

2026:1k

Sida Evaluation

Nordic Consulting Group Denmark

Final Report from the Evaluation of Sida's work with Poverty

Annex 3: Synthesis of Impact and Evidence Use Across Zambia, South Sudan
and Wider Experiences



Authors: Carsten Schwensen, Louise Scheibel Smed, John Rand, Anne-Lise Klausen and Ayla-Kristina Olesen Yurtaslan.

The views and interpretations expressed in this report are the authors' and do not necessarily reflect those of the Swedish International Development Cooperation Agency, Sida.

Sida Evaluation 2026:1k

Commissioned by Sida, Evaluation Team.

Published by: Sida

Copyright: Sida and the authors

Date of final report: 2026-06-16

Art.no.: Sida62873en

urn:nbn:se:sida-62873en

This publication can be downloaded/ordered from www.Sida.se/publications

Annex 3

Central Evaluation of Sida's work with Poverty - Phase II

Synthesis of

Impact and Evidence Use Across Zambia, South Sudan and Wider Experiences

Prepared by Carsten Schwensen, Louise Scheibel
Smed, Anne-Lise Klausen and Ayla-Kristina Olesen
Yurtaslan, Nordic Consulting Group Denmark.

June, 2026.

Abbreviations

ACLED	Armed Conflict Location & Event Data
AHS	Annual Household Surveys
BEAM	Building Effective and Accessible Markets
BGFZ	Beyond the Grid Fund Zambia
CAC	Comprehensive abortion care
CDSS	Community Development Support Services
CoS	Church of Sweden
CSO	Civil Society Organisation
DFID	Department for International Development
DHIS2	District Health Information System
DHS	Demographic and Health Survey
DWAs	District Women's Associations
EDISON	Energy Data and Intelligence System for Off-Grid Networks
EQ	Evaluation Question
ESP	Energy Service Provider
ET	Evaluation Team
FSDZ	Financial Sector Deepening Zambia
FFSWE	Finance for Food Security and Women Entrepreneurs
GBV	Gender-Based Violence
GoS	Government of Sweden
GoZ	Government of Zambia
HDP	Humanitarian-Development-Peace
HMIS	Health Management Information System
HPF	The Health Pooled Fund
HRBA	Human Rights-Based Approach
IDPs	Internally Displaced People
IHME	Institute for Health Metrics and Evaluation
ILO	International Labour Organisation
INCREASE	Increasing Climate Resilience in Energy & Agricultural Systems and Entrepreneurship Project
LCMS	Living Conditions Monitoring Survey
M&E	Monitoring & Evaluation
M4P	Making Markets Work for the Poor
MCDSS	Ministry of Community Development and Social Services
MDG	Millennium Development Goal
MDPA	Multidimensional Poverty Analysis
MIS	Management Information System
MLSS	Ministry of Labour and Social Security
MNDP	Ministry of National Development Planning
MOH	Ministry of Health
MSD	Market System Development
MSI	Marie Stopes International
NEFCO	Nordic Environment Finance Cooperation
NFNC	National Food and Nutrition Council
NGO	Non-Governmental Organisation
OTEA	Organic Trade in East Africa
OVP	Office of the Vice President
PaCC	Peace and Community Cohesion
PMT	Proxy Means Tests
PoG	Passing-on-the-Gift
RALS	Rural Agricultural Livelihoods Survey
RCT	Randomised controlled trial
REAP	Enhanced Agricultural Productivity and Resilience to Climate Change through Solar Powered Irrigation Project
SAP	Strengthened Accountability Programme
SCT	Social Cash Transfer

SDG	Sustainable Development Goal
SEK	Swedish Kroner
SGBV	Sexual & Gender-Based Violence
SRH	Sexual & Reproductive Health
SRHR	Sexual & Reproductive Health Rights
ToC	Theory of Change
ToR	Terms of Reference
TRAC	Tool for Results management and Appraisal of Contributions
UNDP	United Nations Development Fund
UNEG	United Nations Evaluation Group
UNJPSPS	United Nations Joint Programme on Social Protection
USD	United States Dollars
WEE	Women's Economic Empowerment Project
WEESMS	Women's Economic Empowerment through Strengthening Market Systems

ABBREVIATIONS	II
1. INTRODUCTION	1
1.1 REPORT STRUCTURE	2
2. ZAMBIA	3
2.1 EVIDENCE FROM FOUR IMPACT CASE STUDIES	3
2.2 EVIDENCE FROM FOUR ADDITIONAL DECENTRALISED EVALUATIONS	11
2.3 SYNTHESIS OF FOUR IMPACT CASES STUDIES PLUS FIVE ADDITIONAL ZAMBIA CASES	18
2.4 ACCESS TO AND USE OF NATIONAL STATISTICS IN ZAMBIA	19
3. SOUTH SUDAN	22
3.1 EVIDENCE FROM FOUR IMPACT CASE STUDIES	22
3.2 SYNTHESIS OF THE FOUR SOUTH SUDAN IMPACT CASE STUDIES	30
4. SYNTHESIS OF EVIDENCE FROM OTHER SIDA EVALUATIONS	32
4.1 EVALUABILITY OF CONTRIBUTIONS	32
4.2 EXISTENCE AND USE OF EVIDENCE FOR CONTRIBUTION DESIGN	34
5. OVERALL SYNTHESIS ON IMPACT AND EVIDENCE USE	39
5.1 SIDA'S CONTRIBUTION TO IMPACT: CROSS-COUNTRY PATTERNS	39
5.2 PREPAREDNESS FOR GENERATING EVIDENCE FOR IMPACT ASSESSMENT	40
5.3 SHORTCOMINGS IN THE ASSESSMENT, USE, AND QUALITY OF EVIDENCE FOR CONTRIBUTION DESIGN	42
6. REASONS FOR WHY IMPACT EVIDENCE WAS UNDERUSED IN SOME CASES	45
ANNEX - SELECTED SIDA EVALUATIONS	49

1. Introduction

This report presents a cross-country synthesis of how impact evidence has been used in Sida’s contribution management to inform design and evaluation processes. It brings together findings from eight impact cases studies conducted by the evaluation team (ET) in Zambia and South Sudan, complemented by lessons from other Sida evaluations. The objective has been to identify recurring patterns in evidence use, highlight both strengths and missed opportunities, and point to areas for improvements in Sida’s processes to strengthen the evidence base and potential impact.

The assessment draws on (i) four impact case studies from Zambia, (ii) four impact case studies from South Sudan, (iii) five additional decentralised evaluations from Zambia, and (iv) lessons from an additional 18 centralised and decentralised Sida evaluations conducted between 2018 and 2025.

Table 1. Overview of impact case studies in Zambia and South Sudan

Contribution	Sector/subsector
Zambia	
Musika	Agriculture, agribusiness, market system development (MSD)
UN Joint Programme on Social Protection in Zambia (UNJPSP) II	Social protection, social cash transfer, nutrition
Beyond the Grid Fund Zambia (BGFZ)	Renewable energy, MSD
Diakonia: Strengthened Accountability Programme (SAP)	Governance, gender, conservation farming, child labour, extractive resources
South Sudan	
UNFPA Country Programme	Sexual reproductive health and rights (SRHR), gender-based violence (GBV)
Church of Sweden Nexus Pilot Project	Democratic participation, civil society, livelihood, resilience
UNDP Peace & Community Cohesion Project	Peace building/conflict prevention
Health Pooled Fund (HPF)	Health services and health system strengthening

The additional 18 central and decentralised evaluations have been selected from Sida’s publication database applying the following selections criteria: the “Evaluations” category was first selected and combined with five selected “Topics”: “Economic Development”, “Gender Equality”, “Health”, “Humanitarian Assistance” and “Poverty”, reflecting the broader lines of focus of the Sida Poverty evaluation. This search generated 75 hits which were further screened for their relevance by the ET. Ultimately, this resulted in a selection of 18 evaluations (see full list in Annex). In addition, relevant external evaluations, academic and research literature have also been included in the discussions in the report.

This synthesis report constitutes an input to Phase II of the “Central Evaluation of Sida’s work with Poverty” and provides responses to three key evaluation questions:

- What were the key findings on Sida’s contribution to impact (including possible shortcomings)?
- How well were the contributions prepared for generating evidence (for ex-post/impact assessment)? Including identification of possible shortcomings/limitations.
- Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)? E.g. insufficient assessment use of evidence. If yes - why?

Thus, focus is on use and generation of evidence in Sida’s contribution management, with a view to generate learning on how Sida can strengthen this element in the future to enhance impact. The report mainly focuses on “what works”, hence it does not discuss “what to do” questions (e.g. relevance, composition of portfolio etc.), including the Multidimensional Poverty Analysis (MDPA). Likewise, while many good examples have been found, the report has a stronger focus on the shortcomings and the learning from these, in line with the key evaluation questions mentioned above.

1.1 Report structure

Following this introduction, Chapter 2 focuses on findings related to Zambia: Section 2.1 presents findings from each of the four impact case studies conducted by the ET and Section 2.2 presents evidence from the five additional decentralised evaluations conducted in Zambia in the period 2018 to 2025. Section 2.3 synthesises experiences from the Zambia contributions. Section 2.4 includes a discussion of the use of Zambian data and statistics in Sida’s contribution management. Chapter 3 focuses on the four South Sudan impact case studies conducted by the ET: Section 3.1 discusses each of the four impact case studies and Section 3.2 synthesises these experiences. In Chapter 4 is presented a synthesis of evidence from the additional 18 Sida centralised and decentralised evaluations. Chapter 5 presents an overall synthesis on impact and evidence use, based on all impact studies and evaluations covered in this report. Finally, in Chapter 6 is provided some reasons for why impact evidence may have been under-utilised in some cases.

2. Zambia

2.1 Evidence from four impact case studies

Sida's support to the Strengthened Accountability Programme (SAP) in Zambia

The Strengthened Accountability Programme (SAP) aims to empower citizens in resource rich communities to claim their rights to benefit from local extractive industries. Caritas Zambia has implemented the SAP in Zimba District in Southern Province of Zambia.

Overall impact assessment (from impact case study):¹ The strongest and most consistent impact was on household income, driven almost entirely by higher mining earnings, while diversification efforts such as conservation farming and goat rearing showed little or no effect due to low uptake and systemic constraints. The programme contributed to asset improvements, better access to water, increased use of protective gear, reduced child labour in mining, and more gender-equitable decision-making, though it did not reduce the prevalence of GBV.

How well did the contribution prepare for generating evidence (for ex-post/impact assessment)? The SAP impact study was possible only due to substantial reconstruction of the Theory of Change (ToC), new data collection, and methodological workarounds. In other words: SAP did not prepare well for the ex-post impact assessment, but the ET was able to generate credible findings despite the weaknesses. The programme lacked an evaluability ready design; impact indicators were not clearly defined, and the ET had to reconstruct the results chain retrospectively. Baseline data was missing and incomplete; not disaggregated for key groups; and not aligned with the impact/outcomes that SAP later aimed to influence. Monitoring systems focused on activities rather than outcomes, leaving no systematic tracking of cooperative performance, mining incomes, adoption of conservation farming, goat survival, or gender-related changes. Several components were added late and without corresponding indicators, making them inherently difficult to evaluate. SAP II also repeated weaknesses identified in SAP I,² indicating that earlier evidence was not sufficiently used to strengthen evaluability.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)? Learning and adaptation could have been explicit design features in SAP. Given the thin evidence base from working with remote mining communities, SAP could have been framed explicitly as a learning programme, with built-in experimentation and iteration. A design, where some interventions would be expected to fail, and include mechanisms which could either have been dropped or pivoted. Instead, SAP was largely designed as if the model was already known to be

¹ Sida/NCG (2026), Impact study of the Strengthened Accountability Programme, a programme for small-scale miners in Zambia. A case study as part of the Central Evaluation of Sida's work with poverty.

² End of Project Evaluation Report for the Diakonia, Zambia Country Office, Strengthened Accountability Programme (2014-2017), November 2017.

working. With this as the point of departure, Sida could have worked differently with impact evidence at the design stage:

First, evidence on extension, input access, aggregation and risk could have been better used to **prioritise and focus** on a smaller set of interventions with proven or promising impact, instead of working with a broad menu of interventions. Based on this, weaker or less-evidenced components could have been dropped or adjusted.

Second, **conservation farming** could have been designed as a pilot with strong input support, not as a core diversification pathway. SAP added conservation farming as a diversification strategy, expecting it to boost crop income. However, experiences presented in a large body of literature show that conservation agriculture in Zambia and Southern Africa has low and inconsistent adoption unless very specific conditions are met. For instance, Giller et al. (2009)³ show that conservation agriculture often fails to deliver benefits for smallholders because it is labour-intensive and requires sustained input access; Baudron et al. (2012)⁴ find that adoption is highly context-specific and often partial or temporary; and Ngoma et al. (2016)⁵ document that in Zambia, adoption of conservation farming is constrained by labour demands, risk perceptions, and limited access to fertiliser and equipment.

Third, **goat rearing** was treated as a simple asset-transfer solution without including veterinary support, disease-resistant breeds and market access strategies. Prior to SAP design, national and regional evidence already showed systemic constraints in relation to goat rearing, including a declining goat productivity due to climate shocks, disease and market disruptions. Studies conducted both by Indaba Agricultural Policy Research Institute (IAPRI)⁶ and FAO⁷ highlighted that small ruminant productivity in Zambia is constrained by disease control, veterinary services, feed availability and market access.

Fourth, SAP's design over-promised on results of **GBV** interventions (suggested by the partner) and a more evidence-aligned design would have either excluded GBV as an expected outcome or included a dedicated, multi-year GBV component. The SAP design implicitly assumed that gender-equality sensitisation would contribute to reductions in the prevalence of GBV. Global evidence is however unequivocal that short-term sensitisation is insufficient and that GBV prevention requires long-term, multi-component interventions, not short-term sensitisation.⁸

³ Giller, K.E. et al. (2009), *Agriculture, Ecosystems & Environment*.

⁴ Baudron, F. et al. (2012), *Experimental Agriculture*.

⁵ Ngoma, H. et al. (2016), *Food Policy*.

⁶ See e.g. IAPRI (2017), *The Status of Livestock Services in Zambia: Animal Health, Extension and Market Access*; IAPRI (2018), *Livestock Extension and Animal Health Services in Zambia*.

⁷ FAO (2018), *Small Ruminant Production in Zambia: Constraints and Opportunities*.

⁸ See e.g. Heise, L. (2011), *Violence Against Women (synthesised evidence)*; Ellsberg, M. et al. (2015), *Prevention of Violence Against Women and Girls: What Does the Evidence Say? The Lancet (major global review)*.

Sida's support to the United Nations Joint Programme on Social Protection in Zambia (UNJPSP)

United Nations Joint Programme on Social Protection in Zambia (UNJPSP)⁹ aims to reduce multidimensional poverty through social cash transfer (SCT) to the most vulnerable, as well as through strengthening of the institutional framework for social protection through enhancing legislation, unifying SCT projects in Zambia under one framework and strengthening the reliability of the cash transfer system.

Overall impact assessment (from impact case study):¹⁰ The strongest and most consistent impacts were observed in food security, dietary diversity, and child nutrition, particularly when SCT was combined with nutrition interventions such as the 1,000-day pilot where a top-up cash transfer was allocated to support pregnant women and children's diversified diet.¹¹ The programme also contributed to greater engagement in small-scale gardening and some improvements in urban micro-enterprise activity, though it showed no measurable impact on rural incomes or paid employment, and even a negative effect on asset accumulation. Targeting has been effective in rural areas but weaker in urban settings, and misconceptions around nutrition led to unintended negative effects on breastfeeding and vitamin A uptake. Despite reaching one-third of Zambia's population, the SCT's low transfer value and irregular payments limited its ability to counteract rising national poverty, though evidence suggests poverty would have increased even more in the absence of SCT support.

How well has the contribution prepared for generating evidence (for ex-post/impact assessment)?: The SCT is backed by high-quality Randomised Controlled Trials (RCTs), benefitting from a long tradition of research on Zambia's SCT system and operated within a relatively data-rich environment, drawing on national household surveys, administrative data from Zambia's Monitoring Information System (MIS), and a dedicated research component through the 1,000-day SCT+ nutrition pilot, which produced baseline, midline and endline data. This made it possible to assess impacts on food security, dietary diversity and child nutrition with reasonable confidence. However, identified weaknesses in MIS (access to data) inherited from Phase I - problems with payments, grievance systems and coordination - were not fully resolved before Phase II began. As a result, the impact study had to rely heavily on secondary data and external studies, rather than on a robust programme-generated M&E system. Irregular and unpredictable payments have also been a challenge, complicating both programme impact and its measurement.

