

Evaluation title	Evaluation of the Satellite Index Insurance for Pastoralists in Ethiopia (SIPE) Programme: Impact Evaluation of the SIPE Pilot (2017 – 2019)	
Evaluation category and type	Impact Evaluation	
Post Hoc Quality Assessment (PHQA) – overall rating	Satisfactory: 84%	
<p>This study assesses the impact of an innovative insurance programme for pastoralists in arid areas in Ethiopia, named Satellite Index Insurance for Pastoralists in Ethiopia (SIPE) pilot programme. The impact evaluation uses primary data collections, household surveys at baseline and midline, and propensity score matching as the main quantitative methodology for causal attribution. In addition, it also provides qualitative interviews and mixed-methods analysis. The evaluation identified comparable areas where the insurance had not been rolled out yet and used them as a comparison group with a propensity score matching design. The analysis is competently done, the language is accessible, and the report provides targeted recommendations that are actionable. The main limitations on this impact evaluation refer to the timeline of data collection. The follow-up survey was collected only one year after the introduction of the insurance, representing a timeframe too short for permitting to learn any causal effects of providing drought insurance. In addition, during that year, no drought has happened, and therefore the insurance did not pay out since the index insurance threshold was not triggered. Accordingly, without insurance pay-outs, no impacts on food security or behavioural changes were expected. Not surprisingly, the evaluation did not find many statistically significant effects. This study would have benefited from further follow-up data collections, in particular after the occurrence of the drought. Overall, keeping this limitation of the evaluation design in mind, the analysis is technically well crafted, and the exposition of the findings and presentation is appropriate.</p>		
CRITERION 1: REPORT SUMMARY	Rating	Highly Satisfactory 94%
<p>The report summary is well structured, easy to read, and provides a very good representation of the whole report. It presents a clear and easy to understand description of the intervention, the evaluation questions, and the identification strategy. In particular, the report presents well the quantitative methods used in the evaluation without using unnecessary technical terms. The findings, conclusions, and recommendations are presented concisely, with a clear connection to the evaluation questions, allowing the reader to extract the key information and main learnings quickly.</p>		
CRITERION 2: OVERVIEW OF EVALUATION SUBJECT	Rating	Satisfactory 71%
<p>The report presented a good description of the Satellite Index Insurance for Pastoralists in Ethiopia (SIPE) pilot programme, the purpose of the intervention, the unit of intervention and the outcomes of interest. However, the report would have benefited if the theory of change behind the intervention had been presented in greater detail.</p>		
CRITERION 3: EVALUATION CONTEXT, PURPOSE, OBJECTIVES AND SCOPE	Rating	Satisfactory 87%
<p>The study presents a strong description of the context and the scope of the evaluation. It describes the institutions involved and how the results will inform the programme. A more extensive literature review and a description of how the study contributes to existing knowledge gaps would have added greater value also to an academic audience.</p>		
CRITERION 4: METHODOLOGY AND ANALYSIS	Rating	Satisfactory 83%
<p>The evaluation uses Propensity Score Matching (PSM) as the main evaluation design. The report clearly describes the methodology, data, limitations and potential sources of bias. The analysis is well crafted; however, the main weakness of the methodology is the absence of an experimental design, which was not possible due to operational constraints. The report is transparent about the limitations and presents them clearly and repeatedly in the report.</p>		
CRITERION 5: FINDINGS AND RESULTS	Rating	Satisfactory 82%
<p>The analyses in the report are well connected to the evaluation questions and overall framework of the intervention, with no questions left out. Heterogeneity analysis on the gender of the household's head enhances the quality of the report. The main limitation of this study refers to the short evaluation timeframe, which did not allow results to materialise. The follow-up data collection took place only one year after the introduction of the insurance programme,</p>		

in a year when no insurance pay-out took place. As a result, the measurable impacts on food security or behavioural changes were expected to materialise only on a limited scale.		
CRITERION 6: CONCLUSIONS/LESSONS	Rating	Satisfactory 83%
The conclusions are coherent, well structured and in line with the findings and evaluation questions. The lessons reported are specific and well-targeted. The report provides enough details on why the limited duration of the pilot phase hampered the capacity of the impact evaluation to measure the effects of the intervention.		
CRITERION 7: RECOMMENDATIONS	Rating	Satisfactory 88%
The recommendations are actionable and well-targeted, oriented to each key stakeholder. The recommendations would have benefited from greater attention on minorities and vulnerable groups and greater attention to the initial theory of change, showing which causal mechanisms are present and what would be required to improve the programme's impact.		
CRITERION 8: ACCESSIBILITY AND CLARITY	Rating	Highly Satisfactory 98%
The report is well written, clear and easy to read. The language is plain and accessible to a non-specialised audience. The report is well structured, with well-connected sections of the right length and details. Overall the report is well written in a professional tone, with no grammar errors, spelling mistakes. All the acronyms and data are consistently reported and referenced. Key messages are also provided using visual aids such as graphs, tables and boxes to highlight and facilitate access to key information. Most of the information is presented in the text of the report, with the annexes providing greater technical details and information.		

Integration of Gender Equality and Women's Empowerment (GEWE) considerations in the evaluation report based on the UN System-Wide Action Plan (UN-SWAP) Evaluation Performance Indicator (EPI) scorecard	
UN-SWAP EPI – individual evaluation score	Meets requirements: 8 points
The report meets the requirements and attains 8 out of 9 possible points. The evaluation was explicit in assessing the gender differential effect of the intervention, similarly to children as a vulnerable group.	

Post Hoc Quality Assessment – Rating scale and definitions at overall report and criteria levels	
Highly Satisfactory	<i>Definition at overall report level: Evaluation users can rely on the credible and useful evaluation findings provided and can use the evaluation with a high degree of confidence for decision-making. The report is considered an excellent example.</i>
	<i>Definition at criterion level: The criterion is addressed without any gaps or limitations.</i>
Satisfactory	<i>Definition at overall report level: Evaluation users can rely on the quality and credible evaluation findings provided and can use it with confidence for decision-making.</i>
	<i>Definition at criterion level: There are no significant gaps or limitations in addressing the criterion.</i>
Partly Satisfactory	<i>Definition at overall report level: Evaluation users can rely on the evaluation findings provided and may use it for decision-making noting that there are some gaps/shortcomings in the information provided.</i>
	<i>Definition at criterion level: There are some significant gaps or limitations in addressing the criterion.</i>
Unsatisfactory	<i>Definition at overall report level: Evaluation users can use some of the learning from the evaluation, noting that there are significant gaps/ shortcomings in the evaluation findings provided. The report may still contribute to decision making but should be used with caution.</i>
	<i>Definition at criterion level: There are critical gaps or limitations in addressing the criterion. Most of the required parameters are not met.</i>