Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-024-03213-z OPEN

A comprehensive framework for understanding microfinance performance evaluation methods

Javier Sierra [1✉], Victoria Muriel-Patino[1] & Fernando Rodríguez-López[1]

Multiple stakeholders in the microfinance sector seek to foster financial, social, and environmental development in a sustainable way by providing a wide range of products and services for financial inclusion. This heterogeneity is also reflected in the multiple methods to evaluate the performance of financial service providers. For this reason, it can be challenging for stakeholders to identify and understand the evaluation approaches that may be required to meet their needs and help them improve their performance. This study presents a comprehensive review of the existing approaches for the evaluation of the financial and social performance of microfinance. This research provides a systematic and comprehensive classification of microfinance performance evaluation methods, an explanation of evaluation methods and techniques, and a theoretical framework suitable to explain the applicability every method to assess different dimensions of microfinance. These results of this research are useful to help policymakers, donors, and practitioners understand and compare existing evaluation methods. Also, this framework enables the identification of the appropriate evaluation method according to the type of performance being examined and considering how to communicate this information effectively to the market.

[1] Department of Applied Economics, University of Salamanca, Paseo Tomás y Valiente, s/n. 37007, Salamanca, Spain. ✉email: jsierra@usal.es

Introduction

The microfinance sector has experienced rapid growth and development in recent decades (Armendáriz and Morduch, 2010; Hermes and Lensink, 2011). As a consequence, it has evolved towards a heterogenous environment (Chikweche et al., 2022), in which a wide range of institutions with different status and objectives provide a broad spectrum of financial and non-financial services to tackle the needs of socially and financially excluded people (Marconatto et al., 2016). For this reason, financial service providers (FSP) have developed a wide range of products and services to help the poor start productive activities, promote entrepreneurship, manage unexpected events, and smooth consumption, among other crucial issues for poverty reduction. Overall, FSP seeks to foster financial, social, and environmental development in a sustainable way by providing products and services for financial inclusion (Tanin et al., 2019; Fersi and Bougelbène, 2021). This heterogeneity can create some confusion for donors, policymakers, and practitioners, as it is not always easy to identify the different types of institutions operating in the sector, their objectives and characteristics, which can complicate the design of appropriate policies and strategies (Marconatto et al., 2016).

During its inception, microfinance primarily revolved around microcredit offerings in developing nations (Morduch, 1999). Nonetheless, as the sector progressed, there was a significant expansion in product diversity and complexity to cater to the diverse array of customer profiles and their corresponding requirements. In this context, microfinance has also gained increasing significance in enhancing financial inclusion in developed countries (Song et al., 2024; García-Pérez et al., 2017). Consequently, an increasing number of institutions and programs, whether public or private, are integrating microfinance mechanisms into their social, economic, and environmental development initiatives (Ferilli et al., 2024). In this context, microfinance is used today in a multitude of areas with very different objectives, yet consistently serves as a tool to drive positive social transformation. Its utilization spans initiatives concerning gender equality (Cruz Rambaud et al., 2022; Bapolisi et al., 2024), minority integration (Cruz Rambaud et al., 2022), digitization (Johri et al., 2024), mobile banking (Osabohien et al., 2024), youth empowerment (Rokhim et al., 2023), or entrepreneurship promotion (Coronel-Pangol et al., 2023), among other related issues.

In earlier years, microfinance was deemed a socially responsible activity. However, findings that showed that some institutions were moving away from their social mission in order to meet their financial targets suggested that the so-called mission drift phenomenon posed a risk for many FSPs seeking to achieve a set of social and financial aims: the double bottom line (Lopatta et al., 2017). In parallel, the extent to which microfinance is able to foster financial inclusion and improve well-being among the poor became a central point of debate (Lebovics et al., 2016; Bédécarrats et al., 2012). This issue increased pressure for evidence-based interventions in the microfinance sector (Authors, 2020). Stakeholders then began to struggle with a crucial challenge: while evaluation methods from the traditional banking sector were useful for assessing microfinance financial performance, there was a gap regarding practical methods to evaluate social performance and, as a result, data on social outcomes was scarce or simply absent (Spaggiari, 2016). This problem led different stakeholders to collaborate on developing tailored evaluation mechanisms to appraise the effectiveness of microfinance in generating the expected social change. As a result, several evaluation techniques have been developed and refined in recent years to provide social performance assessment with a similar level of consistency and acceptance as that of financial performance (D'Espallier and Goedecke, 2019).

This is reflected in the wide range of quantitative and qualitative methodologies devised to evaluate microfinance's social, economic, and environmental performance. This topic has rapidly gained attention in recent years (Akter et al., 2020), but the literature provides little guidance to practitioners and policymakers, especially regarding the usefulness of existing quantitative and qualitative research methods to assess different performance dimensions. This circumstance is especially relevant for those actors who do not have a deep knowledge of the sector, given that the diversity of institutions and evaluation methods sometimes also leads to the use of terms with very different meanings, which may contribute to generating even more confusion.

Against this background, this study examines the theory and practice of microfinance performance evaluation in an attempt to answer the following research questions: What are the existing approaches for the evaluation of microfinance performance? What are the differences between them? To address these questions, we have undertaken a comprehensive review of the existing approaches for the evaluation of the financial and social performance of microfinance, providing a systematic and comprehensive classification of microfinance performance evaluation methods.

The article makes several contributions to the literature on microfinance evaluation. To the best of our knowledge, this is the first article that presents all the existing methods to comprehensively and systematically evaluate microfinance's financial, social, and environmental performance. In addition, the study includes a theoretical framework to help practitioners, policymakers, investors, and other stakeholders understand existing evaluation methods, compare them, and identify the most suitable approaches to meet their specific needs. Specifically, this framework enables the identification of the appropriate evaluation method to be employed based on the type of performance being assessed, and most importantly, how such information is to be communicated to the market.

The article continues as follows: sections "Standard-based approaches for the evaluation of financial and social performance" and "Customized approaches to evaluate financial and social performance" review the existing methods for the evaluation of microfinance financial and social performance, first the standard-based and then the customized approaches. This is followed by a presentation of a comprehensive theoretical framework to guide microfinance performance evaluation. Finally, the last section presents the main conclusions and suggests further lines of research.

Standard-based approaches for the evaluation of financial and social performance

This section focuses on the forms of evaluation that are based on the review of internal institutional characteristics. First, those that assess and certify the FSP organization and procedures, and then those that seek to prove their compliance with a set of globally accepted standards.

Microfinance assessment and certification. Microfinance assessment is an evaluation approach based on the assumption that the study of the internal organization and procedures of an FSP, cross-checked against a set of commonly accepted standards, can be used as a proxy for the institution's social and financial performance, and therefore reliably judges its trustworthiness (Smart Campaign, 2014). Microfinance certification, a complementary stage to assessment, is designed to identify the achievement of certain levels of performance (according to the accepted standards), benchmarking results, and best practices.

