Enhancing Clients' Livelihoods: Design, Implementation and Impact Evaluation of Decent Work Innovations in Microfinance

Bernd Balkenhol, Markus Frölich, Niels Kemper and Patricia Richter

Microfinance, impact, evaluation, innovation, decent work, child labour, formalization, occupational safety and health, risk management

Mikrofinanzierung, Impact, Evaluation, Innovation, menschenwürdige Arbeit, Kinderarbeit, Formalisierung, Arbeitssicherheit, Risikomanagement

In 2007 the International Labour Organization (ILO) launched the "Microfinance for Decent Work" (MF4DW) action research programme, which was intended to foster the development of innovative financial and non-financial products and services for microfinance clients. A bottom-up approach was pursued where microfinance institutions could propose innovations that promised to have a social impact. The piloting of these interventions was complemented by an impact evaluation. In this article the 17 selected innovations are described as well as the accompanying impact evaluation approaches. Even if final evaluation results are not available yet, lessons learnt during the implementation of the MF4DW initiative seem worthwhile to be made available to other programmes targeting a systematic improvement of microfinance impact.

Die Internationale Arbeitsorganisation (ILO) lancierte im Jahr 2007 das Forschungsprogramm „Microfinance for Decent Work" (MF4DW), um die Entwicklung von innovativen Mikrofinanzprodukten zu unterstützen, die einen Beitrag zur Verbesserung der Arbeitsbedingungen im Umfeld der Mikrofinanzkunden, kurz zu menschenwürdiger Arbeit, versprechen. Mikrofinanzinstitutionen weltweit konnten im Rahmen eines Wettbewerbs innovative Produktideen einbringen, aus denen 17 Innovationen ausgewählt wurden, welche eine positive soziale Wirkung erwarten liessen. Ein zentraler Bestandteil dieses Forschungsprogramms war die begleitende quantitative Evaluation der sozialen Wirkungen der Innovationen. In diesem Artikel werden die ausgewählten Innovationen beschrieben, ergänzt um eine Darstellung des Evaluationsansatzes. Dadurch sollen die Erfahrungen aus der Implementierung dieses Pilotprojekts,

auch wenn die abschliessenden Ergebnisse der Wirkungsanalysen noch ausstehen, für zukünftige Initiativen zur gezielten Verbesserung der sozialen Wirkungen von Mikrofinanzierung nutzbar gemacht werden.

1. Introduction

Microfinance is receiving substantial interest in the development community as well as academic circles as a development tool that is hoped to lessen poverty and vulnerability while even being profitable to the financial institutions. In contrast to other development strategies microfinance claims to succeed without massive aid transfers. The prospect of a positive social impact explains the interest of public institutions in nurturing and developing microfinance.

Recent studies, however, have questioned the impact of microfinance, or at least of microenterprise loans (e.g. *Banerjee, et al*. 2009; *Karlan/Zinman* 2009). Rather than discrediting microfinance, the studies instead reinforce a conclusion that many experts had already reached: microcredit is not sufficient. Indeed, the development of a range of *financial and non-financial services* is needed to further the social impact of microfinance. In fact, this can have a dual impact. Microfinance institutions (MFIs) initiate new products and services not just to improve the livelihoods of their clients, but also to improve the MFI's financial sustainability. Developing successful new products, continuously improving them, could help to expand the frontier of finance and demonstrate that microfinance can make a difference to the employment and income situation of the poor.

To this end, in 2007 the International Labour Organization (ILO) launched the "Microfinance for Decent Work" (MF4DW) action research programme with the financial support of the German government. "Decent work sums up the aspirations of people in their working lives. It involves opportunities for work that is productive and delivers a fair income, security in the workplace and social protection for families, better prospects for personal development and social integration, freedom for people to express their concerns, organize and participate in the decisions that affect their lives and equality of opportunity and treatment for all women and men."[1] It was anticipated that the following aspects of decent work would be most tangibly affected by changes in the access to financial and non-financial services: child labour, vulnerability/risk management, occupational safety and health, formalization, job creation and productive employment, and women's empowerment.