⁹ It is to be noted that UNJPSP was implemented in conjunction with the Girls' Education and Women's Empowerment and Livelihoods (GEWEL) project, and that the two addresses complementary aspects of social protection in Zambia. GEWEL, however, is not covered by this impact assessment.

¹⁰ NCG/Sida (2026), Impact study of United Nations Joint Programme on Social Protection in Zambia. A case study as part of the Central Evaluation of Sida's work with Poverty.

¹¹ It is to be noted that the pilot was implemented in only two districts over 2.5 years, thus its findings may not fully reflect the broader programme.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?: Although SCT is grounded in a long history of social protection work in Zambia, important evidence was not fully used to shape the programme's design:

First, according to literature, the **SCT benefit** - equivalent to only USD 0.25 per day per household - is too small to reduce poverty, yet the design did not provide extensive evidence showing that very small transfers may have limited effects on income, food security and nutrition. This is, however, well established in global reviews, which show that cash provision alone rarely lead to reduced poverty unless transfers are large enough or paired with complementary interventions.¹²

Second, **targeting in urban areas** has been less convincing, despite strong prior evidence that proxy means tests (PMTs) perform poorly in urban settings where incomes are volatile and informal. Various studies had already documented these limitations, yet the programme continued to rely heavily on PMTs without adapting the design.¹³

Third, the programme also underestimated **behavioural responses** to cash transfers, e.g. in relation to breastfeeding and vitamin A uptake, where there were misconceptions among beneficiaries. This reflects a design that did not fully incorporate evidence showing that cash transfers can unintentionally shift consumption and care practices unless accompanied by strong behaviour-change communication.¹⁴

Fourth, the design did not fully address **system readiness**, even though Phase I had already demonstrated critical weaknesses in payment reliability, MIS functionality, grievance systems and coordination. While some literature on social protection systems emphasises that these foundations must be stabilised before scaling,¹⁵ interviews revealed that improving system design and scaling is almost always done in parallel as building/improving systems is seen as a continuous process. Here, it may also be necessary to differentiate between evidence where Sida can actually influence the design of a national (highly political) government owned programme and what may be outside Sida's influence sphere.

¹² See e.g. Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T., & Pellerano, L. (2016), Cash transfers: what does the evidence say? ODI; Davis, B., Handa, S., Hypher, N., Rossi, N., Winters, P., & Yablonski, J. (2016), From Evidence to Action: The Story of Cash Transfers and Impact Evaluation in Sub-Saharan Africa. FAO & UNICEF; Manley, J., Balarajan, Y., Malm, S., et al. (2020), Cash transfers and child nutritional outcomes: a systematic review and meta-analysis. *BMJ Global Health*.

¹³ See e.g. Brown, C., Ravallion, M., & van de Walle, D. (2018), A poor means test? *Econometric targeting in Africa*. *Journal of Development Economics*; Coady, D., Grosh, M., & Hoddinott, J. (2013), Targeting of Transfers in Developing Countries: Review of Lessons and Experience. World Bank; Beegle, K., Coudouel, A., & Monsalve, E. (2018), Realizing the Full Potential of Social Safety Nets in Africa. World Bank.

¹⁴ See e.g. Aker, J. C., Boumnijel, R., McClelland, A., & Tierney, N. (2016), Payment mechanisms and antipoverty programs: Evidence from a mobile money cash transfer experiment in Niger. *Economic Development and Cultural Change*; Manley, J., Balarajan, Y., Malm, S., et al. (2020), Cash transfers and child nutritional outcomes: a systematic review and meta-analysis. *BMJ Global Health*.

¹⁵ See e.g. Barca, V. & Chirchir, R. (2014), Building Social Protection Systems: Integrated Information Management Systems (MIS) for Social Protection. DFID; Pellerano, L. & Barca, V. (2017), Social Protection Delivery Systems: Synthesis Report. Oxford Policy Management.

Sida's support to the Musika Project in Zambia

Musika is a national Non-Government Organisation (NGO), registered in July 2010 with the aim to reduce poverty and create wealth in rural Zambia by stimulating business relationships between the corporate sector and smallholder farmers. In this process, Musika's role has mainly been to serve as a market facilitator.

Overall impact assessment (from impact case study):¹⁶ Musika generated clear, positive poverty-reducing effects for smallholder farmers in supported districts, particularly through improved market access, better access to finance and increased income, livestock holdings, and food security among direct beneficiaries. Other core impact pathways - such as farm productivity, technology adoption, and asset accumulation - showed no significant improvements, and the programme's contribution was far too small to offset worsening national poverty trends. Although Musika aimed to catalyse systemic market change, the evidence indicates that these effects were partial and fragile, with improved service delivery and reduced transaction costs observed mainly among firms already positioned to benefit. Musika helped strengthen some market functions, but the scale and depth of change fell short of its transformative objectives. M&E data did not allow for assessment of gender impact since data were not gender-disaggregated, despite gender targets being established.

How well has the contribution prepared for generating evidence (for ex-post/impact assessment)?: Musika was significantly better prepared for ex-post impact assessment than most other contributions in the Zambia portfolio, but its evaluability was still constrained by several structural weaknesses. Musika generated large, repeated, quantitative datasets over a decade, especially the Annual Household Surveys (AHS), which made a rigorous quasi-experimental assessment possible. However, these data were not originally designed for impact evaluation, and inconsistencies across survey rounds limited their usefulness. Thus, the impact case study succeeded mainly because of the volume and continuity of data, not because Musika had been deliberately designed for impact evaluation.

The strongest aspect of Musika's evaluability was the availability of longitudinal household-level data. The AHS explicitly identify both treated and non-treated households and track them over time, enabling matched Difference-in-Differences analysis. This is rare in Sida's Zambia portfolio and provided a solid foundation for assessing income, food security, assets and market access. The use of the national household survey, the Rural Agricultural Livelihoods Survey (RALS) data for Phase 2 further strengthened the evidence base, even though these data required pseudo-panel construction. However, like AHS, various shortcomings and limitations were also linked to the RALS data set, including inconsistent questionnaires, missing variables, and changes in methodology that prevented direct comparison across years. Some indicators, especially on productivity and technology adoption, could only be analysed to a limited extent due to data gaps.

¹⁶ Sida/NCG (2025), Impact study of Musika, a Zambian rural project. A case study as part of the Central Evaluation of Sida's work with Poverty.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?: Musika’s design was driven by the Market System Development (MSD) narrative, facilitating private sector engagement with smallholders, more than by hard evidence on which facilitation models actually worked in Zambia. If Sida had worked differently with impact evidence at the design stage, this could potentially have improved the contribution design in some ways:

First, the MSD model could have become more **focused and testable**. By the time Phase II of Musika was designed, substantial evidence already existed on which value chains and business models had the strongest potential for inclusive growth in Zambia. Studies by IAPRI and others had shown that smallholder market participation was highly uneven, and that only certain crops and service models consistently delivered benefits for poorer farmers.¹⁷ Early Musika experience also indicated that some facilitation instruments - such as input distribution partnerships and extension models - were more promising than others. A design grounded in this evidence could have led to prioritisation of a smaller set of value chains and facilitation tools with explicit, testable hypotheses.

Second, **partner selection and roles** could have been more evidence-based. Comprehensive research on agribusiness partnerships in Zambia had already shown that only certain types of firms and contractual arrangements reliably deliver inclusive outcomes.¹⁸ This evidence could have informed tighter criteria for selecting private sector partners, clearer expectations for their contributions, and more realistic assumptions about incentives and risk-sharing.

Third, Musika’s **system-change ambitions** could have been tied to measurable indicators. The MSD field had already begun to articulate concrete system-level metrics - such as sustained service provision, improved price transmission, reduced transaction costs, and durable risk-sharing arrangements.¹⁹ Instead of relying on generic “market transformation” language, Musika’s design could have linked its system-change claims to these evidence-based indicators, making the model more evaluable and grounded in observable market dynamics.

Fourth, in terms of **gender**, existing evidence (also generated with support from Sida’s gender help desk) could have helped Musika identify which types of firms and distribution models were more likely to reach women farmers, and to set gender-sensitive criteria for partnerships and matching grants. MSD specific literature, such as the Springfield Centre’s

¹⁷ See e.g. Mason, N. & Jayne, T. (2013), Fertilizer subsidies and smallholder commercialisation in Zambia. Food Policy; Kuteya, A., Chapoto, A., Malabo, M. & Zulu-Mbata, O. (2016), The Politics of Maize in Zambia. IAPRI; Sitko, N. & Jayne, T. (2014), Exploitative briefcase buyers or benevolent middlemen? The role of intermediaries in Zambia’s agricultural markets. Food Security; Chapoto, A. & Chisanga, B. (2016), Medium-Scale Farms and Structural Transformation in Zambia. IAPRI.

¹⁸ See e.g. Sitko, N. & Jayne, T. (2014), Exploitative briefcase buyers or benevolent middlemen? The role of intermediaries in Zambia’s agricultural markets. Food Security; Chapoto, A. & Chisanga, B. (2016), Medium-Scale Farms and Structural Transformation in Zambia. IAPRI.

¹⁹ E.g. Springfield Centre (2014), The Operational Guide for the Making Markets Work for the Poor (M4P) Approach; DCED (2017), Measuring Systemic Change: A Guidance Note; Katalyst Bangladesh (2013–2015), Lessons on systemic change measurement (widely used in MSD design globally).

guidance on gender-responsive market systems²⁰ and Building Effective and Accessible Markets (BEAM) Exchange case studies,²¹ had already shown that market facilitation often reinforces existing gender inequalities unless gender is explicitly integrated into partner selection, business models, and monitoring systems.

Sida's support to the Beyond the Grid for Zambia Project (BGFZ)

Beyond the Grid Fund Zambia (BGFZ). was designed to create access to basic and affordable renewable energy services for energy poor people living outside the national electricity grid. This was to be achieved by incentivising private sector companies' expansion and investment in off-grid energy generation, distribution and service delivery. Overall, BGFZ aimed at enabling access to clean, modern energy services for more than one million Zambians living in peri-urban and rural areas.

Overall impact assessment (from impact case study):²² BGFZ achieved impressive scale and reach, enabling more than one million Zambians - mostly first-time users - to access basic off-grid energy services and experience clear improvements in daily quality of life. However, the evidence base shows that these gains were largely confined to basic energy access, with little verified impact on health, education, or productive economic activity, despite such outcomes being central to the programme's ambitions. While BGFZ successfully demonstrated a viable model for accelerating private-sector entry into off-grid markets, it struggled to effectively reach the poorest and most remote households, and a sustainable last-mile business model remains undeveloped.

How well has the contribution prepared for generating evidence (for ex-post/impact assessment)?: BGFZ was better prepared than most Sida contributions in Zambia because it invested heavily in data systems, but it was not fully evaluability-ready. On the positive side, BGFZ created an unusually strong real-time monitoring infrastructure. The Energy Data and Intelligence System for Off-Grid Networks (EDISON)²³ platform for M&E data required daily reporting from Energy Service Providers, and customer-level outcomes were independently verified by 60 Decibels (an end-to-end impact measurement company), giving the programme a rich evidence base on reach, service quality and customer experience. This made descriptive analysis and verification of outputs possible at scale. However, BGFZ was not designed for rigorous ex-post impact evaluation. The ET could not access raw EDISON data due to system phase-out and legal constraints, forcing reliance on secondary analyses. Moreover, EDISON tracked mainly operational indicators - and to less extent poverty outcomes, productive use, income effects or welfare changes - leaving major impact dimensions unmeasured. Thus, BGFZ generated a lot of monitoring data which were very useful for

²⁰ The Springfield Centre (2014), The Operational Guide for the Making Markets Work for the Poor (M4P) Approach. 2nd Edition (includes guidance on gender-responsive market systems).

²¹ BEAM Exchange (2016), Gender and Market Systems Development: A Practical Framework for Integrating Gender into MSD Programmes. BEAM Exchange, London.

²² NCG/Sida (2025), Impact study of Beyond the Grid for Zambia. A case study as part of the Central Evaluation of Sida's work with Poverty.

²³ Recently, EDISON was launched a new, improved version (Prospect).

performance management but insufficient for evaluating long-term poverty and equity effects.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?: Although BGFZ aimed to expand energy access and stimulate economic activity, the design could have drawn further on the substantial body of pre-existing evidence on energy access, poverty outcomes, productive use, and last-mile distribution in Sub-Saharan Africa:

First, the original design did not define **measurable poverty and equity outcomes** (income, education, health, time-use), despite extensive literature linking energy access to multidimensional poverty.²⁴ BGFZ did not embed such indicators in its design.

Second, the design relied heavily on **self-reported operational data** from private sector partners, even though the literature consistently has warned that private sector reporting in impact-oriented energy programmes requires independent verification due to risks of over-reporting, inconsistent definitions, and weak traceability.²⁵ BGFZ only partially addressed this through 60 Decibels, and only for customer-level indicators.

Third, the design assumed that energy access would catalyse **economic activity**, an expectation not supported by the evidence. A well-established literature shows that electrification alone rarely increases income or enterprise productivity.²⁶ Productive use requires complementary interventions such as access to finance, appliances, training, and market linkages. BGFZ's design did not incorporate these complementary elements, despite their centrality in the evidence.

Fourth, the design did not fully integrate evidence on high rural distribution costs, weak demand, and the need for tailored **last-mile strategies**. Several studies already documented the structural challenges of reaching rural customers with off-grid products.²⁷ By not incorporating this evidence, BGFZ underestimated the scale of distribution constraints and overestimated the speed at which private sector partners could expand into remote areas.

²⁴ See e.g. Khandker et al. (2014), Energy access impacts require explicit measurement frameworks; and Pachauri et al. (2013), Energy-poverty linkages must be operationalised through indicators.

²⁵ E.g. Lighting Africa / World Bank (2010–2015), Off-Grid Solar Market Assessments and Impact Evaluations; GOGLA (2015), Standardised Impact Metrics for the Off-Grid Solar Sector.

²⁶ See e.g. Pueyo, A. (2013) 'The Evidence of Benefits for Poor People of Increased Renewable Electricity Capacity: Literature Review', IDS Evidence Report; Grimm, M., Hartwig, R. & Lay, J. (2013). Electricity access and the performance of micro and small enterprises: Evidence from West Africa. World Bank Economic Review; Bernard, T. & Torero, M. (2015). Social interaction effects and connection to electricity: Experimental evidence from rural Ethiopia. Economic Development and Cultural Change; Lee, Miguel & Wolfram (2020), Does Household Electrification Supercharge Economic Development? American Economic Review; Bernard (2012): productive use requires targeted support and appropriate appliances.

²⁷ See e.g. Roloffs, P., Byrne, R. & Ockwell, D. (2015), Financing sustainable energy for all: Pay-as-you-go solar in Kenya and Tanzania. Energy Research & Social Science; Taneja, J., et al. (2016), Rural energy markets and last-mile distribution. Energy for Sustainable Development; ACE-TAF (2017-2019), Off-Grid Solar Market Assessments for East and Southern Africa.

Synthesis of the four impact cases:

Across the four Zambia case studies, a clear relationship emerges between Sida's contribution to impact, the evaluability of the programmes, and the quality of evidence used at design stage:

Sida's contributions did generate important results: Higher incomes in SAP, improved food security and child nutrition in the SCT programme, better market access and income gains in Musika, and large-scale energy access through BGFZ. Yet these impacts were uneven, and the weaker areas of performance relate closely to weaknesses in contribution designs. SAP's diversification components underperformed, Musika's system-change ambitions remained partial, and BGFZ's broader welfare effects could not be verified. These shortcomings were not simply implementation issues; they reflected design choices that did not fully incorporate existing evidence on what works in agriculture, social protection, market systems and off-grid energy: SAP added conservation farming and goat rearing despite extensive evidence of low adoption and systemic constraints; SCT maintained a transfer size too small to reduce poverty; Musika's design leaned more on the MSD narrative than on empirical evidence about which value chains and facilitation models work in Zambia; and BGFZ assumed that electrification would drive economic activity despite a large literature showing that energy access alone rarely does so. These gaps were often reinforced by over-ambitious intervention designs, pressure to scale quickly, and insufficient integration of academic and policy research.

Taken together, the four Zambian impact case studies showed that Sida's ability to contribute to impact is inseparable from the quality of evidence used at design stage and from the evaluability of the interventions it supports.