While this appraisal could be performed by the FSP themselves with the support of an outside expert or via third parties, certification always requires an independent evaluation from a recognized firm or institution.

There are three main sets of standards used to measure FSP performance: Universal Standards for Social and Environmental Performance Management (Universal Standards, or USPM), Client Protection Standards (CPS), and Social Outcome Indicators (SOI). The USPM is a set of widely used standards designed to help institutions put clients at the center of their activity. In its current version, it includes seven dimensions: social strategy, committed leadership, client-centered products and services, client protection, responsible human resource development, responsible growth and returns, and environmental performance management (SPTF, 2021). These standards have been promoted by the Social Performance Task Force (SPTF) (SPTF, 2016). The CPS is a subset of 7 categories from the USPM, representing the 'minimum standards that clients should expect to receive when doing business with a financial service provider' (Smart Campaign, 2016). For its part, the SOI consists of a list of harmonized indicators to help stakeholders foster strong outcome management in four key categories: business and entrepreneurship; economic poverty, assets, and housing; resilience and vulnerability; and health (SPTF, 2022a).

The microfinance sector is currently involved in a digitalization process that has been accelerated by the COVID-19 pandemic (Pal et al., 2022), but this process is still encountering some challenges (Chiappetta Jabbour et al., 2020). For this reason, a proposal for a draft set of Standards for Responsible Digital Financial Services (DFS) was recently published (SPTF, 2022b). It is on issues such as agent management, algorithm bias, cybersecurity, and data privacy, among other topics relevant to digital microfinance.

These services are offered by four principal firms: MicroFinanza Rating (MFR), MicroRate (MR), M-CRIL, and Inclusion [Social Ratings] (ISR). Together, they offer a diverse set of assessment, certification, and rating elements—often under different commercial names—and are explained in detail below.

Social audits. A social audit is commonly done using social audit tools (SAT) to analyze institutional performance regarding a set of uniform standards. An evaluation that uses these tools can provide comprehensive insights into the strengths and weaknesses of a Microfinance Institution (MFI), helping to improve an organization's overall performance (Wardle, 2017). The Social Performance Indicators (SPI), the most popular SAT in the microfinance sector, was created by a working group led by CERISE. The SPI consists of a spreadsheet designed to help FSPs assess their activity and intentions in relation to the USPM and CPS. The fourth version of the SPI tool (SPI4) is available at no cost on the CERISE website, as is its online version (SPI Online). Both can be applied either by the FSP themselves in a self-assessment or in collaboration with trained and certified consultants or specialized rating agencies, as accompanying assessments (CERISE 2022b).

In addition to the SPI4, CERISE has collaborated with several stakeholders to develop several socially oriented assessment tools, such as ALigning INvestors due diligence and reporting with the Universal Standards (ALINUS), the Social Business Scorecard (SBS), SBS Light (a subset of SBS indicators), the MetODD-SDG, and the Impact-Driven Investor Assessment (IDIA). As shown in Table 1.

Client protection. Client protection is a particularly important part of the field of social performance and, therefore, merits a detailed explanation. The Smart Campaign, an active organization in social performance from 2009 to 2020, developed the Getting Started Questionnaire (GSQ), which allowed institutions to gain recognition through the Client Protection Certification (CPC). The tools and resources of the Smart Campaign were transferred to CERISE and the SPTF (CFF, 2021), which enhanced the GSQ to create a new tool, the Client Protection Self-Assessment Tool (CP SAT), and adapted the CPC to design a new certification method, the Client Protection Pathway (CPP) (SPTF, 2022c).

The CPP is a roadmap designed to help FSP implement the Client Protection Standards, based on a three-step process. The first step (Entry) refers to the commitment to implement client protection. It involves registering on the SPTF website, which creates a public profile for each FSP that shows the corresponding level of achievement as regards client protection. The second step (Progress) focuses on assessing and improving practices. Finally, the third step (Advanced) asks for evidence of progress and achievements. These three steps correspond to three levels defined by a set of eligibility criteria for certifying different levels of achievement, which can be calculated using the Excel CP SAT spreadsheet, based on the CPS. CERISE-SPTF has approved three institutions to certify client protection: MFR, M-CRIL, and ISR. In addition to these methods, a commitment to implement client protection can also be demonstrated via the SPI4, SPI Online, or ALINUS audit tools or by using a social rating, an approach explained in the following subsection. Disclosure documents are accepted if they are less than three years old, while certifications are valid for up to four years (SPTF, 2022c).

Table 1 CERISE social audit tools.

Tool	Use	Target clients	Main advantages
SPI4	Microfinance social audit	Microfinance institutions and other financial service providers	Aligned with USPM; SPI4 benchmark available
ALINUS	Microfinance social due diligence evaluation	Inclusive finance investors	Facilitates reporting and reduces data collection time; benchmarking of results
SBS	Social audit	Social enterprises	Analysis of the implementation practices of the social mission according to the 7P*
SBS light	Due diligence social audit	Impact investors	Sub-set of SBS; identifies and classifies SE according to the 7P
MetODD-SGD	Measure businesses' contribution to the Sustainable Development Goals (SDGs)	Social enterprises and impact investors	Selection of operational indicators covering 73 targets for 16 of the 17 SDGs
IDIA	Fast evaluation of financial intermediary practices	Impact investors, foundations, networks, fund managers, donors	Evaluates impact investment strategies; identifies means to create real impact

Source: (Cerise, 2022a). *7P is a set of seven categories for management benchmark practices: purpose, public, product, human resources policies, ethical principles, profits, partners.

Poverty scorecards. Poverty reduction is one of the main goals for financial service providers seeking to increase financial inclusion. The most reliable sources of poverty information are government-run large-scale representative household surveys (Berhane and Gardebroek, 2011), but these datasets, when available, are not always useful for estimating the poverty rates of directly targeted populations, apart from some other practical limitations (Christiaensen et al., 2011).

One popular instrument to overcome these problems is the Poverty Scorecards (PS), a simple and concise tool that, once calibrated for a certain country or region, can provide a quick proxy indicator of poverty based on a few quick answers from an individual or household. PS surveys can be designed to systematically assess microfinance performance by measuring poverty levels among the FSP's clients. Examples of popular PS include the Poverty Probability Index (PPI) (IPA, 2022), the Simple Poverty Scorecard (SPS) (Schreiner, 2014), and the Multidimensional Poverty Assessment Tool (MPAT) (IFAD, 2014).

Loan portfolio audits. Loan portfolio audits are an evaluation method that exposes the risks inherent in a FSP's loan portfolio, and design procedures to manage them. This evaluation approach, offered by MFR, M-CRIL, and MR, assesses compliance with national regulation and international standards, proves operational and credit quality accuracy, demonstrates consistency with internal practices and policies, and enhances portfolio administration systems (Microfinanza Rating, 2022d). Its final and overall goal is to identify possible underestimates in the figures for the portfolio at risk reported by an FSP (MicroRate, 2022c).

