GEM evaluation, and in Bangladesh this implied excluding 2% of project participants who were male. Project eligibility criteria was determined by having one or two adult cows in 2014, and this was used as a screening question for the comparison group. The figure below summarises the sampling decisions taken and whether the process was based on random sampling or purposive screening and selection.

Figure A1. Sampling strategy overview



Matching: constructing an appropriate counterfactual

The household survey ultimately involved 747 households, 300 from the intervention group (40%) and 447 from the comparison group (60%), who were approached by the survey team during September 2017. The main analysis for the quasi-experimental component is conducted by comparing outcome variable means between the two households groups using propensity-score matching (PSM).

The propensity score is calculated using the following outcome indicators:

- Household size in 2013
- Education of the household head
- Education respondent
- Marital status respondent
- Crop diversity in 2013

- Number of cows owned in 2013
- Land owned in 2013
- Household income engagement in 2013
- 1[household in the 1st quintile wealth index in 2013[1]]
- 1[household in the 2nd quintile wealth index in 2013]
- 1[household in the 4th quintile wealth index in 2013]
- 1[household in the 5th quintile wealth index in 2013]

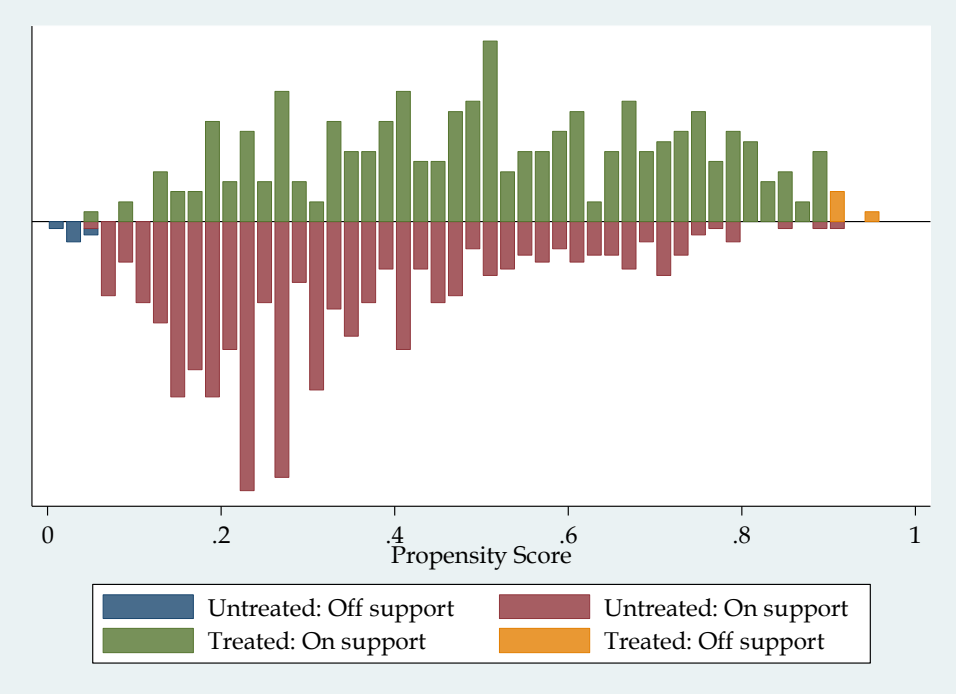
Table A1 below compares intervention and comparison group before matching, providing evidence that on average, *before matching*, households in the intervention group have a larger number of household members, were likely to be poorer, and have greater crop diversification. There were no differences between the two groups on the education level of the household head, the respondent's marital status, the number of cows owned, the area of land owned, or involvement in other sources of income other than crop farming.

Table A1: Descriptive Statistics Before Matching	Intervention mean	Comparison mean	Difference
HH size in 2013	4.247	3.776	0.470***
1[HHH has education greater than primary]	0.180	0.206	-0.026
1[Respondent has education greater than primary]	0.180	0.206	-0.026
1[Respondent is married]	0.933	0.953	-0.020
Crop diversification 2013	3.503	1.987	1.517***
Number of cows owned in 2013	2.857	2.875	-0.018
How much land did you own in 2013?	41.417	37.023	4.394
1[HH involved in sources of income other than crop in 2013]	0.777	0.792	-0.015
1[HH in the 1st wealth quintile]	0.253	0.226	0.027
1[HH in the 2nd wealth quintile]	0.190	0.145	0.045
1[HH in the 4th wealth quintile]	0.190	0.206	-0.016
1[HH in the 5th wealth quintile]	0.120	0.253	-0.133***
Observations (for each variable)	300	447	

In order to control for the observable differences, Propensity Score Matching was implemented. As a result, 9 observations were excluded as being outside the common support area (4 from the intervention group and 5 from the comparison group). Observations are weighted accordingly with their propensity score.

[1] Wealth index calculated using Principal Component Analysis based on household asset ownership and housing characteristics based on information referring to 2013. Questions 201 – 217, and 327 – 328 in Bangladesh survey refer.

Figure A2. Common support area for propensity score matching



Balancing test

Following the matching process, none of the matching variables were significantly different as set out in the table below.

Table A2: Descriptive Statistics After Matching

Variable	Un-matched/ Matched	Treated Mean	Control Mean	%bias	%reduct bias	t-test t	t-test p>t	V_e(T)/ V_e(C)
hhsize2013	U	4.23	3.80	26.1		3.44	0.001	0.73*
	M	4.23	4.11	7.2	72.3	0.88	0.381	0.8
hhh_edu	U	0.18	0.20	-4.2		-0.56	0.574	0.94
	M	0.18	0.18	0.9	78.8	0.11	0.911	1.04
resp_edu	U	0.18	0.20	-4.2		-0.56	0.574	0.94
	M	0.18	0.18	0.9	78.8	0.11	0.911	1.04
maritalstatus_1	U	0.94	0.95	-5.9		-0.79	0.429	1.27*
	M	0.94	0.95	-2.7	53.9	-0.32	0.75	1.1
cropnum2013	U	3.45	2.00	77.7		10.56	0	0.89
	M	3.45	3.42	2	97.4	0.22	0.824	0.91
asset_selected2013_1	U	2.87	2.89	-0.8		-0.11	0.912	0.9
	M	2.87	2.81	2.5	-194.2	0.31	0.753	1.15
hh_land2013	U	39.80	31.13	13.8		1.84	0.067	0.93
	M	39.80	34.67	8.2	40.8	0.98	0.327	0.9
othersinc2013_b	U	0.77	0.80	-5.5		-0.74	0.46	1.06
	M	0.77	0.80	-5.6	-1	-0.68	0.498	1.13
wealth1_2013	U	0.26	0.23	6.6		0.88	0.379	1.09