2.2 Evidence from four additional decentralised evaluations

This section includes experiences from five other comprehensive project/programme evaluations in Zambia, conducted in the period 2018-2025.

Sida's support to the Increasing Climate Resilience in Energy & Agricultural Systems and Entrepreneurship Project (INCREASE) in Zambia

Overall results assessment (from the Midterm Evaluation (MTE) of INCREASE):²⁸

The MTE finds that INCREASE was highly relevant and had clear potential, but overall results were found to be below expectations due to an overly ambitious design, delayed implementation, and weak market and institutional linkages. While farmers reported early gains, especially in cotton and dairy, training roll-out, input delivery, renewable energy components, and market access all faced major setbacks, and gender and youth ambitions were only partially realised. The MTE found it to be highly unlikely that the project would be able to demonstrate its full model or ensure sustainability within the given timeframe.

²⁸ Sida/Nordic Consulting Group (Carsten Schwensen, Louise Smed) (2023), Midterm Evaluation of Increasing Climate Resilience in Energy & Agricultural Systems and Entrepreneurship (INCREASE).

How well has the contribution prepared for generating evidence (for results assessment)?: INCREASE was not well prepared for generating evidence that would support a rigorous ex-post or impact assessment. Although the project had strong ambitions and a broad results framework, its monitoring and learning systems were too weak, fragmented, and delayed to provide reliable outcome-level evidence. The most significant issue was that the M&E system was not functioning effectively. It was only to a limited extent useful for assessing project progress, and the data collected did not provide much insight into actual results achieved. A new M&E software system was introduced but remained cumbersome and not fully operational, leaving the project without a stable platform for tracking change. Data collection was inconsistent and largely activity-based, with implementing partners collecting their own data in silos. This meant that information was not comparable across value chains or provinces, and much of it consisted of simple counts rather than evidence of behavioural, environmental, or market-level change. Finally, INCREASE suffered from early management gaps and staff vacancies, which delayed training, slowed data collection, and weakened documentation. These disruptions further reduced the project’s ability to generate consistent evidence over time.

The **INCREASE** project was an SNV-implemented initiative designed to strengthen the social, economic and environmental resilience of smallholder farmers in Zambia by integrating climate-smart agriculture, renewable energy solutions (including biodigesters), and support to women- and youth-led enterprises. It combines two earlier SNV programmes in Zambia (E4A and SILMS) and works across cotton, dairy and horticulture value chains in five provinces. Sida has funded the project with SEK 78 million for the period January 2020 to December 2022, later extended at no cost to December 2023.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?: INCREASE is the clearest example across all nine Zambia contributions of over-ambitious contribution design relative to existing evidence and capacity. The INCREASE MTE (2023) explicitly noted that “*a more rigorous early value-chain and market-system analysis could have prevented over-ambitious design*”. The contribution design would likely have been improved in at least four ways, if Sida had worked differently with impact evidence: First, early evidence on which **value chains, technologies and practices** actually work in similar contexts could have led to a more focused design (fewer value chains, fewer technologies, clearer geographic focus). Second, instead of a broad “MSD-climate-energy” ambition, Sida could have insisted on use of **evidence-based MSD practice** (e.g. proven facilitation models, realistic partner roles) or even decided to drop the MSD label if the evidence and capacity weren’t there. Third, use of existing evidence on water resources, environmental impacts and climate risks could have been used to design real **human-environment nexus** interventions, rather than a climate-aware agriculture add-on. Fourth, Sida could have required that pilots (in particular those related to biodigesters) would meet minimum evidence and sustainability standards before being scaled.

Sida's support to the Women's Economic Empowerment Project (WEE) in Zambia

Overall results assessment (from the Midterm Evaluation (MTE) of WEE):²⁹ The MTE found WEE to be highly relevant and the project had generated early positive results - especially increased incomes, stronger women's leadership, and successful Passing-on-the-Gift (PoG) cycles (handoff of a goat to another family) - but overall progress was found to be uneven due to weak coordination, inconsistent targeting, delays in training and inputs, and major capacity gaps within the supported District Women's Associations (DWAs). While many vulnerable women were benefiting, others struggled to fully access the project package, and key components such as finance, GBV/SRHR linkages, and climate-resilience activities remained only partially implemented. The MTE therefore found it highly unlikely that WEE would achieve its full intended outcomes without stronger management and more coherent support structures.

How well has the contribution prepared for generating evidence (for results assessment)?:

The WEE project was poorly prepared for generating evidence suitable for an ex-post or impact assessment. The project lacked the basic monitoring structures needed to track outcomes, verify progress, or assess impact. Most importantly, the project had no systematic data collection system in place. In addition, the results framework was overambitious and poorly defined, with indicators that were weakly conceptualised and difficult to measure. Monitoring responsibilities were unclear, and each implementing partner collected its own data in incompatible formats, mostly simple counts rather than outcome-level information. Finally, the project placed unrealistic monitoring responsibilities on District Women's Associations, despite their limited administrative capacity. Weak governance, high staff turnover and limited technical skills further undermined data quality.

The WEE project in Zambia aimed to strengthen women's agency, decision-making power, access to productive resources, and participation in agricultural value chains across eight districts in Zambia. It was implemented by We Effect together with Women for Change, the Zambian Land Alliance and Heifer International, and worked through District Women's Associations to deliver training, crop and livestock support, land rights activities, climate-smart agriculture, and gender-transformative community engagement. Sida funded the project with SEK 75 million for the period November 2019 to July 2023.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?:

WEE was based on a strong concept, but the design was not anchored in an analysis of who benefits and who was at risk. WEE was aiming at being gender-transformative and multi-dimensional, but the WEE MTE (2023) highlighted "gender-transformative intent vs operational reality" and weak data systems. It also noted important negative effects (around 20% of PoG beneficiaries struggling, sometimes selling assets to repay). Based on the MTE, three design elements would likely have been different if Sida had worked differently with impact evidence at the design stage: First, **targeting and risk management** would have been built stronger on evidence: Existing evidence on PoG, women's economic empowerment and intra-household dynamics could have been

²⁹ Sida/Nordic Consulting Group (Carsten Schwensen, Louise Smed) (2023), Midterm Evaluation of Women's Economic Empowerment (WEE).

used to design clear risk thresholds, support mechanisms and exit options for vulnerable women, rather than discovering the 20% struggling group midway through. Second, use of existing evidence on the links between economic empowerment, SRHR and GBV could have been used to design WEE more explicitly as a **multi-dimensional empowerment contribution**, with SRHR/GBV components integrated from the start, and not only added as a narrative afterthought. Third, use of existing evidence on the fragility of DWAs and similar structures (including from INCREASE) could have led Sida to design **organisational development, governance and financial sustainability** as central institutional pillars, and not only peripheral capacity building.

Sida's support to the Women in Financial Inclusion and Capabilities Project (WIN) in Zambia

Overall results assessment (from the Midterm Evaluation (MTE) of WIN):³⁰ The MTE of the WIN project found that the project showed promising and often strong early results, despite a shorter-than-planned implementation window. It effectively supported expansion of women's and youths' access to savings groups, digital financial services and financial literacy, with several partners reaching or surpassing targets ahead of schedule. Beneficiaries reported increased confidence, improved budgeting skills and greater financial security, while digital tools (mobile phones, tablets, apps) proved highly effective for learning and service uptake. The Making Markets Work for the Poor (M4P) approach also fostered collaboration among market actors and NGOs, creating emerging networks and synergies. However, delays in partner onboarding, complex delivery chains and uneven documentation limited transparency and slowed progress.

How well has the contribution prepared for generating evidence (for results assessment)?:

The WIN project was found to be moderately prepared to generate evidence for results assessment, with solid monitoring systems at Financial Sector Deepening Zambia (FSDZ) and detailed results chains for each sub-project. Routine data collection on activities, outputs and early outcomes - especially around financial literacy training, digital service uptake and savings group participation - provided a reasonable evidence base for assessing progress. However, the project's complex M4P structure, multiple partners and flexible, adaptive design meant that documentation was uneven, data aggregation was challenging and reporting to Sida lacked detail. Several sub-projects had limited or inconsistent documentation, making it difficult to track change over time and attribute outcomes to the supported interventions. In addition, while women, girls and youth were central to the project's design, evidence on intersectional vulnerability (poverty, age, disability, remoteness) was limited as the project did not systematically generate evidence

The WIN Project in Zambia was implemented from 2016 to 2020 with the aim to increase financial inclusion for women and youth, helping them improve income, reduce vulnerability, and enhance wellbeing. It was implemented by the FSDZ with a Swedish contribution of SEK 28 million. The project used a M4P approach and worked through private sector actors (banks, mobile providers), partners with NGOs and community groups and focuses on sustainable market-based solutions rather than direct aid.

³⁰ Sida/Niras (2019), Midterm Evaluation of Women in Financial Inclusion and Capabilities Project in Zambia.

on women empowerment and vulnerability outcomes. FSDZ did not consistently capture who among the beneficiaries was most vulnerable (e.g. levels of poverty, disability status, rural/urban disadvantage, barriers faced by adolescent girls etc.). Implementing partners gathered some qualitative insights on how women and youth experienced financial exclusion, but this information was not structured, comparable or integrated into the results framework. As a result, the project could not demonstrate whether the most marginalised women and girls were being reached, nor how different sub-groups benefited. Finally, the project did not include a plan for longitudinal tracking of beneficiaries or for measuring financial inclusion outcomes such as savings behaviour, loan uptake, or financial decision-making.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?: The MTE noted good qualitative and monitoring data, but also that the “*evidence systems struggled to keep up*” with the innovation narrative. If Sida had worked differently with impact evidence at design stage, the contribution design could potentially have been improved in at least three areas:

First, the design could have benefitted from better use of existing evidence on how to build **mechanisms to link financial education and digital services to actual behavioural outcomes**. The design focused heavily on outputs (trainings, group formation, digital tools) but lacked tools to measure behavioural change - such as improved saving habits, reduced vulnerability, or increased financial resilience. Simple follow-up surveys or outcome-tracking tools could have helped to capture these longer-term effects more systematically. WIN could have been designed with phased learning and adaptation built in.

Second, the project could have introduced **systematic vulnerability and inclusion indicators** and more consistently collected data on poverty levels, disability, age-based vulnerability and rural/urban disadvantage. Adding simple, standardised indicators - such as household income category, disability status, adolescent participation - would have allowed the project to better track whether the most marginalised women and girls were being reached.

Third, the **project timeline** (of three years) design could have relied stronger on general experiences from M4P projects (which would generally recommend at least a five-year project period). This would have allowed more systematic evidence generation, better sequencing of activities and more robust tracking of change over time, as well as the ability to generate deeper outcome-level evidence, especially on behavioural change and long-term financial resilience.

Sida's support to the National Legal Aid Clinic for Women Access to Justice Programme (NLACW) in Zambia

Overall results assessment (from the Evaluation of NLACW):³¹ The evaluation concludes that the NLACW Access to Justice Programme delivered strong and often above-target results during the 2019-2023 period, providing high-quality free legal representation, mediation, counselling and legal literacy to thousands of vulnerable women and children across Zambia. NLACW's advocacy contributed meaningfully to strengthening gender-responsive laws and justice sector practices and the organisation was found to be widely respected, well-connected and highly relevant to Zambia's justice needs. However, sustainability challenges remained due to heavy reliance on a few donors, overstretched staff and paralegals, and the need for more strategic planning and clearer approaches to poverty, disability inclusion and long-term funding.

The NLACW Access to Justice Programme in Zambia is a long-standing initiative that provides free legal representation, counselling, and legal literacy to vulnerable women and children, while also engaging in advocacy to strengthen gender-responsive laws and justice systems. Implemented under NLACW's 2019-2023 Strategic Plan, the programme operates across multiple provinces through legal services, community outreach, paralegal networks, and policy engagement, with Sida providing core funding to support its mission of improving access to justice, gender equality, and protection of rights for women and children.

How well has the contribution prepared for generating evidence (for results assessment)?: NLACW was relatively well-prepared for generating evidence for an ex-post or impact assessment compared to many other civil-society programmes in Zambia. The organisation has structured monitoring systems, clear reporting routines, and a results framework that allowed evaluators to document outputs and many outcomes with reasonable confidence. A key strength is that NLACW has long-standing administrative systems for tracking legal cases, client profiles, legal literacy activities, and paralegal work and monitoring and reporting allowed for substantial documentation of results. Case records, activity logs, and legal literacy outreach data provided a solid empirical base for assessing performance against targets. However, the programme was less prepared for generating evidence on deeper behavioural or systemic change. Some indicators were not fully developed and NLACW lacked explicit strategies for monitoring poverty, disability inclusion, and long-term empowerment outcomes. Likewise, data on client vulnerability was limited, as eligibility assessments relied mainly on employment status and general appearance rather than systematic socio-economic profiling.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?: The evaluation of NLACW shows massive reach, unique positioning (the only organisation with a primary focus on women and children), and clear practice change. Yet the contribution is mainly designed for a CSO partner and not framed as a national access-to-justice model. If Sida had worked differently with impact evidence

³¹ Sida/FCG (2024), Evaluation of the National Legal Aid Clinic for Women (NLACW) Access to Justice Programme in Zambia.

in this case, the design could potentially have been improved in at least three ways: First, early evidence from NLACW's own practice (and from global legal empowerment work) could have pushed Sida to build more specific **justice outcomes** - resolution rates, time to resolution, recurrence of GBV, user satisfaction – and **poverty/inclusion targeting** into the design as core metrics, not as evaluation afterthoughts. Second, with impact evidence in hand, Sida could have designed NLACW explicitly as **a model to be scaled or replicated** (e.g. through state systems, bar associations, or paralegal networks), rather than as a stand-alone NGO project. Third, evidence on **funding fragility** could have led to a design that included more diversified financing, institutional partnerships, and capacity building as explicit components, and not only as risks to be monitored later.

Sida support to the International Pregnancy Advisory Services Project (Ipas) in Zambia

Overall results assessment (from the End of Project Evaluation of Ipas):³² The evaluation found that the Ipas programme was highly effective, consistently meeting or surpassing targets for strengthening comprehensive abortion care (CAC) services and expanding community awareness between 2018 and 2022. Facility upgrades, essential commodities, mentoring and service provider training significantly improved service quality, while community engagement and work with traditional, religious and civil-society actors helped shift attitudes and increase acceptance of safe abortion. Ipas filled a critical gap as the main CAC provider in target areas and worked efficiently and cost-effectively. However, long-term sustainability remains uncertain due to reliance on external funding and limited government capacity to maintain material support.

How well has the contribution prepared for generating evidence (for results assessment)?:

The Ipas programme was found to be well-positioned to generate credible evidence for results assessment, with one of the most structured monitoring systems among Sida-funded interventions in Zambia. Its routine data collection was a major asset: Ipas consistently documented service delivery, CAC and post-abortion family planning uptake, client exit interviews, facility upgrades, commodity availability and training across health workers, police, media and community actors. This produced reliable quantitative evidence on outputs and many outcomes, and the organisation actively used this information to refine mentoring, facility support and community engagement, reflecting a solid learning culture. At the same time, the programme was found to be less well-equipped to measure deeper social and behavioural change, such as shifts in community norms, long-term effects of awareness-raising or differential impacts on vulnerable groups, where evidence remained

Sida's support to **Ipas** in Zambia aimed to expand women's and girls' access to CAC and broader SRHR services by strengthening health-system capacity and increasing community awareness across provinces. Implemented over the 2018-2022 period, the support combined training, mentoring, facility upgrades, equipment and drug provision, and community mobilisation to build a more "sustainable abortion ecosystem." Sida's total financial contribution, including a no-cost extension phase that ended in June 2022, was SEK 41 million.

³² Sida/Niras (2022), End of Project Evaluation of the Sida-funded Ipas program in Zambia.

largely qualitative. Thus, overall Ipas was found to be highly evaluable on service-level and institutional results, but less prepared to capture broader societal transformation and long-term behavioural outcomes.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?: The end-of-project evaluation showed that Ipas had strong systems for tracking service-level results, but the programme design placed less emphasis on measuring community-level behavioural change, shifts in social norms, or differential impacts on vulnerable groups. These gaps were not due to weak implementation but to limited indicators, tools and data strategies for capturing social change. As a result, evidence on attitudes, stigma reduction and long-term behavioural outcomes remained largely qualitative and episodic. The programme also lacked structured mechanisms to track how awareness-raising translated into actual service uptake or empowerment outcomes. Thus, while the monitoring system was found to be excellent for operational learning, it was less suited for deeper outcome/impact measurement, meaning that contribution design did not fully integrate evidence needs for long-term social change. With more explicit outcome indicators, baseline data and community-level monitoring tools, these gaps could potentially have been reduced in different way:

First, by including **structured indicators and tools for measuring social and behavioural change**. The programme relied heavily on qualitative insights for shifts in norms, attitudes and stigma. With clearer indicators, baseline data and simple community-level monitoring tools (e.g., periodic attitude surveys, structured community scorecards), Ipas could have generated stronger evidence on how awareness-raising translated into behavioural change.