Truelift pro-poor seal of excellence. The Truelift Pro-Poor Seal of Excellence is a certification approach based on three principles: (1) purposeful outreach to people living in conditions of poverty; (2) services designed to meet the needs of people living in conditions of poverty; and (3) tracking the progress of people living in conditions of poverty. In turn, each of these principles is divided into four categories: (a) intent and strategy; (b) measurement, data quality, and analysis; (c) results achieved; and (d) use of findings. Assessing the degree to which these principles are implemented makes it possible to identify and recognize different levels of performance or milestones: Starting Commitment, Aspirant, Emerging Practitioner, Achiever, and Leader. Only those MFIs that satisfy all 30 indicators are eligible for the Truelift Certificate (Smart Campaign, 2016). These indicators have been included in the SPI4, and currently only MFR and M-CRIL grant specific Truelift certifications (Truelift, 2022).

Code of conduct. A commitment to implementing best practices in an organization can be also demonstrated by a Code of Conduct (CoC) assessment (SPTF, 2022c). Though it may have different structures, a CoC is a process that consists of three general stages: development, implementation, and compliance. It must be structured around clearly defined topics and developed through a consultative process in which all relevant stakeholders participate (SEEP Network and MasterCard Foundation 2015). Therefore, the scope of application and recognition of the CoC is diverse. One particular case is the specific CoC designed to assess the performance of FSP in the European Union: the European Code of Good Conduct for Microcredit Provision (ECGCMP). The ECGCMP is based on a set of standards corresponding to management, governance, risk management, reporting, and consumer and investor relations (European Commission, 2021). MFR certifies the ECGCMP in the European Union, while M-CRIL has been recognized by the Small Industries Development Bank of India (SIDBI) to perform Code of Conduct assessments of FSP in India.

Outcomes management. Outcomes Management is a multi-step institutional procedure of collection, analysis, and use of outcome information, based on a feedback loop to help stakeholders enhance social change. It consists of 10 essential steps divided into four stages: Stage 1 (Planning) includes (1) defining social change goals; (2) selecting indicators to capture social change; (3) selecting the evaluation method to measure social change; and (4) taking budget and human resource decisions to make the evaluation. Stage 2 (Data Collection) includes (5) putting systems in place to collect and capture data; and (6) putting systems in place to check data quality. Stage 3 (Analysis and Reporting of Findings) includes (7) data analysis; and (8) reporting and disseminating the results. Finally, stage 4 (Action) concentrates on (9) using outcomes data; and (10) reviewing the process (Sinha, 2017b).

Additional assessment and certification services. In addition to the services described above, some firms offer additional services designed to meet the specific needs of some stakeholders. MFR and M-CRIL provide evaluation products designed to assess digital services, such as the GSMA Mobile Money Certification (GMMC) in the case of the former and several products for digital credit providers in the case of the latter. MFR, M-CRIL, and ISR also offer due diligence services and country studies, which can be tailored to meet the FSP's needs.

Microfinance rating. The goal of microfinance rating is to evaluate performance by determining whether the MFI meets certain established standards included in a rating scale, providing information about the achieved performance level using a specific grade (The Rating Initiative, 2012). MFR, MR, M-CRIL, and ISR offer rating products, in addition to the assessment and certification services described above.

Institutional rating. Microfinance institutional rating is an assessment method that provides an outside evaluation of the long-term institutional viability and creditworthiness of an FSP. It relies on a comprehensive analysis of a set of key categories related to internal processes, risk management, and performance. MFR, MR, and M-CRIL offer an 'institutional rating', while ISR offers a 'financial sustainability rating' with a focus on financial performance. MR and M-CRIL also provide a rating outlook, which shows the expected evolution of the rating in the upcoming months. As shown in Table 2.

Social rating. The social rating evaluates the degree of success of MFIs in translating their social missions into practice, focusing on up to six crucial areas for social performance: country context, social performance management, social responsibility and client protection, depth of outreach, quality of services, and outcomes. Its aim is to provide useful information to help stakeholders comprehend policies and practices implemented by MFIs, hence identifying possibilities for improvement and the further enhancement of social performance (Clark and Sinha, 2013). All four evaluation firms offer specific products to assess this dimension of microfinance performance. As shown in Table 3.

Credit rating. Credit rating provides an external judgment on the FSP's ability to manage credit risks and on its short-term creditworthiness. As a credit rating is required for a supervised (or 'licensed') FSP to be allowed to receive deposits from the public, rating firms need recognition by national supervisory authorities to offer this service in any given country. Only two microfinance evaluation firms offer credit ratings: MicroFinanza Rating, recognized by the Superintendence of Banks and the

Table 2 Institutional rating scales.

Elements	MicroFinanza	MicroRate	M-CRIL	Inclusion [social ratings]
Commercial name	Institutional Rating	Institutional Rating	Institutional Rating	Financial Sustainability Rating
Areas covered (relative weight between parentheses, if available)	External context (4%) Governance and strategy (28%) Systems and controls (18%) Financial performance (31%) Portfolio quality (19%)	Financial profile Credit operations Portfolio quality Organization Corporate governance and management Microfinance environment	Governance and strategy Management systems Social risk Country risk Funding structure Financial performance	Stakeholder balance (25%) Effective use of financing resources (10%) Capital structure (10%) Asset quality (20%) Management strength (15%) Earnings power (10%) Liquidity (20%)
Rating scale levels (grades between parentheses)	Excellent (α α; α+; α) Good (α−; β+) Fair (β; β-) Poor (γ+; γ; γ-)	High long-term institutional viability, strong financial standing, and excellent performance (α+; α) Good long-term institutional viability, good financial standing, and good performance (α−; β+) Moderate long-term institutional viability, moderate financial standing, and satisfactory performance (β; β-) Low long-term institutional viability, inadequate financial standing, and weak performance (γ+; γ)	Strong (α+) Good (α; α-) Reasonable (β+) Moderate (β) Weak (β−; γ+; γ)	Excellent (78–100) Good (65–77) Adequate (50–64) Weak (30–49) Poor (0–29)

Source: (Microfinanza Rating, 2022c; M-CRIL, 2022a; Inclusion [Social Ratings] 2020; MicroRate, 2022b).

Superintendence of Solidarity-Based Economy in Ecuador, as well as the Supervisory Authority of the Financial System in Bolivia, and MicroRate, recognized by the Superintendence of Banking and Insurance in Peru. As shown in Table 4.