	M	0.26	0.25	2.6	60.6	0.31	0.756	1.05
wealth2_2013	U	0.19	0.15	11.3		1.52	0.13	1.27*
	M	0.19	0.21	-5.7	49.2	-0.65	0.516	0.92
wealth4_2013	U	0.19	0.21	-3.9		-0.52	0.606	0.95
	M	0.19	0.17	6.6	-69.8	0.84	0.403	1.06
wealth5_2013	U	0.12	0.24	-32.1		-4.17	0	0.65*
	M	0.12	0.11	2.4	92.7	0.34	0.734	1.11

* if 'of concern', i.e. variance ratio in [0.5, 0.8) or (1.25, 2]

** if 'bad', i.e. variance ratio <0.5 or >2

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%concern	%bad
Unmatched	0.143	142.3	0	16	6.2	94.9*	1.25	33	0
Matched	0.004	3.15	0.989	3.9	2.6	14.6	0.91	0	0

* if B>25%, R outside [0.5; 2]

Project exposure

Before considering the project's effect on outcomes, it is important to examine whether the respondents reported having participated in the activities implemented under this project. The data confirmed that respondents had experienced a number of training activities spanning farming practices as well as business skills and to support leadership, negotiation and assertiveness techniques, and awareness of women's rights. As shown by the table below, there were statistically significant differences between the intervention and comparison groups for each type of GEM project activity.

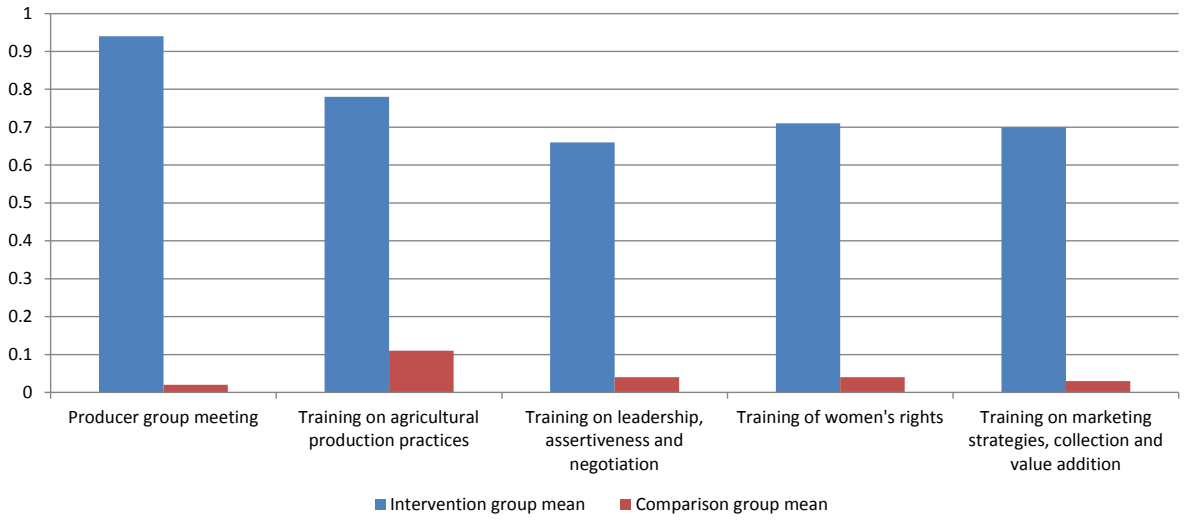
	1	2	3	4	5
Table A3: Project Exposure	Producer group meeting	Training on agricultural production practices	Training on leadership, assertiveness and negotiation	Training of women's rights	Training on marketing strategies, collection and value addition
Intervention group mean	0.94	0.78	0.67	0.71	0.71
Comparison group mean	0.03	0.10	0.03	0.04	0.02
Difference:	0.90***	0.68***	0.64***	0.67***	0.69***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Observations (intervention group)	296	296	296	296	296
Observations (total)	738	738	738	738	738

*Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions.*

These statistical checks confirm that the intervention and comparison groups, while broadly similar in key demographic and geographical aspects, provide a strong contrast in their exposure to the kinds of

interventions GEM sought to provide. This was the basis for the quasi-experimental analysis reported in this evaluation, allowing for the differences between the intervention and comparison groups to be attributed to the GEM project. As with all matching-based statistical analysis, the findings rely on their being minimal unobservable differences between the intervention and comparison group. The evaluation design and household survey sampling process aimed to minimise any risk of bias arising from unobservable differences, and present in this report the most robust available estimates of project impact outside of randomised control trial methodology.