Second, by integrating **systematic tracking of differential impacts on vulnerable groups**. Although the programme aimed to reach adolescents, poor women and other vulnerable groups, the monitoring system did not consistently disaggregate data by age, poverty status or disability. Adding these dimensions to routine data collection would have allowed more precise analysis of who benefits most and who may be left behind.

Third, by **strengthening the link between community engagement and service uptake data**. The design did not include mechanisms to track whether community awareness activities led to increased demand for CAC services. A simple referral-tracking system or community-to-facility linkage indicators would have helped connect behaviour change efforts to measurable outcomes.

2.3 Synthesis of four impact cases studies plus five additional Zambia cases

When the five additional evaluations from Zambia are viewed alongside the four Zambian impact case studies, a strikingly consistent picture emerges across Sida's Zambia portfolio: Sida-supported interventions do deliver important results, but the strength and reliability of those results, the ability to measure them, and the quality of the contribution design depend profoundly on the design model and the quality of the evidence and systems that support it. Where the contributions are grounded in well-tested models with strong

institutional foundations - such as SCT, Musika, NLACW and Ipas - their results are more consistent, credible and sustained. Where the contributions rely on ambitious, multi-component designs without a strong evidence base - such as INCREASE, WEE, SAP's diversification components and BGFZ's productive-use ambitions - their performance is weaker, more fragile or more difficult to verify.

The same divide appears in evaluability. The contributions that generated the strongest and most reliable evidence - again SCT, Musika, NLACW and Ipas - are those that invested early in structured monitoring systems, routine data collection and clear, measurable indicators. These systems made it possible to track outcomes and assess change with confidence. In contrast, INCREASE and WEE mirror the weaknesses seen in SAP and BGFZ: fragmented monitoring arrangements, activity-heavy reporting, inconsistent data collection and indicators that do not match the scale or complexity of the programme's ambitions. WIN sits between these two groups, with some solid monitoring practices but uneven documentation and thin outcome-level evidence. In short, the combined evidence shows that impact and evaluability rise and fall together, and both depend on the same underlying factor: the quality of evidence used at the outset.

The additional evaluations reinforce the same design shortcomings identified in the impact case studies. INCREASE, WEE and SAP all suffered from over-ambitious intervention designs that attempted to combine multiple complex domains (market systems, climate resilience, gender transformation) without the evidence, capacity or timeframes required. WIN and BGFZ illustrate how innovation-oriented contributions can outpace their own evidence systems, generating activity-level data but lacking the tools to track behavioural or systemic change. And across several contributions - INCREASE, WEE, SAP, BGFZ - the design did not fully incorporate existing research on value chains, climate risks, gender dynamics, productive use of energy, or the fragility of local institutions. These gaps were compounded by optimism bias, pressure to scale quickly, and insufficient integration of academic and policy evidence.

Taken together, the nine evaluations show that Sida's contributions achieve the strongest and most credible impacts where contribution are grounded in robust evidence, realistic intervention designs, and functioning monitoring systems. Where designs are overly ambitious, insufficiently evidence-based, or weakly monitored, contributions struggle both to deliver and to demonstrate results.

2.4 Access to and use of national statistics in Zambia

A further structural constraint shaping evidence use in Zambia has been a limited access to, and limited systematic use of, relevant national statistical datasets, RALS, the Demographic and Health Survey (DHS), and the Living Conditions Monitoring Survey (LCMS). These datasets represent some of the most robust sources of nationally representative information available in the country, yet they remain under-utilised in programme design, monitoring, and evaluation across Sida's portfolio and were also difficult to access for the ET (despite support from the Swedish Embassy in Lusaka).

In principle, these national surveys could provide a powerful empirical backbone for decision-making. RALS offers granular insights into agricultural productivity, market participation, asset ownership, and livelihood transitions, precisely the domains in which programmes like Musika, SAP, and BGFZ operate. DHS provides high-quality data on maternal and child health, nutrition, fertility, and service utilization, directly relevant to SCT’s nutrition findings and to broader social protection debates, but also to both Musika’s and Diakonia’s interventions. LCMS tracks poverty dynamics, consumption patterns, inequality, and shocks, offering a national counterfactual against which programme-level results could be benchmarked. Yet in practice, these datasets rarely shape Sida’s individual contribution decisions.³³ Several factors contribute to this gap:

First, access barriers remain significant. Although Zambia’s Central Statistical Office (now the Zambia Statistics Agency, ZamStats) has made progress in data dissemination, many datasets are still difficult to obtain in a timely manner. Access often requires formal requests, lengthy correspondence, or payment of fees. For external partners, including donors, this creates friction that discourages routine use. Even when datasets are technically accessible, metadata, documentation, and codebooks are not always up to date, making analysis more cumbersome.

Second, capacity and incentives within implementing partners limit the use of national datasets. Many NGOs and programme teams rely primarily on their own monitoring systems, which are more immediately available and tailored to their reporting requirements. National datasets, by contrast, require statistical expertise, time, and analytical bandwidth, resources that are often scarce. As a result, programme narratives evolve largely independently of national trends, even when those trends could validate, nuance, or challenge programme assumptions.

Third, Sida’s internal processes do not systematically require triangulation with national data. For example, SCT’s negative breastfeeding and vitamin A findings could have been contextualised using DHS data on feeding practices and supplementation trends. It is noted that the SCT intervention, as a national programme, did include LCMS and DHS data as indicators for measuring people living in extreme poverty, poverty gap ratio and stunted children under five years. There are however no indications that these surveys are used for analytical purposes besides measuring impact up against these indicators.

Musika’s productivity and income effects could have been compared more systematically with RALS data to assess whether observed changes exceeded national trends. While Musika did initially intent to use the RALS and did have internal statistical staff capacity to do so, it became a challenge that the RALS data does not directly identify treatment households and do not follow and interview the same households, thus limiting the comparability over time. SAP’s livelihood gains could have been benchmarked against LCMS poverty dynamics but it would require that Diakonia had staff capacity to fully explore statistical evidence which was not the case. This meant that triangulation with

³³ As mentioned in the introduction, This report focuses on “what works” type of evidence and decisions, and not on relevance and portfolio decisions etc. (which is more linked to the MDPA).

national survey was rarely undertaken, meaning that programme-level evidence remained largely siloed rather than embedded in a broader empirical landscape.

Fourth, national datasets are not produced frequently enough to align with programme cycles, and this reduces their perceived relevance. RALS is conducted every 3-4 years, DHS every 5-7 years, and LCMS on an irregular schedule. Interviewed former and current Sida programme staff in Zambia argued that these intervals are too long to support adaptive management. While this is partly true, it also reflects a missed opportunity, since even infrequent national surveys can provide essential baselines, counterfactuals, and trend data which could strengthen the credibility of programme claims.

The result of this is a paradox: While Zambia possesses unusually strong national data systems for a low-income country, these systems play only a marginal role in Sida's evidence ecosystem. This contributes to several of the underlying patterns identified earlier: The verification deficit is reinforced when results claims are not cross-checked against national statistics; equity blind spots persist when disaggregated DHS or LCMS data - rich in gender, poverty, and geographic breakdowns - are not used more systematically to interrogate who benefits and who is left out.

3. South Sudan

3.1 Evidence from four impact case studies

Sida's support to the Budi Nexus Pilot Project

The Nexus Pilot is a project implemented in Budi County, located in the Eastern Equatoria State in southeastern South Sudan. Some activities were also carried out at the border area between Budi and Ikotos Counties. The Nexus Pilot was implemented by the local faith-based organisation, Community Development Support Services (hereafter CDSS). The project aimed to integrate displaced persons into the host communities, enhance gender equality, include women in peacebuilding and economic activities, and protect them against high levels of violence.

Overall impact assessment (from impact case study):³⁴ The Budi Nexus Pilot was a pilot project, with expectations of evidence generation (for possible scaling). It produced mixed and low-confidence results, with some reported positive changes but limited verifiable evidence. Beneficiary interviews and Community Development Support Services (CDSS) reports suggested perceived improvements in security, women's participation, peace dialogue outcomes, and livelihood gains through VSLAs, but these claims could not be independently confirmed, and only three of five VSLAs functioned effectively. Assertions of reduced cattle raiding, improved gender equality and decrease in the prevalence of GBV lacked triangulation, and severe drought from 2021 likely undermined or obscured project effects. Overall, while the pilot demonstrated potential - especially through CDSS's community-anchored approach - the absence of baseline data, weak monitoring systems, and reliance on self-reported qualitative evidence mean that the scale, attribution and durability of results could not be confidently assessed.

How well has the contribution prepared for generating evidence (for ex-post/impact assessment)?: The Budi Nexus Pilot was very poorly prepared for generating evidence for an ex-post or impact assessment. Among all the impact case studies, this was the least evaluability ready contribution. The project lacked the basic foundations required to document results, test assumptions, and generate learning from a pilot. One fundamental gap was the absence of baseline data. No quantitative baseline was collected for either the Resilience Pilot nor the Nexus Pilot,³⁵ making it impossible to measure change over time or attribute outcomes to the interventions. Monitoring systems were minimal, and reporting was largely descriptive and anecdotal. An end of project narrative report was produced by the local partner organisation, CDSS, but the report was not supported by systematic data collection, outcome indicators, or verification mechanisms. The main evidence sources were self-reporting and some limited field observations, which were of very limited confidence. Thus, many claimed results (e.g., reduced GBV, improved security, reduced cattle raiding) could not be triangulated with

³⁴ Sida/NCG (2026), Impact study of the Budi Nexus Pilot, South Sudan A case study as part of the Central Evaluation of Sida's work with Poverty.

³⁵ Sida/NCG (2026), Impact study of the Budi Nexus Pilot, South Sudan A case study as part of the Central Evaluation of Sida's work with Poverty.

independent data but were largely accepted (also by Sida) as face-value. The evaluation team's own cross-checking of evidence with Armed Conflict Location & Event Data (ACLED), however, did not show full consistence with the reporting.

The project altogether suffered from weak documentation and learning systems, despite being a pilot explicitly intended to generate learning on nexus programming. No structured learning products, monitoring tools, or data protocols were developed. The lack of support and capacity building to the implementing partner and security constraints further limited data collection, meanwhile the core problem was design and lack of resources. The local organisation could not recruit a monitoring and evaluation professional, and the attention from Church of Sweden was not on supporting the capacity of the local organisation to produce rigorous reports. Meanwhile did the recurring conflicts in the area of operation affect access to locations and thereby potentially achievement of results.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?: The Budi Nexus Pilot was not sufficiently grounded in available evidence on conflict, gender, resilience, or nexus programming. While the local implementing partner made strong efforts to implement the pilot project based on its extensive and in-depth local knowledge, the Church of Sweden contribution design was over-ambitious with weak causal logic, risks, and constraints faced by the local implementing partner to generate impact and learning. This represented a missed opportunity for Sida and the wider nexus community. The design of the Budi Nexus Pilot shows substantial gaps in the use and quality of evidence:

First, the design did not draw sufficiently on existing **conflict evidence**, even though Budi County has been extensively studied, documenting entrenched cycles of revenge killings, cattle raiding, and inter-ethnic tensions.³⁶ Other studies similarly show that raiding had become militarised and resistant to short-term dialogue.³⁷ These insights were not reflected in the project's peacebuilding strategy, which relied heavily on community dialogues and committees.

Second, the project's assumptions about **gender and GBV** overlooked well-established evidence showing that GBV prevention in conflict settings requires long-term, multi-component interventions, not short sensitisation sessions.³⁸ It is noted, however, that the local project partner did implement other Church of Sweden and DanChurchAid projects, which extended the collective timeframe for addressing the root causes of GBV and reduce its prevalence.

Third, the **livelihoods and resilience** components did not draw sufficiently on the extensive existing evidence on livelihood fragility, climate shocks, and resilience in South

³⁶ Conflict Sensitivity Resource Facility (2020), Budi County Conflict Analysis.

³⁷ Wild, Hannah; Jok, Jok Madut; & Patel, Rachel (2018). *The Militarization of Cattle Raiding in South Sudan*. London School of Economics.

³⁸ See e.g. Heise, Lori (2011), *What Works to Prevent Violence Against Women?*; UN Women (2015), *Preventing Conflict, Transforming Justice, Securing the Peace*.

Sudan. Both FAO³⁹ and UNDP⁴⁰ have repeatedly emphasised that livelihood gains in Eastern Equatoria are highly vulnerable to drought, displacement, and insecurity, and studies on famine and market collapse⁴¹ and on resilience in protracted crises,⁴² have highlighted the need for sustained, multi-year support and water-secure interventions. These considerations were not built into the pilot design (e.g. of VSLAs, kitchen gardens, and infrastructure components).

The project did attempt to reflect global guidance on humanitarian-development-peace (HDP) nexus programming - emphasising simultaneous interventions of humanitarian action with longer term development programming and peacebuilding - but resources were inadequate, and the emphasis was on the organisational aspects rather than on field implementation dynamics, where the local implementing partner was left to try to make ends meet and address a myriad of challenges with limited support by Church of Sweden. The lack of robust monitoring was a major gap, which led to lack of attention to the pilot project - despite both Church of Sweden and Sida's own policy commitments on the NDP nexus.⁴³

Sida's support to the Health Pooled Fund Programme (HPF)

The Health Pooled Fund (HPF) was a flagship, multi-donor program launched in 2012 to support delivery of essential health services in South Sudan. From the outset, the HPF's long-term goal was to build a government-led health system to provide critical, life-saving services, particularly on maternal and infant health. To this end, the HPF contracted implementing partners to deliver a sub-set of services under the Ministry of Health's 'Basic Package of Health and Nutrition Services' at government health facilities, while working to build the Government's capacities and ability to take over health service delivery in the longer-term.

Overall impact assessment (from impact case study):⁴⁴ The impact case study shows that HPF overall has delivered a substantial, life-saving impact in one of the world's most fragile health systems, significantly expanding access to essential services - especially maternal, newborn and child health - across up to 80% of facilities in supported states. The programme increased service coverage, boosted patient attendance, increased skilled birth attendance, strengthened immunisation, and enabled community-level treatment of millions of childhood illnesses through the Boma Health Initiative. Although impact-level indicators on maternal and child mortality rates could not be measured rigorously due to absence of recent national surveys, existing evidence indicates that the HPF likely contributed to a reduction in preventable maternal and child mortality. Despite these

³⁹ FAO's South Sudan Resilience Strategy (2015-2019).

⁴⁰ UNDP's Community Resilience and Local Governance work (2016-2018).

⁴¹ See e.g. Maxwell, Daniel et al. (2016), Facing Famine: South Sudan's Food Security Crisis. Feinstein International Center.

⁴² Levine, Simon et al. (2017). The Livelihoods-Resilience Nexus in Protracted Crises. Overseas Development Institute (ODI).

⁴³ OECD (2020), DAC Recommendation on the Humanitarian-Development-Peace Nexus; Stockholm International Peace Research Institute (SIPRI) (2019), Connecting the Dots on the Triple Nexus.

⁴⁴ Sida/NCG (2025), Impact study Health Pooled Fund, South Sudan. A case study as part of the Central Evaluation of Sida's work with Poverty.

important gains, the programme fell short of its long-term goal of building a government-led health system, as weak state capacity, chronic underfunding, and repeated national and sub-national humanitarian crisis and conflict. The impact case study also demonstrated that the HPF's service delivery achievements remain vulnerable to reversal without continued donor support, highlighting the tension between meeting urgent needs and fostering systemic resilience in a conflict-affected context.