Microfinance investment vehicles rating. Many FSPs depend on other sources for funding, e.g., borrowing from pooled investment parties that take the form of Microfinance Investment Vehicles (MIV), which include a wide range of investment instruments. They connect the goals of impact investors with FSP funding needs (Dorfleitner et al., 2017). To assess their ability to fulfill the needs of both investors and FSP, a MIV rating evaluates the MIV's capacity to manage impact funds, their performance, and the generated social value, providing reliable information related to the MIV's risk and social and financial performance (Microfinanza Rating, 2022e). This kind of rating follows the MIV disclosure guidelines approved by the Consultative Group to Assist the Poor (CGAP) (CGAP, 2010) and the Principles for Responsible Investment (PRI) (PRI, 2019). MFR and M-CRIL offer MIV ratings, but only MFR provides a detailed and updated disclosure of its evaluation methodology. As shown in Table 5.

Additional rating services. In addition to the four key rating services explained above, some rating agencies offer specific products or services designed to meet the needs of particular stakeholders. MFR provides the Islamic Finance Rating and the Social Performance Roadmap. MFR also offers an Institutional Diagnostic. ISR also offers a specific social enterprise rating and a product with an environmental focus (ESG Rating), while M-CRIL provides corporate social responsibility support. M-CRIL also offers several quantitative, qualitative, and mixed-method services for market research and impact evaluation. Finally, MFR, M-CRIL, and ISR present several institutional diagnostic services, as well as tailored training and capacity-building products. Table 6 summarizes compliance-based approaches.

As regards the relevance of each evaluation method, only Microfinanza, MicroRate, and ISR present a complete list of the evaluations undertaken to date on their websites. As shown in Fig. 1, the most popular evaluation is institutional rating (602 evaluations), followed by social rating (480), client protection (186), credit rating (141), and institutional diagnostic (112). Regarding the number of evaluations undertaken by each rating company, Microfinanza Rating has performed 1205 evaluations, followed by MicroRate (442), and ISR (45).

Customized approaches to evaluate financial and social performance

Case studies and process evaluation. Microfinance evaluation research is dominated by quantitative studies, which either focus on measuring changes experienced by the beneficiaries of microfinance or by assessing FSP performance. However, microfinance services are not provided in a vacuum, and it is crucial to analyze the context in which microfinance institutions operate to fully comprehend the potential social, economic, and environmental effects. In contrast, qualitative research can be used to complement or replace quantitative methodologies, with the overall purpose of providing information regarding the historical, cultural, social, political, and religious context, as well as contributing to understanding how staff-client relationships evolve and determine microfinance performance (Shahjahan Chowdhury et al., 2020). Qualitative research relies on case studies and ethnographic fieldwork methodologies, applying interviews, observation, and focus group discussions to assess grounded theory and perform thematic analyses of specific contexts where the outcomes of the intervention may not be clear or precise or are simply difficult to measure from a quantitative perspective (Yin, 2012). Process evaluation can be understood in many different ways by diverse stakeholders, but the core of this complementary approach is that it helps researchers, managers, and practitioners understand the positive, negative, significant,

Table 3 Social rating scales.

	MicroFinanza Rating	MicroRate	M-CRIL	Inclusion [social ratings]
Commercial name	**Social rating**	**Social rating**	**Social rating**	**Social performance & impact rating**
Covered areas (relative weight between parentheses, if available)	Social performance management system (28%) Client protection and social responsibility (27%) Outreach (25%) Quality of the service (20%)	Definition and monitoring of social objectives Board, senior management, and staff commitment Product design and adequacy Social responsibility to the client Social responsibility to employees Responsible finances	Client protection Mission, governance, and strategy Operations alignment Depth of outreach, client feedback, and awareness	Social mission (15%) Governance (19%) Beneficiary focus (20%) Social outcome or impact (12%) Organization (11%) Environmental concerns (7%) Financial sustainability (16%)
Rating scale levels (grades between parentheses)	Excellent (^{S}AA) Good (^{S}A) Adequate (^{S}BB) Moderate (^{S}B) Weak (^{S}C) Poor (^{S}D)	First class (5 stars) Excellent (4 stars) Good (3 stars) Moderate (2 stars) Lack of social performance or, if any, it is weak (1 star)	Strong social commitment, excellent systems, evidence for strong and comprehensive adherence to social mission and values ($^{\Sigma}\alpha+$) Strong social commitment, very good systems, evidence for good adherence to social mission and values ($^{\Sigma}\alpha$) Strong social commitment, good systems, evidence for good adherence to social mission and values ($^{\Sigma}\alpha$-) Good social commitment, reasonable systems, evidence for reasonable adherence to social mission and values ($^{\Sigma}\beta+$) Good social commitment, moderate systems, evidence of partial adherence to social mission and values. Many aspects of social performance need to be institutionalized ($^{\Sigma}\beta$) Weak social commitment, weak systems, evidence of partial adherence to social mission and values ($^{\Sigma}\beta$-) Weak social commitment, weak systems, no evidence reflecting social mission or values ($^{\Sigma}\gamma+$) Very weak social commitment, very weak systems, no evidence reflecting social mission or values ($^{\Sigma}\gamma$)	Excellent (78–100) Good (65–77) Adequate (50–64) Weak (30–49) Poor (0–29)

Source: (M-CRIL, 2022b; Microfinanza Rating, 2022f; Inclusion [Social Ratings] 2022; MicroRate, 2022d).

Table 4 Credit rating scales.

Elements	MicroFinanza rating	MicroRate
Commercial name	Credit rating	Credit rating
Covered areas	External context	Credit risk
	Governance and strategy	Operational risk
	Organization and operations	Strategic risk
	Asset quality and structure	Financial risk
	Financial structure and management	
	Financial and operational results	
Rating scale grades	AAA	A+; A; A-
	AA	B+; B; B-
	A	C+; C; C-
	BBB	D+; D; D-
	BB	E
	B	
	C	
	D	
	E	

Source: (MicroRate, 2022a; Microfinanza Rating, 2022a).

Table 5 MFR's MIV rating scale.

Elements	MicroFinanza
Commercial name	MIV rating
Covered areas	Governance and management
	Funding and distribution
	Investment strategy and risk management
	Operations
	Fund performance
	Social performance management
	Responsible investment
	Social results
Rating scale grades	Excellent
	Good
	Adequate
	Fair
	Poor

Source: (Microfinanza Rating, 2022e).

insignificant, or unexpected results of an intervention (Dixon and Bamberger, 2022).

Qualitative impact protocol (QuIP). It is an evaluation method to appraise causal devices and infer causality links in program evaluations based on contribution analysis. The QuIP collects evidence of a project's results and effects through narrative causal statements gathered directly from intended project beneficiaries. In this context, beneficiaries reflect on the main changes in their lives and are encouraged to reveal what they recognize to be the primary drivers of these developments, and to whom or what they attribute them (Copestake et al., 2019). However, since a control group is not used to check the attribution of change, it is questionable whether this method can truly be considered as belonging to the impact evaluation category.