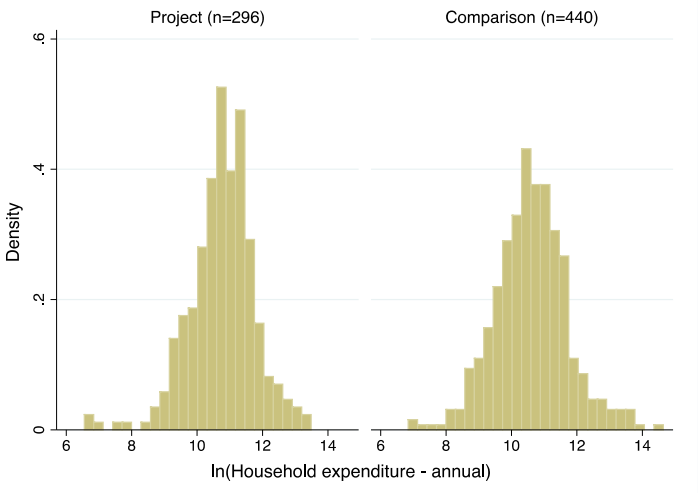
Figure A3. Exposure to GEM-type activities



Household expenditure per annum

Annual spending was estimated by summing a number of activities that were expected to take place less frequently than those included in monthly expenditures. Expenses in the past 12 months included community events and ceremonies, health costs, rent or purchase of land for farming, building materials and repair for the home, clothes and shoe purchases, and purchase of livestock. Clear differences emerged in the average level of household spending between project and comparison groups, as set out in the main report. The distribution of spending took a non-normal form, and as per standard practice the variable was transformed by natural logarithm to present a normal distribution (figure A4). Project impact was found to be statistically significant at the 5% level using log household spending.

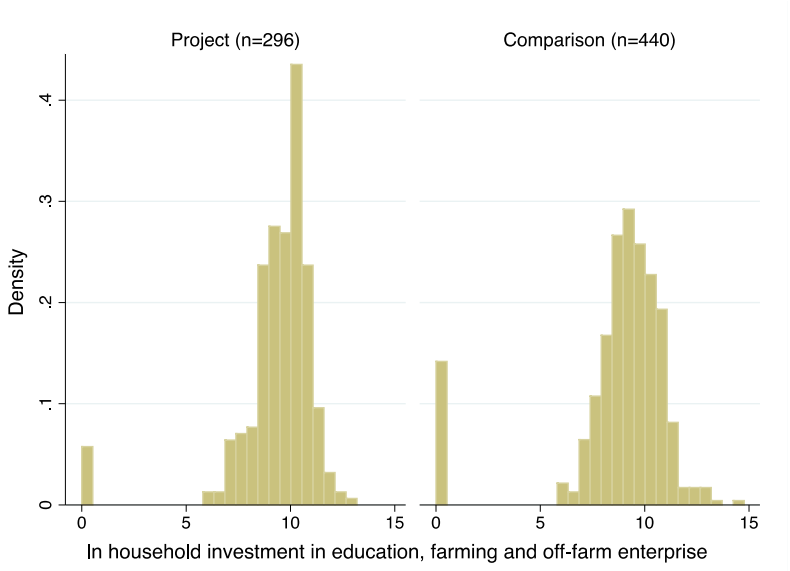
Figure A4. Logarithmic form of annual household spending



Household investment

This variable was defined as annual household spending on three types of activity: education, farming equipment, and off-farm enterprise. Investment was measured in Taka, and statistical analysis transformed the data into logarithmic form to better understand and compare trends. In both quasi-experimental groups, a minority of households reported zero investment (3% in the project group, and 8% in the comparison group). Project impact was found to be statistically significant at the 1% level using log household investment.

Figure A5. Logarithmic form of annual household investment



Female-headed households

With approximately 10% of the sample reporting as female-headed, and after applying the matching methods, the sample size for analysis falls to 77 observations. Statistical power to detect significant project effects is reduced considerably. The results are therefore indicative at best. The data suggests female-headed households had higher annual dairy income of over 15,000 Taka, and greater annual household spending (both statistically significant). Effects are weaker but positive for monthly household spending. No significant effects are detected for food consumption or household investment.

Table A4: Female headed households	Food consumed over past 7 days (per capita per day, calories)	Log household investment	Annual dairy income (Taka)	Log monthly household expenditure	Log annual household expenditure
Intervention group mean (n=38)	94.89	9.15	24109	7.26	10.58
Comparison group mean (n=39)	88.69	8.26	9013	6.88	10.07
Difference:	6.20	0.884	15095**	0.378*	0.508**
(std error)	(10.13)	(0.67)	(7123)	(0.22)	(0.25)
Observations (total)	77	77	77	77	77

ANNEX 2: FOCUS GROUPS

The evaluation incorporated valuable community-level feedback through a series of focus groups across three GEM districts. A total of four communities were targeted, and in each location four focus groups were conducted to ensure the voices of distinct stakeholders were heard: (i) women belonging to GEM dairy Producer Groups (PGs), (ii) village leaders, (iii) husbands of women belonging to GEM PGs, and (iv) women who did not belong to GEM PGs. Approximately 6 to 8 individuals took part in each focus group discussion, from 20 to 85 years of age.

Since men were not included in the household survey, focus groups with husbands was an important means of discovering attitudes and perceptions towards women's empowerment and leadership amongst the male members of their households. The focus groups with husbands, village leaders and women who did not belong to the PGs were also designed to investigate evidence of spillover effects from GEM activities and training amongst the wider community. In each focus group, a trained facilitator led participants through a series of questions and research tools including vignettes, while a second researcher took detailed field notes that were later transcribed for analysis.

The list of locations and dates of focus group meetings is set out in Table B1. Sampling was undertaken at the start of the evaluation in consultation with local partners. Taking the household survey locations as a starting point, the evaluation design team selected communities where community level qualitative data gathering would take place. Site selection for focus group discussions was based on criteria including physical access, extent of damage and recovery from the 2017 floods, and a desire to balance data gathering across river island and mainland communities. Due to timing and resource constraints focus groups only took place in GEM communities and not comparison communities.