How well has the contribution prepared for generating evidence (for ex-post/impact assessment)?: The HPF was relatively well-prepared for an ex-post or impact assessment compared to most interventions in South Sudan, largely because it operated with one of the strongest monitoring systems in the country. The programme generated extensive routine data through the national Health Management Information System (HMIS),⁴⁵ and quarterly and annual reporting by Crown Agents provided consistent, long-term datasets across 12 years of implementation. This allowed credible tracking of service coverage, utilisation and health outcomes, especially related to maternal and child health services. In DFID's Business Case for the first phase of the programme, there was an intent to include strong impact and outcome level measures, including establishing a baseline in 2011 at the outset of the programme; however, these relied on the Government of South Sudan to conduct a (planned) Household Survey, Demographic Health Survey, and a Maternal Mortality Survey in 2011, which did not happen. When these national-level surveys did not take place, the HPF was unable to establish a clear baseline, and instead had to rely on outdated 2006 household survey estimates, limiting the precision of before-after comparisons. Data gaps in national statistics, uneven data quality across states, and reliance on NGO-reported figures further constrained robustness. Overall, the HPF created a strong evidence-base for assessing contribution to health outcomes, but weaknesses in baseline data and systemic data limitations in South Sudan reduced the ability to conduct a fully rigorous impact assessment. It is important here to note an inherent trade-off in investing more in data collection vs. life-saving health services which was reflected in DFID/FCDO's first business case for the HPF: *Any additional data collection must not be burdensome and detract from the health system strengthening aspects of the HPF (p. 43)*. FCDO also used cost/benefit analysis to assess the HPF use of funds.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?:

The evidence base informing HPF's contribution design was generally strong and appropriate for a highly fragile, data-poor context. DFID/FCDO's Business Cases for each phase were unusually thorough, drawing on comparative evidence, lessons from earlier funding mechanisms (such as the Basic Services Fund (BSF), the Sudan Health Transformation Project (SHTP) and the Multi-Donor Trust Fund (MDTF)) and cost-effectiveness modelling using WHO CHOICE data. Although this modelling relied on East Africa regional estimates due to the absence of South Sudan-specific data, it still

⁴⁵ HPF implemented a Quality Supervision Checklist through Ministry of Health in order to strengthen quality of data.

provided a transparent and structured basis for comparing intervention options. HPF's early emphasis on expanding service availability and standardising a basic package of care reflected what was feasible and evidence-based at the time. Subsequent qualitative studies, including Lawry et al. (2017)⁴⁶ and Kane et al. (2018),⁴⁷ later highlighted the importance of demand-side factors - **fear, dignity concerns, gender norms and low trust** - which showed that availability alone would not ensure uptake. These insights helped shape the gradual strengthening of community-level and behavioural components in later phases.

The evidence base on health-system strengthening was also evolving. Kruk et al. (2018)⁴⁸ argued for broader quality-of-care and patient-experience indicators in fragile settings. HPF's monitoring systems focused mainly on outputs, coverage and facility readiness - appropriate for what was feasible - and contributed to improved HMIS reporting, as noted by Jones et al. (2015).⁴⁹ However, the absence of **systematic quality-of-care and patient-reported measures** limited the programme's ability to assess whether expanded services translated into better user experience or more effective coverage. Jones et al. (2015)⁵⁰ also showed that **national-level system strengthening** faced structural constraints, while sub-national progress was more realistic. HPF recognised these challenges, though early ambitions for national harmonisation were somewhat ahead of what the political-economy context could support. Over time, the programme adapted by focusing more on county-level systems.

The **contracting-out model** was selected partly because alternative approaches had performed poorly in South Sudan. Later evidence, such as Bol et al. (2021),⁵¹ suggested that contracting had not yet produced substantial reductions in under-five mortality, though data limitations make firm conclusions difficult. This does not undermine the rationale for contracting but illustrates the challenges of generating robust causal evidence in South Sudan. Finally, the priority-setting approach described by Widdig et al. (2022)⁵² showed that HPF applied clear criteria - cost-effectiveness, coverage, equity and system stabilisation - and coordinated with other donors to avoid duplication. The main constraint was the underlying data, particularly for equity and system-wide effects, which remained incomplete despite HMIS improvements. Thus, overall HPF's design was grounded in the best evidence available at the time. The limitations - around demand-side behaviour, quality of care, political-economy feasibility and the performance of contracted providers - mainly reflect the broader constraints of working in South Sudan rather than

⁴⁶ Lawry et al. (2017), A mixed methods assessment of barriers to maternal, newborn and child health in Gorgrial West, South Sudan.

⁴⁷ Kane et al. (2018), Too Afraid to Go: Fears of dignity violations as reasons for non-use of maternal health services in South Sudan.

⁴⁸ Kruk, M.E. et al. (2018), High-quality health systems in the Sustainable Development Goals era: time for a revolution?

⁴⁹ Jones et al. (2015), Feasibility of health systems strengthening in South Sudan.

⁵⁰ Jones et al. (2015), Feasibility of health systems strengthening in South Sudan.

⁵¹ Bol et al. (2021), Does contracting out of primary health care services to non-state providers in South Sudan reduce child mortality?

⁵² Widdig, et. al. (2022), The political economy of priority-setting for health in South Sudan: a case study of the Health Pooled Fund.

shortcomings in analytical effort, and later studies have refined rather than contradicted the programme's evidence base.

Sida's support to the UNFPA Country Programme

The UNFPA Country Programme for South Sudan aimed to ensure universal access to sexual and reproductive health and rights, and to reduce maternal mortality in South Sudan. UNFPA adopted a systemic approach, partnering with the Government of South Sudan to strengthen national healthcare systems and deliver integrated sexual and reproductive health information and services. In parallel, it provided direct support to crisis-affected populations, particularly women and youth.

Overall impact assessment (from impact case study):⁵³ The UNFPA programme delivered major service-level gains, giving more than one million women and girls access to SRH and GBV services nationwide. Training and deploying midwives increased skilled birth attendance, contributing to reduced maternal mortality, and UNFPA supported significant improvements in fistula treatment, youth-friendly SRH services, and GBV response. However, systemic and long-term impact remained limited. Weak government capacity, chronic shortages of staff and supplies, and low political commitment meant that national systems could not sustain services without continued donor support. Overall, the programme achieved strong, life-saving results, but sustainability and deeper structural change were constrained by the fragility of the context.

How well has the contribution prepared for generating evidence (for ex-post/impact assessment)?: The UNFPA programme was moderately prepared to generate evidence for ex-post or impact assessment. It benefited from relatively strong routine monitoring systems, drawing on HMIS data, UNFPA's own annual reporting, and a substantial independent Country Programme Evaluation that used mixed methods and collected primary qualitative data. These systems provided consistent output and some outcome-level information across all programme years. However, major structural constraints in South Sudan's data environment (ranked 193/195 globally) meant that key outcome indicators (e.g., maternal mortality, contraceptive prevalence, GBV prevalence) relied on outdated national surveys, estimates, or incomplete facility reporting. UNFPA did not generate dedicated baseline data for several outcome areas, and the Country Programme Evaluation did not include an explicit impact assessment. Overall, the programme created sufficient evidence to assess effectiveness, but insufficient conditions for a rigorous impact assessment, especially where national data gaps were largest.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?:

The programme design relied heavily on routine HMIS data and global SRHR guidance, which provided a plausible and operationally realistic basis for identifying priority needs and shaping interventions. This meant that the design rested on credible pathways of

⁵³ Sida/NCG (2026), Impact study of the UNFPA Country Program for South Sudan. A case study as part of the Central Evaluation of Sida's work with Poverty.

change, particularly around expanding service availability, strengthening facility-based provision and ensuring continuity of essential SRHR services in a fragile setting.

However, a wider body of qualitative and contextual research from South Sudan offered insights that were only partially reflected in the programme's design. Studies such as Lawry et al. (2017)⁵⁴ and Kane et al. (2018)⁵⁵ showed that low uptake of maternal and reproductive health services was driven not only by availability but also by entrenched **social norms, fear, concerns about dignity, and low trust in health facilities**. Lawry et al. documented extremely low demand for contraception - over 90 percent of women and men reported not wanting to use contraceptives - and highlighted distance and lack of skilled birth attendance as major barriers to maternal and newborn health. Kane et al. (2018)⁵⁶ similarly found that fears of mistreatment and violations of dignity deterred women from seeking facility-based care, even when services were technically available. These findings pointed to **behavioural, cultural and experiential determinants of service use** that routine HMIS data could not capture.

While the programme acknowledged some of these barriers, its design remained largely supply-driven, focusing on expanding service availability and strengthening facility readiness. The behavioural and socio-cultural evidence emerging from South Sudan was not fully integrated into the programme's theory of change, nor did it significantly shape the design of demand-side interventions or community-level engagement strategies. As a result, the programme's contribution logic did not fully anticipate the extent to which social norms, trust and perceptions of dignity would influence uptake of SRHR services.

In sum, the programme made good use of the routine and institutional evidence available, but did not fully capitalise on the wider research base - both global and South Sudan-specific - that could have strengthened its understanding of demand-side dynamics, quality of care and feasible approaches to outcome measurement in fragile settings. The evidence used was sufficient to support a plausible and contextually grounded design, but not to fully anticipate the social and cultural factors shaping service uptake.

Sida's support to the UNDP South Sudan - Peace and Community Cohesion Project (PaCC)

The Peace and Community Cohesion Project (PaCC) is a community level conflict mitigation programme implemented in five geographic conflict hotspots/clusters. The programme has adopted a conflict clusters approach meaning that peacebuilding interventions were designed in geographic areas with specific conflict dynamics and spanned across several administrative boundaries (i.e. counties and smaller administrative units). PaCC supported divided communities to implement joint projects of common interests to develop positive relationships. To sustain peaceful coexistence, it also helped

⁵⁴ Lawry et al. (2017), A mixed methods assessment of barriers to maternal, newborn and child health in Gorgrial West, South Sudan.

⁵⁵ Kane et al. (2018), Too Afraid to Go: Fears of dignity violations as reasons for non-use of maternal health services in South Sudan.

⁵⁶ Kane et al. (2018), Too Afraid to Go: Fears of dignity violations as reasons for non-use of maternal health services in South Sudan.

establish institutional mechanisms locally. Furthermore, the programme supported institutional peace infrastructure at the national level.

Overall impact assessment (from impact case study):⁵⁷ PaCC Phase 1 produced localised, perception-based improvements in security, trust, and inter-communal relations, and successfully established community peace committees and dialogue platforms across the five conflict clusters. These activities helped reduce tensions in some areas, but the evidence is too limited to attribute changes directly to the project, as findings rely almost entirely on beneficiary perceptions with no independent conflict data or counterfactuals. Gains proved fragile and easily reversed due to recurring violence, and the project made little systemic progress, as weak state institutions and low political will prevented local achievements from being scaled or sustained. Overall, PaCC delivered useful short-term community-level outcomes but limited durable or structural impact in South Sudan's volatile context, which could not in any case be expected given the depth of the political and security crisis in South Sudan.

How well has the contribution prepared for generating evidence (for ex-post/impact assessment)?: PaCC Phase 1 was prepared for generating evidence suitable for an ex-post or impact assessment. The project staff systematically collected output level data in all five locations on a recurrent basis and published these on the project website. However, the project relied almost entirely on a single Summative Evaluation as its only independent data source. The evaluation depended heavily on the output level data and beneficiary perceptions, the Summative Evaluation team had limited field access, and independent research in the project locations was limited due to the recurrent conflicts. A baseline study did provide a basis for the Summative Evaluation and combined with the output data from the monitoring system the evaluation report presents findings, which however could not be triangulated due to lack of other sources.

PaCC was designed to operate in “conflict hotspots”, and all five areas of operation were volatile throughout implementation. The general perception, including by Sida and UNDP, was that PaCC presents a viable approach to engage meaningfully in areas where limited impact and sustainability can be expected. The context and the challenges of generating rigorous data for an impact evaluation is simply one of several constraints faced by the project. This raises the question, if more resources should have been allocated up front to include mechanisms to compensate for this (e.g., remote sensing, third-party monitoring, or longitudinal surveys) and whether resources from donors should be allocated to fund such activities. Such trade-offs are general for projects in conflict affected areas, where resources are scarce, access is limited and populations are affected by a multitude of deep-rooted problems. In the case of PaCC, while the evaluation captured useful qualitative insights, the project did not establish the data systems needed for rigorous impact assessment, leaving attribution weak and long-term effects largely unmeasurable.

⁵⁷ Sida/NCG (2026), Impact study of the UNDP South Sudan, Peace and Community Cohesion project. A case study as part of the Central Evaluation of Sida's work with Poverty.

Have there been shortcomings/gaps in the assessment/use/quality of evidence (for contribution design)?: UNDP identified a considerable number of conflict drivers in each project location during the design of PaCC Phase 1, and the project drew both on internal analysis and made some use of broader contextual, academic, or security research that was already available on South Sudan’s conflict dynamics. However, it was a shortcoming that external and independent data - if and when existing - were not sufficient location specific. Moreover, conflict dynamics changed frequently in the different locations and access to data collection was not possible (or allowed for security reasons). Moreover, the project locations selected by PaCC traversed different administrative boundaries and thereby units of analysis for example for the Conflict Sensitivity Resource Facility (CSRF), which has been setup by the donors to monitor and analyse local conflict patterns usually at county level⁵⁸. All these complexities led the impact case study to note that “*there are no other independent data sources available for the conflict clusters*”. The Summative Evaluation therefore became the primary evidence base for both UNDP and Sida to communicate the “unfiltered truth” despite its methodological constraints. This raises the question whether the contribution after the design stage sufficiently incorporated and triangulated with: i) Independent conflict-trend data (e.g., ACLED, Small Arms Survey); and ii) Studies on the limits of national level peacebuilding in highly militarised contexts. With regards to the latter, PaCC was redesigned in Phase II to focus more on community-level peacebuilding structures and leave out the national level, which failed to deliver results. PaCC has rested on rather optimistic assumptions about the ability of community-level interventions to generate sustained peace within the conflict hotspots in the absence of broader political settlement or functioning state structures.

3.2 Synthesis of the four South Sudan impact case studies

In South Sudan, Sida has often made large, strategic funding decisions in environments where rigorous measurement of evidence would come at a high cost – if at all possible. Donors, including Sida, when they operate in conflict affected settings calculate with trade-offs between their spending on activities including saving lives versus conducting rigorous measurements with advanced and expensive methods, noting that they are well aware of the potential risk of limited impact and sustainability. As a result, even contributions that delivered important service-level gains struggled to demonstrate impact, and several interventions were designed on assumptions that existing research could have challenged.

Across the four South Sudan cases, Sida-supported programmes did achieve important outcomes: HPP and UNFPA contributed to the expansion of life-saving health and SRHR services nationwide; PaCC facilitated local peace dialogues and strengthened community structures; and the Budi Nexus Pilot generated some perceived improvements in security and livelihoods. Yet these achievements were uneven and difficult to verify, and the unevenness follows the same logic across all cases; the tangible, people-oriented outputs and outcomes have in general been more easy to trace (service delivery for example), while

⁵⁸ <https://csrf-southsudan.org/>

the state capacity and institutional or societal change processes (peacebuilding or capacity development of government institutions) have been more difficult to capture.

HPF and UNFPA were the most evaluability-ready, largely because they could draw on the national HMIS system and long-standing reporting structures. Even so, both lacked baseline data, relied on outdated national surveys, and operated in one of the weakest data environments globally. By contrast, the Budi Nexus Pilot was not evaluability-ready. The Budi Nexus Pilot lacked baseline, outcome indicators, independent data sources, and verification mechanisms. Monitoring was narrative and self-reported, and evaluators had no counterfactuals or conflict-trend data to assess change. In the case of PaCC, a baseline study existed (the endline study of an earlier UNDP project) which to some degree covered the PaCC project locations. The systematically collected output data was used by the Phase I Summative Evaluation, both by UNDP and Sida without communicating the limitation of the data and trying to triangulate and verify this with other sources, if and how these would exist. In the case of PaCC this limited the degree to which it could be substantiated whether the PaCC results were attributable to the support.

Taken together, the four South Sudan impact studies showed that impact, evaluability and evidence use are tightly interconnected, just as in Zambia. But in South Sudan the relationship runs in the opposite direction. The country's deep rooted conflict situation (myriads of local and national level conflicts), lack of good data availability, and limited access to project locations have constrained designs, even with adaptive project implementation (as was the case of PaCC), and the contributions have struggled both to deliver and to demonstrate impact. Even the strongest interventions produced service-level gains that were difficult to attribute or sustain, while the weakest lacked the foundations needed for learning. The learning points anticipated in the Budi Nexus Pilot is a real shortcoming, while the lack of impact of donor interventions in conflict affected settings is a real threat in general for countries like South Sudan as international development cooperation programmes thin out when they cannot document results and sustained impact.