Financial diaries. One of the most important challenges faced by the poorest households results from their low level of income, combined with uncertainty about when such income can be expected, including a lack of ability or skills to budget or manage money. Financial Diaries consists of a series of monthly

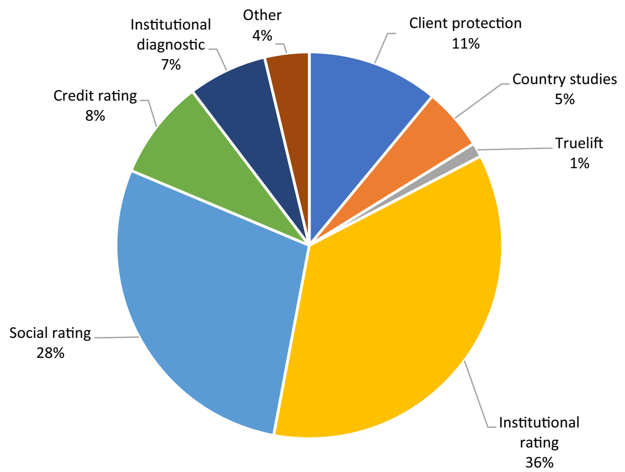

Fig. 1 Relevance of the different standard-based evaluations.

Table 6 Compliance-based approaches offered by MicroFinanza, MicroRate, M-CRIL, and inclusion.

Approach	Method	MicroFinanza	MicroRate	M-CRIL	Inclusion [social ratings]
Assessment	Social audit	X		X	X
	Client protection	X		X	X
	Due diligence	X		X	X
	Loan portfolio audit	X	X	X	
	Code of conduct	X		X	
	Digital services				
	Outcomes management			X	
	CSR support			X	
	Social performance roadmap	X			
	Country studies	X	X	X	
Certification	Truelift	X		X	
	Digital services	X			
Rating	Institutional rating	X	X	X	X
	Social rating	X	X	X	X
	Credit rating	X	X		
	MIV rating	X		X	
	Islamic finance rating	X			
	ESG rating				X
	Social enterprise rating				X

Source: Rating agency websites.

interviews for some period, usually a year, to help households understand how they use money and find ways to improve their budget (Collins et al., 2009). Applications of this methodology suggest that the poor may benefit from initiatives and public policies to address their income uncertainty, hence increasing their economic stability (Biosca et al., 2020; McHugh et al., 2019).

Efficiency studies. The evolution of the microfinance industry has helped advance management practices. In recent times, many FSPs have transformed towards improved commercialization, seeking to reduce their dependency on external subsidies and aiming to achieve financial self-sufficiency (D'Espallier et al., 2017). Studies on the efficiency of microfinance are increasingly crucial, in that they provide information about the use of resources to enhance the overall mission of microfinance. This type of research is largely supported by two main estimation methodologies: Stochastic Frontier Approach, a parametric technique to estimate efficiency boundaries applying econometric inference, and Data Envelopment Analysis, a nonparametric method to construct efficiency frontiers applying mathematical optimization (Fall et al., 2018).

Impact evaluation: quasi-experimental and experimental methods. Despite the potential benefits of microfinance to expand financial inclusion and help the poor (Morduch, 1999; Karlan and Appel, 2012), some programs, products, or services that might seem promising may not actually generate the expected results effectively (Karlan and Appel, 2016; Roodman, 2011). Therefore, it is crucial to understand whether and how microfinance fulfills expectations in order to accurately identify the mechanisms by which beneficiaries respond to intervention (Banerjee and Duflo, 2012). In this regard, impact evaluation aims to identify and quantify the extent to which changes in client well-being are attributable to a specific microfinance program, product, or service.

Two major research methods address this issue: experimental and quasi-experimental. In experimental settings, program allocation is based on a random distribution of potential beneficiaries who are divided into a treatment group (who experience the intervention) and a control group (who are not affected by the intervention and form the counterfactual) (Glennerster and Takavarasha, 2013). This type of evaluation has become very popular in the field of Development Economics, and microfinance. The Jameel Abdul Latif-Poverty Action Lab (J-PAL), Innovations for Poverty Action (IPA), and the International Initiative or Impact Evaluation (3ie) are all examples of actors using this methodology. MFR has developed an impact assessment methodology to apply a similar framework (Microfinanza Rating, 2022b).

Nevertheless, it may not be always possible or desirable to use it for different technical, economic, ethical, and political reasons (Bédécarrats et al., 2019). When the valuation setting is not purely experimental, some quasi-experimental techniques have been applied to try to mimic randomization effects, including matching methods, regression discontinuity analysis, instrumental variables, control function methods, and quantile regressions (Glewwe and Todd, 2022).

Microfinance performance evaluation and the theory of change

Having examined the existing methods for the evaluation of microfinance, the next step involves identifying the usefulness and applicability of each approach to meet different stakeholders' needs and objectives. To achieve this, we applied a two-pronged framework defined by the theory of change (TOC) and the nature of the framework of reference used to analyze the success of the intervention (Authors, 2021). A TOC describes the causal process that leads from intentions to an ultimate objective and can be expressed both graphically and as a narrative. It is structured into five stages that describe a causal process to generate a positive social, economic, or environmental change, which in the case of the microfinance sector, are as follows:

i. Inputs include both intent (the specific goals that FSP wants to achieve, including their social mission) and design (internal organization, structure, and practices).
ii. Activities refer to how products and services will be managed to meet clients' needs.
iii. Outputs represent the direct results or expected consequences of activities.
iv. Outcomes indicate unexpected or uncontrolled changes that are plausibly associated with microfinance.
v. Finally, impact shows changes in well-being that can be attributed to microfinance, all other things excluded.

Although it may not always be explicit in all interventions it is possible to identify an underlying theory of change in every scenario (Jackson, 2013). Since the TOC symbolizes the coherent structure of the chronological steps necessary to achieve an expected change, it also represents a hierarchy of results: the first three steps represent measures of implementation, outcomes signal changes that go beyond outputs and finally, impact refers to effects that are a direct consequence of microfinance (Sinha, 2017a).

As explained in the previous section, the framework of reference used to make microfinance performance evaluations can be based either on commonly agreed standards (USPM, CPS, SOI, etc.) or on customized research settings, and evaluation approaches can be classified as standard-based or customized. Figure 2 shows a map of the existing practical methods for the evaluation of microfinance performance within the theoretical framework. The commercial names of the evaluation tools and examples of institutions and firms that do these evaluations are in parentheses. Names of institutions and firms are in italics.

The framework of reference also provides non-exhaustive information for three important variables to be considered when selecting an evaluation strategy: cost, time, and the potential to generalize the results. Standard-based methods, in general, are more accessible in terms of implementation costs and duration, especially if they are applied by FSP themselves, possibly with the support of an expert, instead of by a third party. Moreover, since these methods rely on commonly agreed indicators, the results of the evaluations permit comparisons between institutions and benchmarking. In contrast, non-standard-based methods usually require a longer period for full implementation, mainly because of the need to assess changes at different points in time. There can also be important differences regarding implementation costs, depending on the number of clients or beneficiaries participating in the study, the size of the research team, the nature of the product or service subject to evaluation, and the time required to properly assess the performance of the intervention. Significantly, although non-standard-based methods are useful to evaluate specific contexts or projects, their outcomes can be hard to generalize, making it difficult to extrapolate the results to different settings.