Table A5: Schedule of Four Community Focus Group Discussions (October 2017)

Date	Village and District	Brief description of village
2 – 3 Oct	Panchasona, Sirajgonj	Mainland village, no boat travel required, was within 5 mins walk of a road. Some village services identified including a madrasah, small enterprise and petty trading, but overall a quiet scene. Only NGO active was MMS as part of GEM. The significant event of the past 12 months was flooding, which destroyed the rice fields.
4 – 5 Oct	Bara Ghorjan, Sirajgonj	*Char* village, required travel by boat and is not within 5 mins walk of a road. More dispersed settlement, but nonetheless was busy with enumerators identifying children and women, a doctor and school teacher. It had a modest market scene of a small grocery shop and public services appeared less well established. Significant events of the past 12 months did not mention flooding.
6 – 7 Oct	Paschim Deluabari, Gaibandha	*Char* village, required travel by boat and is not within 5 mins walk of a road. Small enterprises and petty trading was visible to fieldwork enumerators, and a primary school was located there. Only NGO active in the village was part of GEM/REECALL. Enumerators noted it was a quiet village. There had been positive changes noted in the past 12 months linked to training, livestock practice, and GEM activities. But floods were the most significant recent event, submerging cropping fields and causing damage to houses and part of the road network.
8 – 9 Oct	Taluk Shabaj, Rangpur	Mainland village, no boat travel required, was within 5 mins walk of a road, with proximity to the river. Services in the village included a high school, market and a dispensary. Only SEED was active in the village as part of GEM/REECALL project. The significant event of the past 12 months was flooding, with crops unable to be grown due to the fields being underwater.

ANNEX 3: SYSTEMIC CHANGE INTERVIEWS

The Systemic Change Interview approach (SCIA) draws on Outcome Harvesting, Most Significant Change and Collaborative Outcome Reporting approaches. The overall aim is to use project stakeholders to identify changes and outcomes; identify gaps in information needed to explain the changes (specifically in relation to a market systems project); and build up a list of questions for specific stakeholders based on these information gaps. Key informant interviews can then be conducted with these stakeholders to verify and elaborate upon these changes, after which the change descriptions are revised and updated.

Step 1: Identify the levels and domains of the Theory of Change to be investigated

GEM being a markets project, has a substantial focus on systemic change and has a strategy to work with government departments, private sector, media and civil society as well as smallholder producers to facilitate change in the dairy value chains. When such change happens it may not be limited to a specific intervention group or direct beneficiaries – the aim is to facilitate positive changes that benefit agricultural producers in targeted value chains, whether or not they are direct project participants. These changes are therefore difficult to capture from a household survey that compare intervention and comparison groups; or from qualitative techniques at the community level if they are only being conducted in project areas.

Techniques such as Outcome Harvesting, Most Significant Change and Collaborative Outcome Reporting are useful approaches that enable the collection of data for complex projects such as market systems programmes. If used well they can also be a means of capturing unintended and negative outcomes of a project intervention. For this reason the GEM evaluation is using SCIA to understand the outcomes of the project in relation to market systems change. During the Theory of Change workshop held in Gaibandha between 17th and 19th September a number of research questions were identified by project stakeholders, which will also be collected through SCIA.

Step 2: Design questions to enable collection of change statements from project stakeholders and documentation

Step 2a: Questions to ask stakeholders to prompt thinking and discussions

At the Theory of Change workshop we asked participants to record and share their experience of what had changed in the last three years, specifically changes regarding their relations with other stakeholders. Participants were asked to give examples of one positive and one negative change. To help participants think about these changes they were given a guiding question: In the last three years, what positive and negative changes have you experienced in working with the private sector and local agribusinesses, government departments, civil society or smallholder farmers?

Step 2b: Get participants to complete templates for each change identified

For each of the changes they identified, participants were asked to complete a template to capture details on what the change was, when and where it took place, who was involved in the change and what were the consequences of the change. Some of the answers were vague or focussed mostly on personal change in income and farming practices, however a total of 23 change examples were collected.

Step 2c: Outcome trawl of project documentation

A 3rd year MEAL report for GEM in Bangladesh was conducted in 2016/17. Outcomes identified in this report were also collected and added to the outcome descriptions collected in the Theory of Change workshop. The same information for step 2b was documented for these outcomes.

Step 3: Grouping, selecting and categorising change statements

Step 3a: Grouping change statements

At this point, a number of the changes identified were overlapping so it was necessary to group and categorise these outcome descriptions and collate the information into one outcome description, this led to the formulation of 19 outcome descriptions:

Small-scale dairy producers are spending more on inputs leading to lower than anticipated profit margins.	PG member households spending income on child education and health
Insufficient availability of livestock services at community level, i.e. AI, input and veterinary services, vaccines.	Households experiencing increased conflict between male and female members
Increased number of stakeholders providing input supply services – including feed and medicine/vaccinations and AI.	Increased workload for women entrepreneurs at household level
Cost of fodder has increased substantially	Increased decision making power for women at household level
Increase in percentage of owned cows receiving full vaccination courses and treatment	Increased social acceptance among men that women can earn and contribute towards family income
Increase in the percentage of heifer cows becoming pregnant through AI for improved breeds	PG members are collectively demanding services from government departments (District Livestock Offices)
Milk quality and production has increased	Bangladesh Government prepared draft of National Dairy Development Policy (NDDP) in 2016.
Household income has increased	Bangladesh Bank financed BDT 200 Crore (2 billion) fund for dairy farmers at 5% interest rate.
PG member households able to purchase more land	Loan scheme of Bangladesh Bank not implemented well

Step 3b: Selecting change statements

The next step was to determine which of these 18 outcomes we wanted to investigate further. For this the main criteria was that it related to (1) the research questions identified in the Theory of Change workshop; (2) the market services stream of the Theory of Change; or (3) was a systemic change that wouldn't be captured from the household survey or community focus group discussions.

Change statement	ToC Research Question	Market services stream	Not in survey or FGDs	Selected for SCIA
Small-scale dairy producers are spending more on inputs leading to lower than anticipated profit margins.		X		Yes
Insufficient availability of livestock services at community level i.e. AI, input and veterinary services, vaccines.	X	X		Yes
Increased number of stakeholders providing input supply services – including feed and medicine/vaccinations and AI.		X		Yes

Cost of fodder has increased substantially		X		Yes
Increase in percentage of owned cows receiving full vaccination courses and treatment	X	X		Yes
Increase in the percentage of heifer cows becoming pregnant through AI for improved breeds	X	X		Yes
Milk quality and production has increased	X			No
Household income has increased	X			No
PG members' households able to purchase more land			X	No
PG member households spending income on child education and health	X			No
Households experiencing increased conflict between male and female members				No
Increased workload for women entrepreneurs at household level				No
Increased decision making power for women at household level				No
Increased social acceptance among men that women can earn and contribute towards family income	X			No
PG members are collectively demanding services from government departments (District Livestock Offices)		X	X	Yes
Bangladesh Government prepared draft of National Dairy Development Policy (NDDP) in 2016.			X	Yes
Bangladesh Bank financed BDT 200 Crore (2 billion) fund for dairy farmers at 5% interest rate.		X	X	Yes
Loan scheme of Bangladesh Bank not implemented well		X	X	Yes

It was decided that of the 18 change statements, based on the criteria 10 would be taken forward for the SCIA component of the GEM evaluation in Bangladesh.