4. Synthesis of evidence from other Sida evaluations

Across the diverse set of other Sida evaluations reviewed (18 in total) - spanning market systems development, women's economic empowerment, financial inclusion, agriculture, climate resilience, humanitarian health, and SRHR - some consistent lines of experience emerge.

4.1 Evaluability of contributions

The review of other Sida evaluation reports pointed to a persistent weakness of evaluability at the contribution design stage. Many contributions lacked clear ToCs, measurable indicators, baseline data, and plans for assessing systemic change. This made it difficult not only to evaluate impact later, but also to use evidence for real-time decision-making. In the Women's Economic Empowerment through Strengthening Market Systems (WEESMS) Project in Bangladesh, the ToC was gender-aware but not gender-sensitive, and it was not linked to market constraints.⁵⁹ The GROW project in Liberia included no plan for systemic-change assessment despite MSD ambitions. Marie Stopes International (MSI) Afghanistan lacked a ToC altogether, and its results framework was too narrow to capture the programme's broader ambitions.⁶⁰ Agricultural and climate-resilience programmes faced similar issues: Zimbabwe's Enhanced Agricultural Productivity and Resilience to Climate Change through Solar Powered Irrigation Project (REAP) had no baseline for water-table trends or market access, despite these being central to its intervention logic.⁶¹ The Organic Trade in East Africa (OTEA) programme lacked poverty indicators, even though poverty reduction was a stated goal.⁶² Mali's loan-guarantee facility had no evaluability plan for behavioural change among banks or borrowers. As a consequence of these shortcomings, learning has often been retrospective and incomplete, and where the ability to adapt based on evidence has been constrained by the absence of evaluability foundations.⁶³

Monitoring systems were often not designed to generate the evidence needed for adaptive design. Monitoring systems were typically output-heavy, meaning they tracked activities and immediate results but not the deeper behavioural or systemic changes that the interventions aimed to influence. This manifested as: Lack of baseline data or inconsistent baselines; indicators that were not specific, not measurable, or not aligned with the ToC; limited tracking of behavioural change, institutional uptake, or systemic effects; weak or

⁵⁹ Sida/NIRAS Sweden AB (2021), End-term Evaluation of Women's Economic Empowerment through Strengthening Market Systems (WEESMS) Project, Bangladesh. Sida Decentralised Evaluation 2021:36.

⁶⁰ Sida/FCG Sweden & Tana Copenhagen (2023), Evaluation of MSI Reproductive Choices Afghanistan, with a focus on the project funded by Sweden. Sida Decentralised Evaluation 2023:12; NIRAS Sweden AB (2021), Decentralised Evaluation of GROW Liberia Phase 1 - a Swedish-funded Market System Development Project in agricultural value chains in Liberia 2013-2020. Sida Decentralised Evaluation 2021:1.

⁶¹ Sida/FCG Sweden (2022), Evaluation of Enhanced Agricultural Productivity and Resilience to Climate Change through Solar Powered Irrigation in Zimbabwe. Sida Decentralised Evaluation 2022:10.

⁶² Sida/NIRAS Sweden AB (2019), Evaluation of Sida's Support to the Project "Organic Trade and Value Chain Development in East Africa" 2014-2019. Sida Decentralised Evaluation 2019:23.

⁶³ Sida/Tana Copenhagen (2019), Mid-Term Evaluation of the Sida and USAID Loan Portfolio Co-Guarantee and the Mali Finance for Food Security and Women Entrepreneurs (FFSWE) Programme. Sida Decentralised Evaluation 2019:10.

absent monitoring of unintended consequences, especially gender backlash, market distortions, or exclusion effects; heavy reliance on self-reported partner data, often without verification. This meant that even when interventions were well-intentioned, the contributions lacked the evidence needed to adjust course, validate assumptions, or refine design choices.

Examples of this include: WEESMS lacked systematic tracking of systemic change, despite its MSD orientation. MSI Afghanistan's results framework had limited indicators making it difficult to assess contributions to higher-level outcomes. REAP Zimbabwe relied heavily on qualitative evidence and UN Women Afghanistan had fragmented monitoring, with limited evidence of systematic use of results for adaptive management.⁶⁴ A related weakness was a misalignment between intervention design and measurement needs. WEESMS did not link market system constraints to intervention design, complicating assessments of systemic change. Loan guarantee facilities had loan-level data but limited evidence on borrower outcomes such as income, employment, or business performance. These gaps suggest that evidence needs were not well-embedded in the design phase, reducing the ability to assess results later.⁶⁵

In addition to this, across the reviewed evaluations, a consistent concern is a heavy reliance on partner-generated and reported data without independent verification. This is found to create a verification deficit, weakening the quality of the evidence base and undermining the credibility of reported results, thereby limiting the ability to make informed decisions. Many contributions depended heavily on partners' own data systems, which varied widely in quality and rigour. This created a structural vulnerability: Weak partner systems translated directly into weak evidence. Examples include: SME databases in WEESMS that were incomplete and required evaluators to reconstruct sampling logic; Banks in guarantee programmes that did not systematically track borrower performance, job creation, or financial behaviour; SRHR partners who collected service statistics but not behavioural or rights-based indicators; OTEA with fragmented data systems and inconsistent reporting across countries and without verifying certification or market-linkage results; in Ethiopia's Enat Bank facility, borrower outcomes were not independently verified;⁶⁶ MSI Afghanistan relied heavily on partner reporting, with limited external validation; Zimbabwe's REAP programme did not independently measure productivity or market-access claims; Mali's Finance for Food Security and Women Entrepreneurs (FFSWE) programme did not validate loan performance or borrower outcomes.

⁶⁴ Sida/FCG Sweden & Tana Copenhagen (2023), Evaluation of MSI Reproductive Choices Afghanistan, with a focus on the project funded by Sweden. Sida Decentralised Evaluation 2023:12; Sida/Watkins F (2018), Evaluation of UN Women Country Programme in Afghanistan (2018). Sida Decentralised Evaluation 2019:14

⁶⁵ Sida/NIRAS Sweden AB (2021), End-term Evaluation of Women's Economic Empowerment through Strengthening Market Systems (WEESMS) Project, Bangladesh. Sida Decentralised Evaluation 2021:36; Sida/Tana Copenhagen (2019), Mid-Term Evaluation of the Sida and USAID Loan Portfolio Co-Guarantee and the Mali Finance for Food Security and Women Entrepreneurs (FFSWE) Programme. Sida Decentralised Evaluation 2019:10.

⁶⁶ Sida/FCG Sweden (2023), Evaluation of Guarantee Instrument and Technical Assistance Offered to Enat Bank to Increase Access to Capital for Women-Owned/Managed Small and Medium Enterprises in Ethiopia. Sida Decentralised Evaluation 2023:12.

Taken together, the contributions were not consistently prepared to generate the type of evidence needed for robust design, adaptive management, and credible results assessment. The gap between ambition and evidence systems was a persistent weakness across the reviewed evaluations. The weaker cases lacked robust M&E frameworks, had unclear or overly broad objectives, or relied heavily on qualitative or anecdotal evidence. The overarching pattern from these evaluations is that while the evaluators were mainly able to produce credible findings, they often had to compensate for critical data gaps rather than building on strong programme systems.

4.2 Existence and use of evidence for contribution design

Despite differences in sector, geography and modality, the synthesis reveals a shared structural weakness: Contributions were rarely designed on the basis of sufficiently strong, targeted, and usable evidence. This has affected the quality of intervention choices, the realism of expected results, and the ability to adapt during implementation. Most interventions began with some form of analysis - market system assessments, gender analyses, climate-risk assessments, conflict scans, sector studies - yet these insights rarely shaped the design logic or strategic choices.

Market systems programmes (WEESMS, GROW) were found to include analyses that were technically sound but not translated into prioritised, sequenced interventions, and often lacked gender-sensitive and political-economy depth. In WEESMS (Bangladesh), the programme narrative centred on the idea that strengthening market systems would automatically create opportunities for women. Yet the evidence generated by the programme's own market and gender analyses showed that the identified constraints were structural, deeply gendered, and unlikely to shift through light-touch facilitation alone. Despite this, the programme continued to operate on the assumption that systemic change would emerge organically from partner engagement. Similarly, GROW Liberia adopted a classic MSD narrative assuming that targeted value chains would transform once key actors were incentivised to innovate. The programme invested heavily in sector analyses, yet the evidence was not used to adjust partner selection or value-chain focus early enough, despite early monitoring data showing limited partner uptake, weak crowding-in, and minimal signs of systemic change. The programme's ambition for systemic change was not matched by a willingness to pivot based on emerging data and rather than revisiting the underlying assumptions, the programme largely maintained its original narrative for several years before adjusting.⁶⁷

⁶⁷ Sida/NIRAS Sweden AB (2021), End-term Evaluation of Women's Economic Empowerment through Strengthening Market Systems (WEESMS) Project, Bangladesh. Sida Decentralised Evaluation 2021:36; NIRAS Sweden AB (2021), Decentralised Evaluation of GROW Liberia Phase 1 - a Swedish-funded Market System Development Project in agricultural value chains in Liberia 2013-2020. Sida Decentralised Evaluation 2021:1.

Evaluations of financial-sector programmes showed some of the same dynamics. Loan-guarantee facilities such as the Enat Bank LGF (Ethiopia) and the Mali FFSWE guarantee lacked robust market diagnostics on borrower segments, collateral constraints, and financial behaviour, leading to unrealistic expectations of uptake. Instead, these designs built on the assumption that banks would expand lending to underserved groups - women entrepreneurs, smallholders - if risk-sharing instruments were provided. Evidence from financial-sector diagnostics and global experience show, however, that guarantees alone rarely shifted lending behaviour unless accompanied by changes in incentive structures, product design, and institutional capacity. The evaluations of these contributions found that while these contextual realities were known they were not integrated into contribution designs, because the narrative of “unlocking finance through guarantees” remained dominant. In addition, although the Enat Bank loan-guarantee facility in Ethiopia had access to portfolio data and gender analysis, the design did not incorporate evidence on collateral constraints, informal-sector barriers, or the need for technical assistance to enable uptake.⁶⁸

The Financial Sector Deepening (FSD) Mozambique - despite producing high-quality political-economy and market analyses - these were not consistently translated into strategic prioritisation or regulatory engagement and the contribution also struggled with a convincing narrative. The programme’s strategic direction continued to emphasise innovation and market development even when evidence suggested that regulatory bottlenecks and political constraints were the primary barriers to financial inclusion.⁶⁹

The evaluation of energy and climate programmes show similar patterns: Solar-irrigation and productive-use-of-energy programmes (e.g., REAP Zimbabwe) were often designed around the global narrative that access to renewable energy would catalyse income generation. However, evaluations found that the evidence base for such impacts was mixed, and in some cases weak. The programme underestimated water-table decline, overestimated farmers’ ability to commercialise production, and assumed that energy access would translate into market access - despite evidence to the contrary.⁷⁰

Agriculture and certification programmes also reveal optimism bias: The OTEA programme was built on the narrative that organic certification would increase incomes and market access for smallholders. Yet decades of sector evidence - much of it available at design stage - showed that certification benefits were uneven, costly to maintain, and often captured by exporters rather than farmers. The evaluation conducted found that these lessons were not integrated into the programme’s intervention logic, which

⁶⁸ Sida/Tana Copenhagen (2019), Mid-Term Evaluation of the Sida and USAID Loan Portfolio Co-Guarantee and the Mali Finance for Food Security and Women Entrepreneurs (FFSWE) Programme. Sida Decentralised Evaluation 2019:10; Sida/FCG Sweden (2023), Evaluation of Guarantee Instrument and Technical Assistance Offered to Enat Bank to Increase Access to Capital for Women-Owned/Managed Small and Medium Enterprises in Ethiopia. Sida Decentralised Evaluation 2023:12.

⁶⁹ Nordic Consulting Group (2021), Final Evaluation of Financial Sector Deepening Mozambique (FSDMoc) 2016-2020. Sida Decentralised Evaluation 2023:23.

⁷⁰ Sida/FCG Sweden (2022), Evaluation of Enhanced Agricultural Productivity and Resilience to Climate Change through Solar Powered Irrigation in Zimbabwe. Sida Decentralised Evaluation 2022:10.

continued to assume that certification would drive poverty reduction.⁷¹ Likewise, agricultural interventions (REAP Zimbabwe) were found to lack environmental and hydrological evidence, resulting in design choices that did not fully anticipate water scarcity or climate variability. In relation to Zimbabwe's solar-irrigation programme, climate-risk evidence was available, but the design underestimated water-table decline and overestimated market access.⁷²

In humanitarian and SRHR contexts - where evidence is often harder to generate - similar patterns emerged. Governance and gender programmes (UN Women Afghanistan) lacked political economy analysis, leading to over-ambitious reliance on weak government institutions.⁷³ SRHR programmes (MSI Afghanistan) had strong service delivery data but weak analysis of social norms, vulnerability patterns, and barriers to access, limiting the ability to design for gender mainstreaming, rights-based approaches, and targeting of the most vulnerable groups. In MSI Afghanistan, the programme narrative emphasised that awareness-raising and service expansion would reduce GBV and improve reproductive health outcomes. However, global evidence shows that awareness-raising alone rarely shifts entrenched social norms, and that service uptake depends on a complex mix of security, mobility, household dynamics, and provider trust. The evaluation conducted found that these contextual factors were known but not fully incorporated into programme logic.⁷⁴

Across all these cases, the underlying issue was not the absence of evidence, but the absence of a systematic approach for identifying and making use of evidence to inform decisions. In many cases, critical evidence existed but was not used as the primary driver of design choices.

Poverty/equity analysis are often weak

In agricultural and market-systems programmes, such as OTEA and REAP, gender analysis was found to be limited or superficial. While designs acknowledged that women often face structural barriers in accessing land, inputs, and markets, the interventions did not systematically address these constraints. Monitoring frameworks often lacked gender-disaggregated indicators, making it difficult to assess whether women benefited proportionately or whether interventions inadvertently reinforced existing inequalities.⁷⁵

⁷¹ Sida/NIRAS Sweden AB (2019), Evaluation of Sida's Support to the Project "Organic Trade and Value Chain Development in East Africa" 2014-2019. Sida Decentralised Evaluation 2019:23.

⁷² Sida/FCG Sweden (2022), Evaluation of Enhanced Agricultural Productivity and Resilience to Climate Change through Solar Powered Irrigation in Zimbabwe. Sida Decentralised Evaluation 2022:10.

⁷³ Sida/Watkins F (2018), Evaluation of UN Women Country Programme in Afghanistan (2018). Sida Decentralised Evaluation 2019:14.

⁷⁴ Sida/FCG Sweden & Tana Copenhagen (2023), Evaluation of MSI Reproductive Choices Afghanistan, with a focus on the project funded by Sweden. Sida Decentralised Evaluation 2023:12.

⁷⁵ Sida/NIRAS Sweden AB (2019), Evaluation of Sida's Support to the Project "Organic Trade and Value Chain Development in East Africa" 2014-2019. Sida Decentralised Evaluation 2019:23; Sida/FCG Sweden (2022), Evaluation of Enhanced Agricultural Productivity and Resilience to Climate Change through Solar Powered Irrigation in Zimbabwe. Sida Decentralised Evaluation 2022:10.

Financial-sector programmes also exhibited equity “blind spots”. While loan-guarantee facilities and financial-inclusion initiatives often aimed to support women entrepreneurs, the reviewed evaluations found that the designs rarely addressed the informal-sector barriers women face - such as lack of collateral, limited financial literacy, or discriminatory lending practices. Without targeted strategies or disaggregated monitoring, it was difficult to determine whether guarantees actually expanded access for women or simply subsidised lending to existing clients.