Conclusion

In light of the heterogeneity and rapid evolution of the microfinance industry, it can be challenging for stakeholders to identify existing methods to assess microfinance performance,

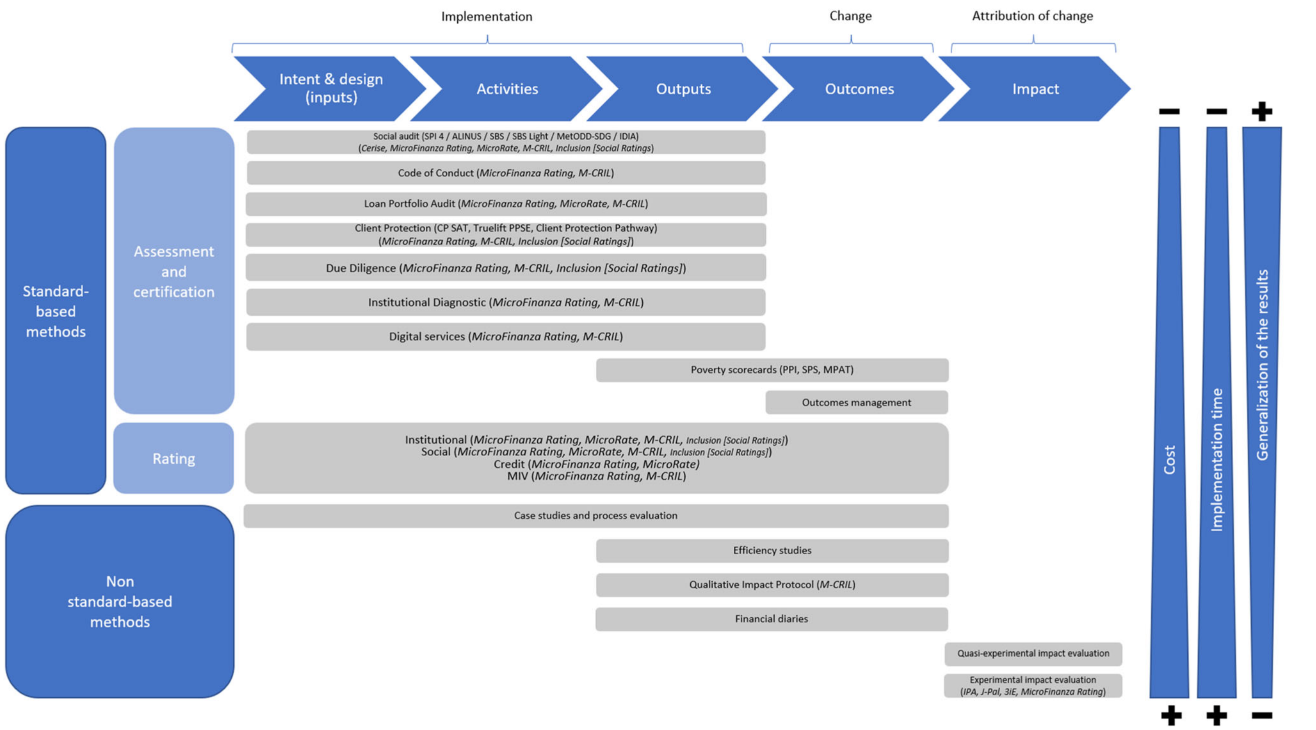

Fig. 2 Practical methods for the evaluation of microfinance performance.

comprehend its characteristics, and decide what method is most suitable to help them achieve their financial and social objectives. In contrast, the theoretical framework presented in this study explains how each method addresses the different steps in the microfinance theory of change, making it possible to compare the different techniques and help stakeholders identify and decide what method best fits their needs. At the same time, it provides an estimation of the differences between methods in terms of implementation costs and time and the possibility of generalizing the results.

This article presents the first attempt to map practical methods for the evaluation of microfinance performance. To that end, the theoretical framework presents a comprehensive approach that can include all products and services, with the caveat that it was impossible to analyze the different categories and their corresponding methods in depth within the confines of this study. Neither has there been room to provide a detailed analysis of the differences between the various methods in each of the categories, which makes it difficult to offer a critical view of the techniques. Another limitation lies in the inability to accurately determine the precise cost of different methods and the time required for their implementation. This analysis would require a much more detailed assessment of each evaluation method, which would cause the article's length to increase excessively. At the same time, information about the cost of each method is not available, making it impossible to accurately determine the cost of implementing the different techniques. Undoubtedly, these are two fundamental variables that can guide decision-making on which method to choose, depending on the needs of each actor and the available budget.

Our study adds to the body of knowledge on microfinance performance evaluation, helping to clarify the fuzzy thinking and complexity that can characterize the microfinance sector and existing evaluation methods. Further research in this area would benefit from richer data on the use of evaluation methods, especially if information regarding the results of the assessments and their implications is made available. In this regard, further evidence on the outcomes of microfinance can help analyze the relationship between the use of different evaluation methods and the positive or negative effects of various programs, products, and services. In this sense, being able to more precisely understand each evaluation method's capacity to contribute to achieving the results expected by organizations would be highly useful in assisting different institutions in making decisions about the most suitable method for their needs. At the same time, the dynamism of the sector continues to increase the number of products and services, especially in the field of digitization. Undoubtedly, it is possible that new performance evaluation methods may emerge in the coming years, so it would be interesting to study their compatibility with the model presented in this article.

Data availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

Received: 19 June 2023; Accepted: 21 May 2024;
Published online: 14 June 2024

References

Akter S, Uddin H, Tajuddin AH (2020) Knowledge mapping of microfinance performance research: a bibliometric analysis. Int J Soc Econ 48(3):399–418. https://doi.org/10.1108/IJSE-08-2020-0545

Armendáriz B, Morduch J (2010) The economics of microfinance. 2nd edn. The MIT Press, Cambridge, Massachusets, 10.1086/523604

Banerjee A, Duflo E (2012) Poor economics. A radical rethinking of the way to fight global poverty. 1st edn. PublicAffairs

Bapolisi WA, Makelele J, Ferrari G, Kono-Tange L, Bisimwa G, Schindler C, Merten S (2024) Engaging men in women's empowerment: impact of a complex gender transformative intervention on household socio-economic and health outcomes in the Eastern Democratic Republic of the Congo using a longitudinal survey. BMC Public Health 24(1):1–11. https://doi.org/10.1186/S12889-024-17717-5/TABLES/5