Step 3c: Categorising change statements

Once the changes to be investigated are selected we can categorise the changes into four groups: positive intended; positive unintended; negative intended and negative unintended. You can see in the table below that the majority of outcomes are positive intended and that there is a split of six positive and four negative outcomes. This categorising process helps to check if we have collected a range of outcomes and that we are not only checking on positive outcomes. Ideally in this process we would have identified more unintended (both positive and negative) outcomes. The collection of change statements should be amended to better facilitate the collection of unintended changes. More time dedicated to harvesting outcomes in the Theory of Change workshop may have helped and also the framing of questions on what has changed in a way that is more likely to create a thought process about unintended changes.

Figure 1: Outcomes categorised into positive, negative, intended and unintended

Unintended Outcomes	Positive Outcomes		Intended outcomes
		• Increased number of stakeholders providing input supply services – including feed and medicine/vaccinations and AI. • Increase in percentage of owned cows receiving full vaccination courses and treatment • Increase in the percentage of heifer cows becoming pregnant through AI for improved breeds • PG members are collectively demanding services from government departments (District Livestock Offices) • Bangladesh Government prepared draft of National Dairy Development Policy (NDDP) in 2016. • Bangladesh Bank financed BDT 200 Crore (2 billion) fund for dairy farmers at 5% interest rate.	
	• Small-scale dairy producers are spending more on inputs leading to lower than anticipated profit margins. • Insufficient availability of livestock services at community level, i.e. AI, input and veterinary services, vaccines. • Cost of fodder has increased substantially • Loan scheme of Bangladesh Bank not implemented well		
	Negative outcomes		

Step 4: Drafting initial details of the change statement description, identifying information gaps and associated informants

Step 4a: Detailing the outcome description

This step involves adding more detail to the existing outcome description. Currently the description will be limited to the information collected in step 2; what the change was, when and where it took place, who was involved in the change and what were the consequences of the change. Based on a review of a paper written by ITAD[2], a number of failings in market programme evaluations were identified. These generally referred to concepts that were missing or lacking from M4P evaluations including scale, sustainability and systemic change. These areas have therefore been incorporated into a template for the outcome description so as to ensure capture of relevant information in these areas. Below is an outline of the sub-headings in the outcome description template. Content is added under each sub-heading by the evaluator based on what they already know about the project, from what they have heard or read in the workshops and documentation.

Figure 2: Sub-headings from outcome description template

Description: (What was the change; when and where did it happen; who was involved in the change; and what was the consequence of the change?)
Outcome identified from: (Name of stakeholder, report, reference)

[2] Ruffer, T. & Wach, E. (2013). *Review of M4P Evaluation Methods and Approaches.* ITAD. Commissioned by UK DFID.

Outcome verified by: (Name of stakeholders interviewed for verification)
Contribution: (What did GEM do that contributed towards this outcome? What other actors were significant contributors?)
Evidence of systemic change: (Crowding in, copying, replication, sector growth, backward and forward linkages)
Scale: (What is the estimated reach of this outcome? How many smallholder farmers/indirect beneficiaries is it likely to impact? How was this identified?)
Sustainability: (Static or dynamic? Commercial viability, investment by private or public sector, innovation, organisational capacity).
Impact on women: (How have women been specifically involved/impacted by this change?)
Negative consequences: (What are the actual or potential negative impacts of this change? Do these differ for men and women?)

Step 4b: Identify information gaps and associated informants

For ease of explanation, steps 4a and 4b are separated here but in practice these steps were combined. While detailing the outcome description (step 4a) a number of gaps in the information become apparent and it is in this step that we record all these gaps and questions we have about the outcome and determine where we can get this information from. Some of the missing information may be answered from the community level qualitative research and some from the household survey but we will also need to interview other project stakeholders to better understand these outcomes and to understand their perspective.

Step 5: Drafting Key Informant Interview questions

Once the information gaps have been identified in step 4 it is then possible to group these gaps/questions according to the associated informant that is required to be interviewed. The information gaps will need to be structured into an interview checklist

Step 6: Conducting and documenting key informant interviews

Key informant interviews were conducted by the qualitative research consultant in-person and via telephone. Interviews took place in September and October 2017 with the following key informants:

No.	Type of Stakeholder
1	SKS - GEM Partner
2	MMS - GEM Partner
3	SEED - GEM Partner
4	Artificial Insemination (AI) provider (Rangpur)
5	Feed Supplier -ACI (Rangpur)
6	District Livestock Officer (Rangpur/Gaibandha)
7	PRAN Dairy (Rangpur/Gaibandha)
8	Rupali Bank (Rangpur)
9	Department of Livestock Services - Policy Contact (Dhaka)
10	Secretary, Multi-stakeholder Forums
11	OXFAM Bangladesh

ANNEX 4: ADDITIONAL TABLES OF RESULTS

Table A6a: Dairy income indicators

	5	6	7	8	9	10	11
	Produced milk in the last week	Total quantity of milk PRODUCED in the last week (litres)	Total quantity of milk SOLD in the last week (litres)	Total value of sales of milk SOLD in the last week (Taka)	Average revenues per milking cow (Taka)	Dairy income – annual (Taka)	Buyer of milk was milk collection centre for producer group
Intervention group mean	0.72	19.46	16.30	645.13	469.00	33546.83	0.20
Comparison group mean	0.65	7.25	3.73	176.73	138.22	9190.13	0.00
Difference:	0.07	12.21***	12.49***	467.30***	327.99***	24299.67***	0.20***
	(0.09)	(2.34)	(2.35)	(106.95)	(72.74)	(5736.43)	(0.02)
Observations (intervention group)	141	102	99	99	99	99	285
Observations (total)	263	184	180	180	180	180	691