Humanitarian and SRHR programmes showed similar gaps. The evaluation of MSI Afghanistan and other health-service initiatives noted that while vulnerable groups were mentioned in programme objectives, there was limited analysis of how intersecting vulnerabilities - such as disability, displacement, or adolescent status - shaped access to services. As a result, programmes struggled to identify which groups were most vulnerable or to tailor interventions accordingly.⁷⁶

Cost-effectiveness evidence is rarely used

Across the reviewed evaluations, cost-effectiveness analysis remains the exception rather than the norm. In climate-resilient agriculture and energy programmes, the evaluations found that economic viability was often assumed rather than tested. Solar-irrigation initiatives did not fully assess whether farmers could generate sufficient returns to cover maintenance costs, especially in areas facing declining water tables or limited market access. Without such analysis, it was difficult to determine whether solar pumps were a sustainable investment or whether other technologies might offer better value for money. Organic-certification programmes similarly lacked cost-effectiveness assessments. The OTEA evaluation noted that certification processes were expensive and time-consuming, yet the long-term sustainability of certification systems - and their ability to deliver meaningful income gains - was not rigorously analysed.⁷⁷ As a result, it remained unclear whether certification represented a cost-effective pathway to poverty reduction. Loan-guarantee facilities provide another example. The evaluations of the Enat Bank LGF and the Mali FFSWE programme found that guarantees were introduced without comparing their cost-effectiveness to other financial-sector interventions, such as technical assistance, digital-finance solutions, or direct support to microfinance institutions. Without such comparisons, it is difficult to assess whether guarantees were the most efficient mechanism for expanding access to finance.⁷⁸

⁷⁶ Sida/FCG Sweden & Tana Copenhagen (2023), Evaluation of MSI Reproductive Choices Afghanistan, with a focus on the project funded by Sweden. Sida Decentralised Evaluation 2023:12.

⁷⁷ Sida/NIRAS Sweden AB (2019), Evaluation of Sida’s Support to the Project “Organic Trade and Value Chain Development in East Africa” 2014-2019. Sida Decentralised Evaluation 2019:23.

⁷⁸ Sida/Tana Copenhagen (2019), Mid-Term Evaluation of the Sida and USAID Loan Portfolio Co-Guarantee and the Mali Finance for Food Security and Women Entrepreneurs (FFSWE) Programme. Sida Decentralised Evaluation 2019:10; Sida/FCG Sweden (2023), Evaluation of Guarantee Instrument and Technical Assistance Offered to Enat Bank to Increase Access to Capital for Women-Owned/Managed Small and Medium Enterprises in Ethiopia. Sida Decentralised Evaluation 2023:12.

Adaptation happens but not systematically

Among the reviewed evaluations, some stand out for their adaptive strengths. GROW Liberia, for example, used early monitoring data and market feedback to refine its sector strategies, adjust partner engagement, and recalibrate expectations around systemic change.⁷⁹ MSI Afghanistan underwent significant adaptation after 2021, responding to a radically altered operating environment by restructuring service delivery models, renegotiating partnerships, and adjusting its ToC⁸⁰ FSD Mozambique also demonstrated adaptive capacity, using political-economy insights to shift its focus toward regulatory engagement and away from interventions that lacked traction.⁸¹

However, several other contributions did not adapt despite clear evidence of weaknesses. In WEESMS Bangladesh, uptake of interventions remained low for extended periods, yet the programme continued to pursue the same facilitation strategies without revisiting its assumptions about partner incentives or gendered market constraints.⁸² The OTEA programme did not adjust its approach even as certification bodies struggled with capacity and sustainability issues that undermined the programme's core logic.⁸³ REAP in Zimbabwe underestimated climate-driven water scarcity, yet the programme did not recalibrate its expectations around solar-irrigation viability or explore alternative technologies.⁸⁴ In Mali's FFSWE programme, evidence accumulated that commercial banks had limited appetite for smallholder lending, yet the programme did not pivot toward microfinance institutions or other actors better positioned to reach the target group.⁸⁵

⁷⁹ NIRAS Sweden AB (2021), Decentralised Evaluation of GROW Liberia Phase 1 - a Swedish-funded Market System Development Project in agricultural value chains in Liberia 2013-2020. Sida Decentralised Evaluation 2021:1.

⁸⁰ Sida/FCG Sweden & Tana Copenhagen (2023), Evaluation of MSI Reproductive Choices Afghanistan, with a focus on the project funded by Sweden. Sida Decentralised Evaluation 2023:12.

⁸¹ Nordic Consulting Group (2021), Final Evaluation of Financial Sector Deepening Mozambique (FSDMoç) 2016-2020. Sida Decentralised Evaluation 2023:23.

⁸² Sida/NIRAS Sweden AB (2021), End-term Evaluation of Women's Economic Empowerment through Strengthening Market Systems (WEESMS) Project, Bangladesh. Sida Decentralised Evaluation 2021:36.

⁸³ Sida/NIRAS Sweden AB (2019), Evaluation of Sida's Support to the Project "Organic Trade and Value Chain Development in East Africa" 2014-2019. Sida Decentralised Evaluation 2019:23.

⁸⁴ Sida/FCG Sweden (2022), Evaluation of Enhanced Agricultural Productivity and Resilience to Climate Change through Solar Powered Irrigation in Zimbabwe. Sida Decentralised Evaluation 2022:10.

⁸⁵ Sida/Tana Copenhagen (2019), Mid-Term Evaluation of the Sida and USAID Loan Portfolio Co-Guarantee and the Mali Finance for Food Security and Women Entrepreneurs (FFSWE) Programme. Sida Decentralised Evaluation 2019:10.

5. Overall synthesis on impact and evidence use

When the findings from the Zambia and South Sudan impact cases are viewed alongside the reviewed evaluations, some deeper and more consistent patterns emerge.

5.1 Sida's contribution to impact: Cross-country patterns

Across the Zambia and South Sudan impact cases, Sida's contributions clearly generated important benefits for target groups, yet the depth, consistency, and sustainability of these impacts varied remarkably across contexts. In Zambia, programmes tended to operate in comparatively stable institutional and market environments, with stronger administrative systems, more predictable implementation conditions, and a longer history of sectoral engagement. These factors created a foundation in which Sida's support could translate more reliably into measurable improvements. Income gains among smallholders in Musika, improvements in food security and child nutrition through SCT, and the expansion of basic energy access under BGFZ all illustrate how Sida's contributions produced tangible and verifiable results when embedded in systems with at least some degree of stability and data availability.

However, even in Zambia, impacts frequently fell short of the ambitions articulated in programme documents. Many interventions achieved strong effects along one or two pathways but did not deliver the broader systemic or transformative changes envisioned. For example, Musika's support improved market access and incomes for participating farmers, yet the wider systemic change it aimed at - such as durable shifts in market functions, incentives, and service provision - remained partial and uneven. Similarly, BGFZ succeeded in expanding access to off-grid energy but did not generate the expected improvements in productive use, income generation, or educational outcomes. SCT contributed to enhanced nutrition and food security not to reduced poverty due to low transfer size, persistent system weaknesses and a general increase in poverty in Zambia. These cases show that even in relatively stable environments, programme ambitions often exceeded what the underlying evidence base and implementation conditions could support.

In South Sudan, Health and SRHR programmes have contributed to stabilising service delivery and maintaining essential care in an extremely fragile environment. Supported facilities generally showed improved readiness, more consistent staffing and better availability of basic services, and HMIS reporting became more regular over time. These achievements pointed to plausible outcomes, even though the scale of change/impact could not be measured. The main constraint is the weakness of South Sudan's national data systems, which limits population-level measurement of outcomes such as maternal mortality, contraceptive prevalence and GBV prevalence. As a result, programmes have relied heavily on routine facility data, which are useful for tracking outputs but insufficient for rigorous attribution or ex-post impact assessment. Larger programme designs have made good use of the routine and operational evidence available - HMIS data, global SRHR guidance, lessons from earlier pooled funds and cost-effectiveness modelling.

However, they made only partial use of the wider evidence base which showed that low uptake of maternal and reproductive health services is shaped by fear, dignity concerns, gender norms and low trust, not just by availability. These insights were acknowledged but not fully integrated. Similarly, global evidence on quality of care and feasible outcome measurement in fragile settings was not systematically reflected in monitoring systems.

The broader review of Sida evaluations reinforces these cross-country patterns. Across regions and sectors, Sida-supported programmes often generated important benefits: improved market access, expanded financial inclusion, strengthened service delivery, enhanced resilience, and increased participation of women and marginalised groups. Yet these gains were frequently uneven, fragile, or insufficiently linked to higher-level outcomes such as sustained income growth, systemic transformation, institutional strengthening, or long-term resilience. In many cases, programme ambitions exceeded what the evidence base could support. Designs often assumed that complex behavioural or systemic changes would occur without the enabling conditions required to sustain them. Moreover, the lack of robust monitoring systems limited the ability to track whether early gains were translating into deeper or more durable forms of change. Weaknesses in outcome measurement, limited attention to equity effects, and insufficient tracking of systemic dynamics meant that programmes could demonstrate activity-level progress but struggled to show whether this progress was leading to the intended long-term impacts.

Taken together, these patterns highlight a central insight: Sida's contributions can and do generate results, but the strength and sustainability of these results depend heavily on the quality of the underlying evidence, the realism of the design assumptions, and the robustness of the monitoring systems. Where these elements are strong, results tend to be clearer, more consistent, and more resilient. This underscores the importance of grounding programme ambitions in realistic assessments of context, evidence, and system readiness, and of ensuring that monitoring systems are capable of capturing not only outputs but also the deeper forms of change that Sida seeks to influence.

5.2 Preparedness for generating evidence for impact assessment

Across the Zambia and South Sudan impact cases, a striking pattern emerges: even where programmes generated substantial amounts of data, they were rarely designed with ex-post impact assessment in mind with only few exceptions (e.g. SCT). Evaluability was often an afterthought rather than a foundational design principle. This meant that, when impact assessments were eventually commissioned, evaluators frequently had to reconstruct ToCs, rebuild baselines, and rely on secondary data sources rather than programme-generated evidence. The Zambia cases illustrate this clearly. SAP, for instance, lacked the basic architecture required for credible impact assessment and an impact target was not even defined. Baseline data were no longer available in its raw format, incomplete and misaligned with the outcomes the programme later sought to influence - at least what Caritas Zambia tried to influence - and outcome indicators were either missing or insufficiently defined. As a result, the impact evaluation depended heavily on retrospective reconstruction and new data collection rather than on a coherent monitoring system established from the outset.

Musika represents a different but equally instructive challenge. Although it accumulated one of the richest datasets in the portfolio - spanning a decade of household surveys - these data were not originally designed for impact evaluation. Inconsistencies across survey rounds, missing variables, and methodological shifts limited comparability and required the evaluators to invest significant effort in harmonising the data. The programme's evaluability therefore stemmed more from the sheer volume of data than from deliberate design choices. BGFZ, meanwhile, invested heavily in real-time operational monitoring, but its systems were oriented toward tracking sales, installations, and service quality rather than poverty and welfare outcomes. This left major impact dimensions unmeasured, despite the programme's stated ambitions around economic and social transformation. Even the SCT, which benefited from a long tradition of external research and RCTs, continued to rely heavily on secondary studies because internal MIS weaknesses persisted across phases. The programme's own monitoring systems were not sufficiently robust to support a comprehensive impact assessment without external supplementation. This is somehow addressed by the World Bank's MIS but it is only in recent year that it has been scaled nationally and therefore assessment over time is still a challenge. While the ET did request access to the full data set this was not granted and only a snapshot from 2022 was provided.

In South Sudan, the evaluability challenges were even more acute. Many programmes lacked credible baselines altogether, and monitoring systems were fragmented, irregular, and heavily dependent on partner self-reporting. ToCs were often implicit rather than explicit, making it difficult to trace contribution pathways and assess whether observed changes could plausibly be attributed to Sida's support. In such contexts, even when programmes delivered outputs or intermediate outcomes, the absence of systematic outcome tracking meant that higher-level effects remained largely speculative. The volatile operating environment further complicated matters: staff turnover, insecurity, and shifting implementation arrangements disrupted data collection and undermined institutional memory. As a result, impact assessments in South Sudan often relied on qualitative inference or triangulation rather than on structured, programme-generated evidence.

The broader set of Sida evaluations reviewed reinforces these findings. Across sectors and countries, the absence of clear ToCs, measurable indicators, and robust monitoring systems consistently limited the ability to assess impact, learn from implementation, and adapt programme design. Many contributions were designed with ambitious objectives but without the evaluability foundations needed to test whether these ambitions were realistic or achieved. In several cases, monitoring systems focused heavily on activities and outputs, with little attention to behavioural change, systemic effects, or sustainability. Gender and equity indicators were often missing or weakly integrated, making it difficult to assess who benefited and who was left behind. In other words, monitoring systems existed but were not used strategically; data were collected but not analysed or analysed but not fed back into decision-making.

Taken together, these patterns point towards a wider systemic issue: evaluability is rarely embedded as a core design principle within Sida's programming. Instead, it is often treated

as a technical add-on, addressed only when an evaluation is commissioned. This limits not only the ability to assess impact ex-post but also the potential for adaptive management during implementation. Without clear ToCs, aligned indicators, credible baselines, and functioning monitoring systems, programmes cannot generate the evidence needed to understand what works, for whom, and under what conditions.

5.3 Shortcomings in the assessment, use, and quality of evidence for contribution design

Across the Zambia and South Sudan impact cases, the most pervasive finding is that evidence - whether strong, weak, or entirely absent - has not been consistently used to inform contribution design. The Zambia cases illustrate a situation where good evidence has often existed but not been fully utilised. In several sectors, substantial bodies of knowledge were available on conservation agriculture, goat rearing, market systems development, and social protection design. Yet these insights were only partially integrated into contribution choices. Designs often proceeded as if the evidence base were thinner or more ambiguous than it actually was, resulting in interventions that were either over-ambitious or misaligned with contextual realities. The consequence of this was not the absence of impact, but rather that impact became narrower, more fragile, and less transformative than intended.

In South Sudan, the pattern was reversed: contributions were frequently designed as if weak or unverified evidence were stronger than it actually was. The Budi Nexus pilot is a clear example of this dynamic. It was treated as a scalable model despite minimal verification, weak monitoring arrangements, and no credible counterfactual. Similarly, large-scale investments in health and protection relied heavily on partner-reported data without independent validation, and assumptions about behaviour change, service uptake, and system capacity were rarely tested. In such cases, the issue was not simply the absence of evidence, but the over-confidence placed in evidence that was not robust enough to support the scale or ambition of the intervention. This created a situation in which contributions moved forward on the basis of untested assumptions, making it difficult to determine whether observed changes were attributable to Sida's support or to broader humanitarian dynamics.

Across both countries, many contributions also suffered from evidence gaps that could have been filled with modest effort. Baseline surveys, gender-disaggregated indicators, market analyses, and system diagnostics were often missing, even though they would have been feasible to generate early in the programme cycle. The absence of these foundational elements has limited the ability to design interventions grounded in a realistic understanding of context and later made it difficult to assess whether programmes were achieving their intended outcomes. These gaps were not inevitable but reflected a lack of systematic attention to evaluability and evidence needs during the contribution design stage. In the case of Zambia, strong national datasets such as RALS, DHS and LCMS are available, yet they have played only a limited role in Sida's contribution design and evidence use. Access barriers, uneven documentation, and limited analytical capacity among partners mean that programme teams rely mainly on their own monitoring systems rather than

situating results within national trends. Sida's internal processes also do not require systematic triangulation with national statistics, thus opportunities to benchmark programme outcomes have often been missed. Although national surveys are infrequent, they still offer valuable baselines and counterfactuals that could strengthen contribution analysis, but it relies on data accessibility and possibility to break the survey data down on specific locations since many interventions are isolated to specific locations.

A broader pattern emerging from all reviewed evaluations is the dominance of narrative over evidence in contribution choices. Many contributions have been built around compelling development narratives - systemic change, women's economic empowerment, productive use of energy, financial inclusion, climate-smart agriculture, organic certification, or GBV prevention. These narratives often carried strong intuitive appeal and aligned with Sida's strategic priorities. However, they sometimes overshadowed contradictory evidence or gaps in the underlying logic. Designs frequently began with a conceptual storyline and then selected interventions to fit that storyline, rather than grounding the design in empirical evidence about what was likely to work in the specific context. One consequence of this stronger reliance on narratives than on evidence has been that interventions often have tended to persist even when early evidence has shown limited traction. Designs have remained overly ambitious, assuming systemic and behavioural change without sufficient enabling conditions. In this situation, adaptation has often become reactive rather than proactive, occurring only when problems became too visible to ignore.

Verification deficits have further compounded these issues. Across many contributions, partner-reported results were accepted without triangulation, creating blind spots around who was being reached, who was being excluded, and what types of change were realistically achievable. This was particularly evident in South Sudan, but similar patterns appeared in MSD programmes, financial inclusion initiatives, and gender-equality interventions. Without independent verification, programmes risked basing decisions on incomplete or overly optimistic representations of progress.

Equity considerations, despite being central to Sida's institutional commitments, have often been implicit rather than operationalised. While the contributions in general referenced gender equality, human rights, and inclusion in their design documents, they have often lacked the mechanisms needed to assess who benefited and who did not. Monitoring systems frequently failed to capture differences in outcomes across gender, age, disability, poverty status, and geographic location. As a result, the evaluations were often unable to answer basic distributional questions. This limited both accountability and adaptive management, as programmes lacked the evidence needed to adjust strategies in response to emerging inequities.