Bédécarrats F, Baur S, Lapenu C (2012) Combining social and financial performance: a paradox? Enterp Dev Microfinance 23(3):241–58. https://doi.org/10.3362/1755-1986.2012.024
Bédécarrats F, Guérin I, Roubaud F (2019) All that glitters is not gold. The political economy of randomized evaluations in development. Dev Change 50(3):735–62. https://doi.org/10.1111/DECH.12378
Berhane G, Gardebroek C (2011) Does microfinance reduce rural poverty? Evidence based on household panel data from northern Ethiopia. Am J Agric Econ 93(1):43–55. https://doi.org/10.1093/AJAE/AAQ126
Biosca O, McHugh N, Ibrahim F, Baker R, Laxton T, Donaldson C (2020) Walking a Tightrope: using financial diaries to investigate day-to-day financial decisions and the social safety net of the financially excluded. 689(1): 46–64. https://doi.org/10.1177/0002716220921154
Cerise (2022a) Les Outils de Référence de CERISE: Une Grande Famille. https://cerise-sptf.org/
Cerise (2022b) SPI4 Cerise. https://cerise-spm.org/en/social-performance/
CFF (2021) The Smart Campaign. https://www.centerforfinancialinclusion.org/about/what-we-do/the-smart-campaign
CGAP (2010) Microfinance investment vehicles disclosure guidelines. www.cgap.org
Chiappetta J, Jose C, De Camargo Fiorini P, Ndubisi NO, Queiroz MM, Piato ÉL (2020) Digitally-enabled sustainable supply chains in the 21st century: a review and a research agenda. Sci Total Environ 725(Jul):138177. https://doi.org/10.1016/J.SCITOTENV.2020.138177
Chikweche T, Lappeman J, Egan P (2022) Marketing financial services in Africa: exploring the heterogeneous middle-class consumer across nine countries. J Financ Serv Mark 29(1):1–16. https://doi.org/10.1057/S41264-022-00179-4/FIGURES/6
Christiaensen L, Lanjouw P, Luoto J, Stifel D, Christiaensen L, Lanjouw P, Lanjouw P, Luoto J, Stifel D (2011) (2011) Small area estimation-based prediction methods to track poverty: validation and applications. J Econ Inequality10(2):267–97. https://doi.org/10.1007/S10888-011-9209-9
Clark H, Sinha F (2013) Social rating guide. The Rating Initiative
Collins D, Morduch J, Rutherford S, Ruthven O (2009) Portfolios of the poor. Princeton University Press
Copestake J, Morsink M, Remnant F (2019) Attributing development impact: the qualitative impact protocol case book. Practical Action Publishing, Rugby, UK, 10.3362/9781780447469
Coronel-Pangol K, Paule-Vianez J, Orden-Cruz C (2023) Conventional or alternative financing to promote entrepreneurship? An analysis of female and male entrepreneurship in developed and developing countries. Int Entrepreneurship Manag J 20(1):163–87. https://doi.org/10.1007/S11365-023-00906-W/TABLES/4
Cruz Rambaud S, López Pascual J, Moro-Visconti R, Santandreu EM (2022) Should gender be a determinant factor for granting crowdfunded microloans? Humanit Soc Sci Commun 29(1):1–13. 10.1057/s41599-022-01475-z
Cruz R, Salvador, Pascual JL, Santandreu EM (2023) A socioeconomic approach to the profile of microcredit holders from the hispanic minority in the USA. Financ Innov. 9(1):1–25. https://doi.org/10.1186/S40854-022-00422-W/TABLES/9
D'Espallier B, Goedecke J (2019) Social performance measurement in microfinance. In: A research agenda for financial inclusion and microfinance. In: Hudon M, Labie M, Szafarz A (eds). Edward Elgar. pp. 62–74
D'Espallier B, Goedecke J, Hudon M, Mersland R (2017) From NGOs to banks: does institutional transformation alter the business model of microfinance institutions? World Dev 89:19–33. https://doi.org/10.1016/j.worlddev.2016.06.021
Dixon V, Bamberger M (2022) Incorporating process evaluation into impact evaluation: what, why and how." WP 50. https://www.3ieimpact.org/evidence-hub/publications/working-papers/incorporating-process-evaluation-impact-evaluation-what?utm_source=Subscribe+to+3ie&utm_campaign=a82a00ffcb-EMAIL_CAMPAIGN_april_newsletter&utm_medium=email&utm_term=0_58acc0a9f7-a82a00ffc
Dorfleitner G, Röhe M, Renier N (2017) The access of microfinance institutions to debt capital: an empirical investigation of microfinance investment vehicles. Q Rev Econ Financ 65(Aug):1–15. https://doi.org/10.1016/J.QREF.2016.06.005
European Commission (2021) European code of good conduct for microcredit provision. https://doi.org/10.2767/47915
Fall F, Akim AM, Wassongma H (2018) DEA and SFA research on the efficiency of microfinance institutions: a meta-analysis. World Dev. 107:176–88. https://doi.org/10.1016/j.worlddev.2018.02.032
Ferilli GB, Palmieri E, Miani S, Stefanelli V (2024) The impact of FinTech innovation on digital financial literacy in Europe: insights from the banking industry. Res Int Bus Financ 69(Apr):102218. https://doi.org/10.1016/J.RIBAF.2024.102218
Fersi M, Bougelbène M (2021) Financial and social efficiency analysis of islamic microfinance institutions. Int J Emerg Market (ahead-of-print). https://doi.org/10.1108/IJOEM-02-2020-0197/FULL/PDF
García-Pérez I, Muñoz-Torres MJ, Fernández-Izquierdo MÁ (2017) Microfinance literature: a sustainability level perspective survey. J Clean Prod 142(Jan):3382–95. https://doi.org/10.1016/J.JCLEPRO.2016.10.128
Glennerster R, Takavarasha K (2013) Running randomized evaluations: a practical guide. 1st edn. Princeton University Press
Glewwe P, Todd P (2022) Impact evaluation in international development. Theory, methods, and practice. World Bank Group, Wasahington, DC, 10.1596/978-1-4648-1497-6
Hermes N, Lensink R (2011) Microfinance: its impact, outreach, and sustainability. World Dev. 39(6):875–81. https://doi.org/10.1016/j.worlddev.2009.10.021
IFAD (2014) The multidimensional poverty assessment tool. International Fund for Agricultural Development
Inclusion [Social Ratings] (2020) Methodology report: financial sustainability rating system. 2020. https://inclusionsocialratings.org/product/methodology-report-financial-sustainability-rating-system/
Inclusion [Social Ratings] (2022) Social performance & impact rating methodology. https://inclusionsocialratings.org/product/spir-methodology-english/
IPA (2022) Poverty Probability Index. https://www.povertyindex.org/
Jackson ET (2013) Interrogating the theory of change: evaluating impact investing where it matters most. J Sustain Financ Invest 3(2):95–110. https://doi.org/10.1080/20430795.2013.776257
Johri A, Asif M, Tarkar P, Khan W, Rahisha, Wasiq M (2024) Digital financial inclusion in micro enterprises: understanding the determinants and impact on ease of doing business from world bank survey. Humanit Soc Sci Commun 11(1):1–10. https://doi.org/10.1057/s41599-024-02856-2
Karlan D, Appel J (2012) More than good intentions. Plume, London
Karlan D, Appel J (2016) Failing in the field. Princeton University Press
Lebovics M, Hermes N, Hudon M (2016) Are financial and social efficiency mutually exclusive? A case study of vietnamese microfinance institutions. Ann. Public Cooperative Econ 87(1):55–77. https://doi.org/10.1111/apce.12085
Lopatta K, Tchikov M, Jaeschke R, Lodhia S (2017) Sustainable development and microfinance: the effect of outreach and profitability on microfinance institutions' development mission. Sustain Dev 25(5):386–99. https://doi.org/10.1002/SD.1663
M-CRIL (2022a) Microfinance institutional rating. M-CRIL
M-CRIL (2022b) Social rating. https://www.m-cril.com/services-2/ratings-certifications/ratings-inclusive-financial-service-providersfsps/social-rating/
Marconatto D, Cruz LB, Pedrozo EA (2016) Going beyond microfinance fuzziness. J Clean Prod 115:5–22. https://doi.org/10.1016/J.JCLEPRO.2015.12.070
McHugh N, Biosca O, Baker R, Ibrahim F, Donaldson C (2019) Innovating on methods to understand the relationship between finances and wellbeing, 129–44. https://doi.org/10.1007/978-3-030-05535-6_6
Microfinanza Rating (2022a) Calificación de Riesgos. Metodología. https://www.mf-rating.com/products/credit-rating/
Microfinanza Rating (2022b) Impact assessment. https://www.mf-rating.com/products/impact-assessment/
Microfinanza Rating (2022c) Institutional rating. https://www.mf-rating.com/products/institutional-rating/
Microfinanza Rating (2022d) Loan portfolio audit. https://www.mf-rating.com/products/loan-portfolio-audit/
Microfinanza Rating (2022e) MIV rating. https://www.mf-rating.com/products/miv-rating/
Microfinanza Rating (2022f). Social rating. https://www.mf-rating.com/products/social-rating/
MicroRate (2022a) Credit rating. http://www.microrate.com/productos/15/
MicroRate (2022b) Institutional rating. http://www.microrate.com/productos/14/
MicroRate (2022c) Loan portfolio assessment. http://www.microrate.com/productos/17/
MicroRate (2022d) Social rating. http://www.microrate.com/productos/16/
Morduch J (1999) The microfinance promise. J Econ Lit 37(4):1569–1614. https://doi.org/10.1257/jel.37.4.1569
Osabohien R, Jaaffar AH, Akpa AF, Jakovljevic M (2024) Mobile money, medical cost anxiety and welfare of individuals within the reproductive age in Malaysia. Humanit Soc Sci Commun 11(1):1–10. https://doi.org/10.1057/s41599-024-02767-2
Pal A, Dey S, Nandy A, Shahin S, Singh PK (2022) Digital transformation in microfinance as a driver for sustainable development. Handbook of Sustainability Science in the Future. Springer. pp. 1–21
PRI (2019) Seven guidelines for investors in inclusive finance. https://www.unpri.org/thematic-and-impact-investing/seven-guidelines-for-investors-in-inclusive-finance/4043.article
Rokhim R, Lubis AW, Novianti N (2023) The efficiency of education and microcredit programs on young adults' independence. J Small Bus Entrepreneurship 35(5):677–96. https://doi.org/10.1080/08276331.2022.2035170
Roodman D (2011) Due diligence: an impertinent inquiry into microfinance. Centre for Global Development
Schreiner M (2014) The process of poverty-scoring analysis. Microfinance.com
SEEP Network, MasterCard Foundation (2015) Effective codes of conduct: a guide for microfinance associations. https://seepnetwork.org/files/galleries/1554_SEEP-Code_of_Conduct-ENGLISH.pdf