Table A6b: Household income indicators: food consumption, income from farming and change in household wealth

	1	2	3	4
	Value food consumed in last 7 days, daily per person equivalent	Log(Value food consumed in last 7 days, daily per person equivalent)	Change in household wealth	ln(Household income from crop farming)
Intervention group mean	102.09	4.54	0.01	9.14
Comparison group mean	91.40	4.44	0.05	8.64
Difference:	10.74***	0.10***	-0.04	0.52**
	(3.53)	(0.04)	(0.04)	(0.23)
Observations (intervention group)	296	296	296	232
Observations (total)	738	738	738	429

Table A6c: Household income indicators: income from non-farming sources and household expenditure

	5	6	7	8	9
	ln(Income from sources other than crop)	Household expenditure - annual	ln(Household expenditure - annual)	Household expenditure - monthly	ln(Household expenditure - monthly)
Intervention group mean	6.70	73574.27	10.76	2858.19	7.35
Comparison group mean	6.84	65986.62	10.54	2779.19	7.42
Difference:	-0.15	7506.34	0.21**	57.96	-0.07
	(0.16)	(6648.20)	(0.09)	(695.83)	(0.08)
Observations (intervention group)	244	296	296	296	296
Observations (total)	608	738	738	738	738

Table A7: Quality of income indicators

	1	2	3	4	5
	Stability of income (Months in a year you have at least one cow giving milk)	Frequency of income (Receive income daily or weekly)	Predictability of income (Received amount of income as expected)	Timeliness of income (Received income when expected)	Sufficiency of income (Selling assets to cover household expenses)
Intervention group mean	4.99	0.07	0.43	0.84	0.46
Comparison group mean	4.03	0.01	0.43	0.88	0.42
Difference:	1.02**	0.06***	-0.00	-0.04	0.04
	(0.45)	(0.02)	(0.06)	(0.04)	(0.04)
Observations (intervention group)	231	229	232	232	296
Observations (total)	428	415	429	429	738

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions.

Table A8: Dairy farming practices

	1	2	3	4
	Ever used a separate feed and water manger in the last 12 months	Ever given your livestock concentrated feed in the last 12 months	Ever used artificial insemination (AI) services in the 3 years	Vaccinated your cow in the last 12 months
Intervention group mean	0.78	0.94	0.77	0.93
Comparison group mean	0.41	0.60	0.15	0.46
Difference:	0.37***	0.34***	0.61***	0.46***
	(0.04)	(0.04)	(0.04)	(0.04)
Observations (intervention group)	296	296	296	296
Observations (total)	738	738	738	738

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions.

Table A9: Participation in household decision making indicators

	Farming practice	How to spend money made from farming	Education of children	Going to see the doctor	Who takes care of family members	Personal travel to relatives	Personal participation in community group	Family planning	Who works to earn money
Intervention group mean	0.75	0.69	0.86	0.91	0.97	0.90	0.93	0.82	0.65
Comparison group mean	0.59	0.56	0.78	0.80	0.97	0.75	0.55	0.77	0.38
Difference:	0.16***	0.13***	0.08**	0.10***	0.00	0.15***	0.38***	0.05	0.27***
	(0.04)	(0.05)	(0.04)	(0.03)	(0.01)	(0.04)	(0.04)	(0.04)	(0.04)
Observations (intervention group)	296	296	296	296	296	296	296	296	296
Observations (total)	738	738	738	738	738	738	738	738	738

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions.

Table A10a: Household unpaid care work indicators

	1	2	3	4
	Men support increased: care of elderly or disabled	Men support increased: care of children	Men support increased: cooking, cleaning or washing	Men support increased: collecting water or firewood
Intervention group mean	0.48	0.54	0.27	0.31
Comparison group mean	0.28	0.32	0.10	0.12
Difference:	0.20***	0.23***	0.17***	0.19***
	(0.04)	(0.04)	(0.04)	(0.04)
Observations (intervention group)	296	296	296	296
Observations (total)	738	738	738	738

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions.

Table A10b: Household unpaid care work indicators

	1	2	3	4
	Ask help: care of elderly or disabled	Ask help: care of children	Ask help: cooking, cleaning or washing	Ask help: collecting water or firewood
Intervention group mean	0.46	0.54	0.33	0.47
Comparison group mean	0.38	0.47	0.23	0.28
Difference:	0.09**	0.08*	0.11**	0.19***
	(0.04)	(0.05)	(0.04)	(0.05)
Observations (intervention group)	290	291	290	291
Observations (total)	725	724	722	723

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions.

Table A10c: Household unpaid care work indicators

	1	2	3	4	5
	Acceptability of men doing care work: care of elderly or disabled	Acceptability of men doing care work: care of children	Acceptability of men doing care work: cooking, cleaning or washing	Acceptability of men doing care work: collecting water or firewood	No satisfied with the division of labour
Intervention group mean	0.50	0.52	0.21	0.28	0.02
Comparison group mean	0.43	0.44	0.09	0.12	0.04
Difference:	0.07	0.08*	0.12***	0.15***	-0.02*
	(0.05)	(0.05)	(0.03)	(0.04)	(0.01)
Observations (intervention group)	296	296	296	296	296
Observations (total)	738	738	738	738	738

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions.

Table A11a: Women's empowerment and leadership indicators

	1	2	3	4
	Women's economic empowerment (*If a woman earns more money than her husband, it's almost certain to cause problems*)	Acceptability of market role (*On the whole, men make better entrepreneurs than women*)	Care responsibilities (*If the wife is working outside the home, the husband should help her with household chores*)	Collective action (*It is better for women to work together to solve problems than working alone*)
Intervention group mean	0.62	0.33	0.95	0.94
Comparison group mean	0.67	0.47	0.82	0.89
Difference:	-0.06	-0.14***	0.13***	0.05*
	(0.04)	(0.04)	(0.04)	(0.03)
Observations (intervention group)	296	296	296	296
Observations (total)	738	738	738	738

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. Empowerment measured positively when respondent says 'disagree' or 'strongly disagree' to first two statements, and 'agree' or 'strongly agree' to last two statements.