Cost-effectiveness evidence was also rarely used. Despite substantial financial commitments across the portfolio, very few contributions assessed whether their chosen approach represented the most efficient use of resources or whether alternative models might deliver similar outcomes at lower cost. This has limited Sida's ability to prioritise interventions, allocate resources strategically, and assess value for money. It has also

constrained the ability of evaluations to provide actionable recommendations, as they could not determine whether programme outcomes justified the investment.

Adaptation, while present in some cases, was not systematic. Some contributions demonstrated a strong capacity to adjust course in response to emerging evidence or shifting political conditions. Others continued along their original trajectory even when evidence clearly signalled the need for change. This inconsistency resulted from different factors: rigid programme designs which left limited room for strategic shifts; results frameworks tied to predefined outputs, making adaptation administratively difficult; incentives that rewarded delivery over learning; and monitoring systems that tracked activities rather than mechanisms of change. In the absence of clearly defined decision-making triggers, iterative planning cycles, and learning-oriented monitoring systems, adaptation has often tended to depend on individual programme teams rather than being a structural feature of Sida's approach.

Taken together, these patterns suggest some systemic challenges. While evidence is often present, it is not consistently embedded in contribution design, decision-making, and adaptation practices. Strong evidence is often under-used, weak evidence is over-interpreted, and missing evidence is rarely prioritised for generation. Thus, the issue is not simply the availability of evidence, but rather the incentives, capacities, and governance structures that determine whether and how evidence is being used. At the same time, evaluability is not treated as a design requirement, and monitoring systems are not consistently aligned with contribution pathways. This tends to reduce the overall effectiveness of Sida's investments and limits the potential for learning across contributions. Strengthening evidence governance - through clearer standards for pilots, mandatory evaluability checks, systematic triangulation, and stronger integration of cost-effectiveness and equity analysis - would likely not only improve accountability but also enhance the strategic value of Sida's investments.

6. Reasons for why impact evidence was underused in some cases

Interviews with Sida staff and others pointed towards explanations for why impact evidence was under-utilised in a number of the contribution cases studied.

Delivery pressures and limited time and space for reflection

During interviews, Sida staff explained that the contribution management system has placed considerable pressure on desk officers to keep contributions moving through the system and meet administrative system requirements. Even when evidence has pointed to weaknesses in a contribution's logic or feasibility, the costs - administrative, political, and temporal - of re-visiting the contribution design has often felt too high, leading to situations where evidence was acknowledged but applied only superficially.

Limited incentives for pursuing and using impact evidence

Some Sida staff pointed to some deeper constraints in the incentive structures that shape day-to-day contribution management within Sida. Even when staff recognised the value of robust impact evidence, the organisation's internal drivers were often found to reward other behaviours. Some contribution managers felt that they were primarily assessed on delivery, disbursement, and compliance with administrative requirements, not on the quality of evidence used to inform design or adaptation. Some also found that the system offered few incentives for taking evaluative risks - such as redesigning a contribution mid-stream, challenging a long-standing narrative, or commissioning deeper analytical work - because such actions would slow processes, trigger additional approvals, or complicate reporting. In practice, this meant that even when impact evidence was available, staff perceived limited organisational reward for engaging with it, and in some cases even disincentives for doing so.

Lack of clarity about how to access and make use of impact evidence

Some staff expressed that they simply did not feel confident about how to access and use impact evidence in practice. Even when evaluations, studies, or external datasets are available, the pathway from "having evidence" to "applying evidence" is not always clear. Sida's internal guidance documents tend to emphasise what should be done but offer less concrete support on how to translate evidence into design choices, partner dialogue, or adaptive management. It is however noted that recent briefs have provided clear guidance on partner dialogue which has previously been a challenge.⁸⁶

Sida HQ produces thematic and methods brief which provide examples of good practices, and guidance on what works within a thematic area or a specific method. According to Sida staff these briefs do provide useful evidence on specific topics, but the usability is sometimes limited since they still need to be adapted to the local contexts. Experiences from an assessment of evidence showed that few HQ thematic and methods briefs are

⁸⁶ E.g. Sida (2024), Preventing and responding to gender-based violence, Thematic Support Unit.

referenced in appraisals. Interviews indicated, that consulting national programme officers with long experience was often considered more relevant than considering HQ documents.

Help desks are also established within strategic topics such as gender, environment, human rights, conflict, peace building etc. to support both Sida staff but also partners in considering mainstreaming of strategic perspectives and priority areas. Staff members can request support from help desks to e.g. conduct an analysis of gender or conflict sensitivity for a contribution. Thematic networks are also meant to share newest evidence and best practice examples across geographic areas. Thus, structures are present in Sida to foster knowledge sharing but their usefulness depends very much on the persons in the networks and it is less clear to which extent material shared through these networks is being applied in contribution management.⁸⁷

Nevertheless, contribution managers may hesitate to challenge partners, revise assumptions, or propose design changes because they are unsure of the procedural implications, or the organisational appetite for such shifts. This uncertainty is reinforced by limited structured spaces for peer learning, few practical examples of evidence-informed redesign, or hands-on analytical support. In practice, “not knowing how to go about it” have often become a rational reason to default to continuity, even when stronger evidence points in another direction.

Dependency on strong development narratives

Some staff explained that they feel Sida’s strategic contribution priorities are in some cases framed around strong development narratives - such as systemic change, service delivery, women’s economic empowerment and climate resilience - that provide coherence and direction but can overshadow focus on use of contextualised empirical findings. Once a narrative becomes embedded in a contribution, it tends to shape expectations, indicators, and partner relationships in ways that make it difficult for contradictory evidence to shift the trajectory. As a result of this narrative “lock-in”, contributions may continue along their original path even when emerging evidence suggests that assumptions were unrealistic or that certain pathways are not delivering the expected results.

Similar findings were found in national audit reports conducted of both Swedish development aid (2024) and humanitarian aid (2025). They point to a lack of clarity and transparency on how evidence is applied in decision making and how previous results inform new decisions. Often decision making seems to be based on political priorities or other considerations than evidence of results.⁸⁸

Strong reliance on partner assessment and reporting

Sida staff explained that they tended to rely strongly on partner analysis, reporting and internal monitoring because these sources were readily available and aligned with reporting

⁸⁷ The national auditor (2024), Sweden’s International Development Aid – monitoring, evaluation and reporting, 2024:4.

⁸⁸ The national auditor (2024), Sweden’s International Development Aid – monitoring, evaluation and reporting, 2024:4; The national auditor (2025), Saving lives and alleviating suffering, Sida’s work with humanitarian aid, 2025: 33.

requirements, whereas external evidence requires time, analytical capacity, and institutional continuity that are not always present. Sida's partner assessments are quite solidly founded on evidence and while the process has been criticised for being too cumbersome with humanitarian partners and sometimes overlapping with different departments assessing the same organisation,⁸⁹ staff members largely indicate that selection of partners is founded on solid evidence. This is also confirmed by the Sida assignment on evidence commissioned in July 2025 and again in November 2025. Application of external evidence on results is less evident and systematically conducted and instead partner reporting forms the main evidence base with limited external/independent evidence. Effectiveness and efficiency are rarely assessed with reference to external evidence, and this element is dedicated much less attention in the Sida appraisal system. Interviews further revealed that limited analytical bandwidth in some Sida teams compounded this challenge, making it difficult to identify and interpret large datasets even when they are accessible.

The contribution management system Tool for Results management and Appraisal of Contributions (TRAC) has long been criticised for not being suitable for project follow-up and for being too cumbersome⁹⁰ and several attempts to update it have taken place. In the last quarter of 2025, TRAC was being updated to comply with changes in the Contribution Rules from July 2025.⁹¹ According to interviews, a decision has been taken to replace TRAC fully by 2027.

Evidence constraints in fragile and data-poor contexts

The fact that Sida does not have staff permanently on the ground - or that staff change frequently - in fragile contexts (such as in South Sudan) was found to create other challenges and less systematic follow-up on contributions.⁹² Without a continuous in-country presence, Sida lacks the everyday proximity that would allow staff to observe programme realities and identify early signs of implementation drift. Instead, desk officers must rely on remote communication, periodic missions, and partner-generated documentation, which limits their ability to interrogate data quality or contextualise reported results. The absence of staff on the ground also reduces Sida's ability to build informal knowledge networks, maintain continuity across programme cycles, and detect discrepancies between reported progress and actual change on the ground. In practice, this creates a situation where Sida may have no alternative but to use weak, unverified, or anecdotal evidence simply because it is the only information available within the operating context. Without staff on the ground to validate findings and sense-check assumptions, Sida's ability to critically assess evidence is also further constrained. Still, in very fragile or conflict-affected contexts, even though Sida has an in-country presence, access to project areas may still be limited.

⁸⁹ The national auditor (2025), *Saving lives and alleviating suffering, Sida's work with humanitarian aid*, 2025: 33.

⁹⁰ SIPU (2016), *Evaluation of the effectiveness of Sida's work with Contribution management*.

⁹¹ Sida (2025), *Rule for contribution management*, 1st July 2025.

⁹² Also a finding in the national auditor (2025), *Saving lives and alleviating suffering, Sida's work with humanitarian aid*, 2025: 33.

Lack of evaluability requirements in contribution design

A further explanation provided is that Sida's contribution management do not require use of evidence nor evaluability at the design stage. There is no mentioning of evidence in the Contribution Rule and while it has been encouraged previously it has not been a requirement, so it has largely rested on the individual programme officer to decide to what extent external evidence has been mentioned.⁹³ Contributions can be approved without baselines, counterfactual logic, gender-disaggregated indicators, impact targets or system diagnostics. When these elements are missing at the outset, they are rarely added later, leaving programmes without the data needed to assess impact or adjust course. Even when staff are committed to evidence-driven decision-making, they may simply not have the information required to act on that commitment. Interviews with Sida staff further indicated that measurement of impact will often be pushed from one phase to another in the recognition that it takes time to achieve. In smaller interventions as the Diakonia SAP an impact target was not even defined since it was considered outside the sphere of project influence and by not defining the target the project would not be measured accordingly.

Taken together, these dynamics indicate that the under-use of impact evidence at Sida is not primarily a technical problem. It tends to reflect deeper incentives structures and design practices that make it difficult for evidence - however strong - to systematically shape decisions. Strengthening evidence use therefore may require not only better access to data and information, but also changes to the systems, requirements, and incentives that govern how contributions are designed, approved, monitored, and adapted.

⁹³ Sida (2025), Rule for contribution management, 1st July 2025.

Annex - Selected Sida Evaluations

Sida Decentralised Evaluations, Zambia:

FCG Sweden (2023), **Evaluation of the National Legal Aid Clinic for Women (NLACW) Access to Justice Programme in Zambia**. Sida Decentralised Evaluation 2024:04.

NIRAS Sweden AB (2022), **End of Project Evaluation of the Sida-funded program in Zambia**. Sida Decentralised Evaluation 2022:29.

Topsøe-Jensen B., Nakaponda B., Tempo S., Ericsson F. (2019), **Mid-term Evaluation of Women in Financial Inclusion and Capabilities Project in Zambia**. Sida Decentralised Evaluation 2019:21.

Nordic Consulting Group (2022), **Increasing Resilience in Energy and Agriculture Systems and Entrepreneurship (INCREASE)**. Sida Decentralised Evaluation 2023:20.

Nordic Consulting Group (2022), **Mid-Term Evaluation of the Women Economic Empowerment Project in Zambia**. Sida Decentralised Evaluation 2023:21.

Other Sida Centralised and Decentralised Evaluations:

NIRAS Sweden AB (2021), **Ipas project: Expanding women's and girls' access to comprehensive abortion and contraception care in Nampula and Zambezia provinces in Mozambique**. Sida Decentralised Evaluation 2022:33.

Nordic Consulting Group (2020), **End of Project Evaluation of the One UN-SRHR for girls and young women in Mozambique - Rapariga Biz Program, 2016-2019**. Sida Decentralised Evaluation 2023:15.

FCG Sweden (2021), **End of Project Evaluation of Building Coffee Farmers Alliances in Tanzania (CFAT 2013-2021)**. Sida Decentralised Evaluation 2022:9.

NIRAS Sweden AB (2021), **End-term Evaluation of Women's Economic Empowerment through Strengthening Market Systems (WEESMS) Project, Bangladesh**. Sida Decentralised Evaluation 2021:36.

NIRAS Sweden AB (2021), **Decentralised Evaluation of GROW Liberia Phase 1 – a Swedish-funded Market System Development Project in agricultural value chains in Liberia 2013-2020**. Sida Decentralised Evaluation 2021:1.

FCG Sweden (2023), **Evaluation of Guarantee Instrument and Technical Assistance Offered to Enat Bank to Increase Access to Capital For Women-Owned/Managed Small and Medium Enterprises in Ethiopia.** Sida Decentralised Evaluation 2023:12.

FCG Sweden & Tana Copenhagen (2023), **Evaluation of MSI Reproductive Choices Afghanistan, with a focus on the project funded by Sweden.** Sida Decentralised Evaluation 2023:12.

Ternström B., Ternström I., Davies A. (2018) **Evaluation of Sida Humanitarian Method Partners.** Sida Decentralised Evaluation 2019:12.

NIRAS Sweden AB (2022), **Evaluation of Sida's support towards food security as part of a wider food system.** Sida Decentralised Evaluation 2022:2.

NIRAS Sweden AB (2019), **Evaluation of Sida's Support to the Project "Organic Trade and Value Chain Development in East Africa" 2014-2019.** Sida Decentralised Evaluation 2019:23.

NIRAS Sweden AB (2025), **Evaluation of Sweden's support to health systems in Somalia through PHAS, SPIDER and WHO.** Sida Decentralised Evaluation 2025:06.

FCG Sweden (2022), **Evaluation of Enhanced Agricultural Productivity and Resilience to Climate Change through Solar Powered Irrigation in Zimbabwe.** Sida Decentralised Evaluation 2022:10.

Watkins F. (2018), **Evaluation of UN Women Country Programme in Afghanistan (2018).** Sida Decentralised Evaluation 2019:14.

Nordic Consulting Group (2021), **Final Evaluation of Financial Sector Deepening Mozambique (FSDMoç) 2016-2020.** Sida Decentralised Evaluation 2023:23.

Tana Copenhagen (2019), **Mid-Term Evaluation of the Sida and USAID Loan Portfolio Co-Guarantee and the Mali Finance for Food Security and Women Entrepreneurs (FFSWE) Programme.** Sida Decentralised Evaluation 2019:10.

FCG Sweden & Tana Copenhagen (2022) **Evaluation of Guarantee Instrument and Technical Assistance Offered to Enat Bank to Increase Access to Capital for Women-Owned/Managed Small and Medium Enterprises in Ethiopia.** Sida Decentralised Evaluation 2023:2.

NIRAS Sweden AB (2019), **Mid Term Evaluation of the Sida and USAID Loan Portfolio Co-Guarantee and The Mali Finance for Food Security and Women Entrepreneurs (FFSWE) Programme.** Sida Decentralised Evaluation 2019:10.

Verian & Conflict Management Consulting (2026), **Central Evaluation of Conflict Sensitivity in Sida's Development and Humanitarian Cooperation.** Sida Evaluation 2026:2.

Final Report from the Evaluation of Sida's work with Poverty

Main evaluation method: mixed-methods, synthesis analysis, literature review.

Positives: Across Zambia and South Sudan, Sida-supported interventions generated tangible benefits for poor populations, including improved incomes, food security, and access to essential services. In Zambia, interventions performed better as the systems are relatively stable while in South Sudan contributions have helped sustain basic services under extreme and fragile conditions.

Shortcomings: The impact results from Sida's work were uneven and remained limited in scale and transformative impact. Impact measures were often constrained by underuse of existing evidence in contribution designs and weak evaluability. Heavy reliance on partner-reported data, limited triangulation, and insufficient use of national data systems further created a verification gap and constrained deeper analysis.

SWEDISH INTERNATIONAL DEVELOPMENT COOPERATION AGENCY

Visiting address: Rissneleden 110, 174 57 Sundbyberg
Postal address: Box 2025, SE-174 02 Sundbyberg, Sweden
Telephone: +46 (0)8-698 50 00. Telefax: +46 (0)8-20 88 64
E-mail: sida@sida.se Web: sida.se/en