Shahjahan Chowdhury M, Ahmmed F, Hossain MI (2020) Methodological dilemma in microfinance research: applicability of a qualitative case study design. Qual Rep. 25(2):271–90
Sinha F (2017a) Guidelines on outcomes management for financial service providers. Key terms and 10 STEPS for practical, cost-effective, outcomes management. social performance task force. https://sptf.info/images/Guidelines-on-Outcomes-Management-for-FSPs.pdf
Sinha F (2017b) Making the case for outcomes management to financial service providers. What is outcomes management? How do financial service providers use outcomes data? Social Performance Task Force
Smart Campaign (2014) A guide to client protection assessments. Smart Campaign
Smart Campaign (2016) Client protection certification standards. Smart Campaign
Song X, Li J, Wu X (2024) Financial inclusion, education, and employment: empirical evidence from 101 countries. Humanit Soc Sci Commun 11(1):1–11. https://doi.org/10.1057/s41599-024-02630-4
Spaggiari L (2016) Guidelines on outcomes management for investors. European Microfinance Platform and Social Performance Task Force
SPTF (2016) Universal standards for social performance management. Version 2.0. Social Performance Task Force
SPTF (2021) The universal standards for social and environmental performance management. Social Performance Task Force
SPTF (2022a) List of harmonized social outcome indicators. https://sptf.info/working-groups/outcomes
SPTF (2022b) Responsible DFS standards. https://sptf.info/working-groups/dfs-standards
SPTF (2022c) Social performance task force. Client Protection. https://sptf.info/
Tanin TI, Mobin MA, Ng A, Dewandaru G, Salim K, Nkoba MA, Razak LA (2019) How does microfinance prosper? An analysis of environmental, social, and governance context. Sustain Dev 27(6):1001–22. https://doi.org/10.1002/SD.1952
The Rating Initiative (2012) Rating guide
Truelift (2022) Get assessed. https://sealofexcellence.wordpress.com/get-assessed-how-to-apply-for-truelift-recognition/
Wardle L (2017) The universal standards for social performance management implementation guide. Version 2. Social Performance Task Force
Yin RK (2012) Case study methods. In: APA Handbook of Research Methods in Psychology, vol 2: research designs: quantitative, qualitative, neuropsychological, and biological. American Psychological Association. pp. 141–155

Acknowledgements
The authors would like to thank their colleagues Maricruz Lacalle-Calderón and Begoña Gutiérrez-Nieto for their comments and suggestions received during the research process.

Author contributions
JS, VMP, and FRL conceived the study wrote the manuscript, and reviewed the manuscript. J.S. acted as the corresponding author.

Competing interests
The authors declare no competing interests.

Ethical approval
This study was performed in line with the principles of the Declaration of Helsinki. Ethical approval was not required as the study did not involve human participants.

Informed consent
This article does not contain any studies with human participants performed by any of the authors.

Additional information
Correspondence and requests for materials should be addressed to Javier Sierra.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.