Table A11b: Women's empowerment and leadership indicators

	1	2	3	4
	Perception of self-esteem (*I feel I have a number of good qualities*)	Confidence (*I continue to work on hard tasks even when others oppose me*)	Business skills/confidence (*Even when my farm is doing well I keep my eyes open in case I find a way to improve it*)	Leadership skills (*I feel comfortable speaking out at a meeting of men and women to help decide on infrastructure (roads, water supplies, well) to be built in my community*)
Intervention group mean	0.91	0.64	0.84	0.66
Comparison group mean	0.66	0.39	0.75	0.39
Difference:	0.24***	0.24***	0.09**	0.27***
	(0.04)	(0.05)	(0.04)	(0.04)
Observations (intervention group)	296	296	296	296
Observations (total)	738	738	738	738

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions. Means reported on proportion of women who 'agree' or 'strongly agree' with statement.

Table A12: Gender based violence indicators

	1	2	3
	NOT accept any violence	Received any violence in the last 12 months	Reported, in case of violence
Intervention group mean	0.39	0.63	0.34
Comparison group mean	0.35	0.84	0.19
Difference:	0.06	-0.21***	0.14**
	(0.06)	(0.05)	(0.06)
Observations (intervention group)	160	158	101
Observations (total)	388	381	278

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01;
PSM estimates are bootstrapped with 1000 repetitions.

Table A13: Exposure to 2017 floods indicators

	1	2	3	4	5
	Affected by floods to some extent or severely in 2017	Affected by floods SEVERELY in 2017	Sufficiency of income: selling assets to cover household expenses	How much cash did you raise from all sales? (Taka)	Preparedness: No seek weather info to help prepare for potential floods
Intervention group mean	0.92	0.44	0.46	49983.48	0.09
Comparison group mean	0.72	0.24	0.42	39872.43	0.10
Difference:	0.19***	0.20***	0.04	9183.51	-0.01
	(0.04)	(0.04)	(0.05)	(5684.17)	(0.02)
Observations (intervention group)	296	296	296	135	296
Observations (total)	738	738	738	278	738

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions.

Table A14: Consequences of 2017 floods indicators

	Evacuated home temporarily	Number of days evacuated	Anyone in the household suffer illness or injury due to the flood water	Crop or milk production was disrupted	Lost cows due to the flood	How much cash compensation did you receive? (Taka)	Value of any compensation you received in kind (food, clothes) (Taka)
Intervention group mean	0.35	8.89	0.45	0.59	0.03	116.97	197.44
Comparison group mean	0.23	16.84	0.42	0.46	0.05	17.71	185.20
Difference:	0.12***	-5.80**	0.04	0.14***	-0.03	102.21**	7.95
	(0.05)	(2.45)	(0.05)	(0.05)	(0.02)	(48.90)	(43.74)
Observations (intervention group)	271	95	271	271	271	271	271
Observations (total)	597	171	597	597	597	597	597

Standard errors clustered at village level in parentheses; * p < 0.1, ** p < 0.05, *** p < 0.01; PSM estimates are bootstrapped with 1000 repetitions.

ANNEX 5: LIMITATIONS OF FIELDWORK AND ANALYSIS

All fieldwork faces limitations, and this evaluation exercise was no exception. The timing of data collection coincided with the monsoon season and, following heavy rainfall and flooding earlier in the year, this made for logistical challenges in travel and reaching pre-determined survey locations. As far as possible, the survey team relied on local partners' knowledge of the area and transportation links to reach project participants even in some difficult to access villages. This enabled reasonable coverage of diverse geographical areas, notably the river islands (*char* villages). A pre-agreed sampling technique for alternative locations in comparison unions (sub-district level) also ensured minimal bias where sampled villages and households either could not be reached or contacted. The use of tablets minimised risks to fieldwork operations, ensuring secure collection and transmission of data and promoting extensive travel to enhance coverage of the three districts chosen for the evaluation exercises.

This evaluation exercise deliberately relied on a combination of quantitative and qualitative methods and data, in order to enhance both our understanding of the results and context, and the validity of findings. The analysis has relied on quasi-experimental methods in order to allow for causal inference and determine the impact of the project on those it reached. The statistical rigour of these methods depends to a great deal on the quality of data and the robustness of the matching approach used. Household data was subject to a series of quality assurance checks both during fieldwork and once the data was collated. The matching approach also underwent a number of statistical procedures to select the right matching algorithm, and replication materials can be made available on request. Ultimately, the statistical analysis relies on the comparison group being an appropriate counterfactual for the project intervention, and the matching process ensures as far as possible this is the case based on observable characteristics. There remains the possibility of unobserved variables introducing bias and reducing the validity of results. However, the triangulation of statistical analysis provides an important source of triangulation to sense-check and corroborate findings. Further investigation of the 'risk of bias' is set out in table A15.

The qualitative analysis does not attempt to allow for causal inference, but does provide valuable context, triangulation, and nuance to the statistical analysis. Further, qualitative data has shed light on important aspects of programme implementation and the policy environment in which Oxfam and local partners operate. This data is essential to understanding how and why the project leveraged impact (or failed to). Focus group data was designed to ensure the lived experiences of project participants was adequately captured in the evaluation, both as a unique source of knowledge and reflection on how the programme delivered improvements (or did not) and in order to reflect Oxfam's values through the co-production of insight and learning for future programming. The quality of focus group and interview data necessarily relies on highly skilled facilitation and the use of research tools to engage participants in a meaningful and authentic way; and the application of vignettes for community-level discussion was one attempt to do so.

Lessons learned from this evaluation exercise include the following. The combination of household survey with in-depth community group discussions and stakeholder interviews, building on a participatory theory of change workshop to elicit local priorities for evaluation questions, was a viable evaluation model. It does, however, require resources and time in order to deliver high quality outputs, and this may not be feasible or cost effective for all programmes.

Not all quasi-experimental impact evaluations are the same. Choices made during sampling, selection of the comparison group, and at the analysis stage are crucial in assessing the overall level of confidence in the results. This document provides a framework to assess the risk of bias against ten predetermined parameters, specifically for ex-post quasi-experimental impact evaluations. Lower overall risk, provides higher confidence in the results.

Table A15: Risk of Bias Assessment

	Title	Description	Assessment	Description
Sampling				
1	Random sampling	Score LOW risk if: • Sampling is conducted using probability random sampling methods on a clearly established sample frame. Score MEDIUM risk if: • Sampling is conducted using probability random sampling methods at geographical level (e.g. village level), and use random sampling to select respondents within the geographical area. Score HIGH otherwise.	Low	Sampling was conducted using probability random sampling methods on a clearly established sample frame for both intervention and comparison households. The sample frame for intervention households came from the project lists for the three districts. The sample frame for comparison group households came from union-level administrative lists sourced from local government, for the unions selected as appropriate counterfactual regions. Household level eligibility was determined by criteria replicating original project selection criteria (including owning a cow in 2014).
2	Representativeness of project participants	Score LOW risk if: • Project participants have been involved for the entire duration of the project and have been involved in the project with the same level of exposure. • Project participants have been exposed to a variety of different activities, some may have dropped out from some activities, but sampling is conducted on the entire list of project participants. Score MEDIUM risk if: • Project participants have been exposed to a variety of different activities. Sampling is conducted only among those project participants that have been enrolled for the entire duration of the project or that have been enrolled in all the activities. These are not less than 80% of the entire list of project participants OR it is clear the results apply only to a particular group of project participants. Score HIGH otherwise.	Low	Broadly speaking there has been good continuity with project participation. Some participants may have joined after the project start date but they appeared small in number when reviewing project listings. Sampling was conducted on the entire list of active participants in 2017.

	Title	Description	Assessment	Description
Selecting comparison group				
3	Potential for contamination (spillovers)	Score LOW risk if: • The units for comparison group are selected in geographical areas where it is not reasonable to expect for the project to have had spillover effects. • The project also implemented some activities (which are not considered the most relevant under analysis) which are expected to have had an impact also in the comparison group. (e.g. the project implemented campaigns using radio and other digital media, but these are only a minor component of the activities implemented). The report makes clear which impact is assessed (added-value of other components, taking into account exposure to those minor components) Score HIGH risk if: • Units for the comparison group are selected within the same geographical area as the intervention group, and it is reasonable to expect that project activities had spillover effects. (e.g. comparison observations within the same village, for awareness raising projects)	Low	Units for comparison group were selected in geographical areas where it is not reasonable to expect for the project to have had spillover effects. Radio-based interventions were a minor component of the activities implemented.
4	Self-selection of project participants	Score LOW risk if: • The comparison group is exploiting an experiment or natural experiment. • Units are randomly selected at community level both in the intervention and comparison group. • The selection process for the comparison group is mimicking the same selection process used by the project. Score MEDIUM risk if • If the self-selection is corrected during the matching procedure (e.g. controlling for group participation at baseline) Score HIGH risk if: • Project participants were selected or self-selected based on idiosyncratic or unobservable characteristics, and the selection of comparison respondents is done randomly from neighbouring geographical sites.	Medium	Small holder farmers directly involved in the project are farmers that decided to enrol in cooperatives or producer groups. It might be possible that there are unobservable characteristics that are making these farmers different (e.g. aptitude). The matching procedure will attempt to correct for this by controlling for group participation at baseline.

	Title	Description	Assessment	Description
5	Other interventions in the comparison group	Score LOW risk if: • There are no other actors in the area (e.g. INGOs, NGOs, governmental programmes) • Other actors are conducting activities which are not linked to the project's theory of change Score MEDIUM risk if: • Other actors are conducting similar activities in both the intervention and the comparison group Score HIGH risk if: • Other actors are conducting similar activities, partially related in the comparison communities • Other actors are conducting activities in the comparison communities, which are not the same, but are partially related to the project's theory of change.	Low	Project exposure indicators suggest comparison group did not receive interventions from GEM partners or other NGOs.
Analysis				
6	Representativeness	Score LOW risk if: • During analysis or matching procedure less than 10% of the sample in the intervention group is excluded. Score HIGH risk if: • During analysis or matching procedure more than 10% of the sample in the intervention group is excluded.	Low	Less than 10% of intervention group sample is excluded.
7	Robustness checks	Score LOW risk if: • Magnitude and statistical significance of the results are approximately consistent with different econometric models Score HIGH risk if: • Results are not consistent with different econometric models and sub group analysis.	Med	Triangulation with qualitative data and other variables suggest results are consistent. However, no alternative econometric models have been reported (hence not assessed as low risk).
8	Triangulation	Score LOW risk if: • Results are triangulated and consistent with other evaluation methods within the same evaluation. • Results are triangulated and consistent with other data on the same project but from different evaluations. Score HIGH risk if: • Results are not consistent or triangulated with other evaluation methods.	Low	Triangulation with qualitative data from focus groups and key informant interviews.
9	Multiple hypothesis testing	Score LOW risk if: • Multiple hypothesis tests apply Benjamini-Hochberg or Bonferroni tests.	Med	No controls applied for multiple hypothesis testing. But pre-analysis plan was prepared and followed.

	Title	Description	Assessment	Description
		• The evaluation drafted a pre-analysis plan prior data analysis, and followed the plan. Score HIGH otherwise		
10	Clustering	Score LOW risk if: • Clustering is applied • Clustering was tested but rejected as providing higher standard errors than non-clustering estimates. Score HIGH otherwise.	Low	Clustering applied at village level.
Other				
11	Other	Any other issue reported by the evaluator.	Evaluation took place during an exceptional year with severe flooding. As far as possible the evaluation design tried to minimise bias but validity of results must be seen in light of the circumstances around which data gathering took place. The impact of floods is judged to have had a more severe effect on intervention households, so would make it more difficult to detect project impacts. As such, the results can arguably be seen as conservative estimates of the impact of GEM activities.	