In this article we describe in detail the microfinance innovations of the MF4DW programme as well as their implementations and briefly sketch the evaluation results with a particular focus on lessons learnt from the evaluation process. This article therefore intends to reach two goals: On the one hand, to illustrate innovative microfinance products that aim to address work-related challenges of microfinance clients, and on the other hand, to illustrate the lessons learnt from the process of impact evaluations in order to develop good-practice evaluations and shape future evaluations that provide precise knowledge about impacts and their heterogeneity. Assuming that no one-size-fits-all product exists, we believe that a sequential process of product and process innovations with subsequent evaluations followed by further product improvements is needed to weed out ineffective products and continuously develop a menu of solutions and products that have a

1 http://www.ilo.org/global/topics/decent-work/lang--en/index.htm.

positive impact on microfinance clients and their families. The quantitative impact evaluations are still on-going such that no definite recommendations can yet be made about the potential of these innovations. Yet, many lessons can already be drawn about designing innovations so that they can be evaluated, and the difficulties in implementing or marketing the innovation. The final impact publications will be made available towards the end of 2012 on the ILO website.[2]

2. What is the Microfinance for Decent Work action research programme?

2.1 The basic idea

The MF4DW action research programme originated in 2005 when, in the context of the International Year of Microcredit, the ILO adopted a policy on "Microfinance for Decent Work". In this the ILO acknowledges that microfinance can make a powerful contribution to decent work in a variety of ways, for example through:

- creating conditions for wage and self-employment;
- reducing vulnerability;
- promoting gender equality and
- empowering the poor.[3]

To live up to the policy statement, the ILO set out to create knowledge about such contributions and to encourage microfinance institutions to gear their operations towards effective work improvements, notably by developing innovative products and services. In 2006 an action research project ("Microfinance for Decent Work") was designed to experiment with innovations in microfinance products, services and operating modes and to track the impact on the welfare of clients, focusing on the different aspects of decent work.

The action research had three objectives:

- generate knowledge: to throw light on the effects of modifications in the services offered by MFIs: do they really make a difference to jobs, incomes, risk management strategies used by poor households? The underlying idea was that MFIs, rather than going about their business as usual, can go out of their way to enhance clients' livelihoods.
- demonstrate a business case: amongst these innovations the MF4DW project aimed at identifying those that are also a good business case for the institution itself (win-win situation). The underlying idea being that in the longer term MFIs will continue only with innovations that generate some revenue.
- prepare better policy: if these effects can be demonstrated, then MF4DW would also make a case to governments and aid agencies for smart policies so as to provide start-up funding to MFIs for those innovations that seem to work.

2.2 The focus on decent work

One of the first steps in the design of MF4DW was to unpack the concept of decent work and break it up into measurable dimensions. ILO tools such as Decent Work Country Pro-

2 http://www.ilo.org/employment/areas/social-finance/WCMS_168033/lang--en/index.htm.
3 ILO 2005: ILO policy statement: Microfinance for Decent Work, GB. 94/ESP/3, 294th Session.

files,[4] the Key Indicators of the Labour Market[5] (KILM) classification and indicators used by the Statistical Information and Monitoring Programme on Child Labour (SIMPOC)[6] were reviewed to extract suitable indicators for micro level data generation. Another consideration was to ensure a certain spread across decent work themes so as to enlist the interest and support of the full range of ILO technical programmes and demonstrate the potential usefulness of financial inclusion across the work of the ILO. As a result, the MF4DW defined the following decent work areas:

- Child labour: earlier observations signalled that poor households are sensitive to changes in earnings of family members, changes in time allocation within the family and changes in the demand for unremunerated labour and hence more vulnerable to engaging in child labour; child labour could also be measured quite reliably in an indirect way, using data about school attendance.
- Risk management and over-indebtedness: income shocks occur as a result of catastrophes or accidents constituting major threats to poor households' survival; the demand for appropriate coping strategies including savings and insurance products is thus widespread and growing, notably in the absence of social protection. Vulnerability can also result from poor decisions about debt. Households can be worse off if they contract debt in unmanageable proportions. Over-indebtedness is a risk that MFIs can respond to with financial literacy training or products that minimise cash changing hands (e.g. leasing); moreover, financial literacy training and similar innovations are likely to produce relatively rapid changes in financial attitudes that could still materialize in the life time of the project.
- Formalisation: a few MFIs induce clients to progressively formalize. It is possible to break down the advancement of poor clients out of the informal economy into several steps and to measure the progression. Again, awareness-raising services of this kind can be expected to lead to tangible and rapid responses on the part of clients.
- Occupational safety and health: workplaces in the informal economy are often unsafe or even dangerous. Accidents at work have been observed to diminish incomes and increase poverty. Responses on the part of clients may take time, but can be observed relatively easily.
- Job creation: one of the most frequent criticisms of microfinance is that it fails to create jobs. It is true, changes in the demand for labour are notoriously difficult to track given the complexity of MFIs' typical household/enterprise client and the substitutions occurring between family and non-family as well as between remunerated and unpaid labour. Yet, this happens to be the decent work aspect that ILO member countries are most keen to learn about. This impact is not always easy to measure; moreover, it is likely to have a substantial leadtime.
- Women empowerment: the evidence on improvements in women's opportunities and gender equality as a result of microfinance programmes is mixed. There are instances where women were able to interact in markets and enhanced their intra-marital position, but there are also cases where women were used to obtain access to a loan without having much say and control over its use. As with job creation the impact on gen-

4 http://www.ilo.org/integration/themes/mdw/lang--en/index.htm.
5 http://www.ilo.org/empelm/what/WCMS_114240/lang--en/index.htm.
6 http://www.ilo.org/ipec/ChildlabourstatisticsSIMPOC/lang--en/index.htm.

der equality is expected to require some gestation period before stable changes in attitudes, values and behaviour can be detected.

3. MF4DW: putting the innovations in place

3.1 Selection of participating MFIs

Most MFIs are convinced that what they are doing is inherently and automatically good for their clients. Most also consider that they have enough to handle, ensuring their financial performance, satisfying their shareholders and donors, and reaching out to an ever growing number of clients. Many MFIs feel that they cannot be expected to go out of their way to take care of their clients. Providing financial services to the poor should be sufficient evidence of their "social" performance. The MF4DW project was targeted at a subset of MFIs that take a more pro-active and committed view of "social" performance, those MFIs that see themselves as genuine social enterprises.

Logically, there had to be a voluntary commitment on the part of participating MFIs, hence a call for proposals was carried out in 2008. Over 60 MFIs responded, out of which 25 were selected based on the following criteria:

- track record of 5 years of operations,
- minimum of 4 branches and at least 10,000 clients,
- a degree of innovativeness, as shown in the range of their products and services on offer,
- established social mission and particular interest in decent work,
- at least positive operational self sufficiency (OSS >100),[7]
- a computerized management information system (MIS), and
- affiliation to a national association of MFIs or an international network.

16 MFI partners eventually continued throughout the entire experimental research until 2012.

3.2 The decent work diagnostic

At the outset of the action research, each participating MFI conducted a diagnostic survey among 200 of its clients to determine their most pressing work-related challenges. At this stage, it was vital to actively involve the MFIs from the board to the client level to get their full commitment and follow through the whole research.

The diagnostic was carried out in 2008 covering 4,748 clients in 22 MFIs. The analysis of the client data showed that child labour, occupational safety and health, formalisation, job creation and productive employment, risk management/over-indebtedness, and women's empowerment, are key challenges for microfinance clients to obtain decent work. *Box 1* summarises the main findings from the diagnostic across all MFIs. Disaggregated by MFI, the results showed a diverse picture of the decent work status of microfinance clients – not only by regions of the world but also by type of institutional set-ups. In addi-

7 Meaning that operational revenue exceeds operational expenses, in a given period and without adjusting for subsidies and inflation. It is the minimal requirement for a MFI that eventually seeks to be self-financing.

tion, a number of MFIs were surprised by the extent of decent work challenges among their clients.

- 90% of clients were self-employed.
- The self-employed clients had created, on average, two jobs in their businesses.
- Majority of clients' employees were paid (cash, in kind or combination).
- Child labour (5-14 years) constituted 5% of total employment created by clients.
- 11% of clients reported dangerous working conditions or injuries.
- 54% of client business activities were informal, 41% paid taxes.
- 8% of clients reported to have borrowed from other sources to repay an MFI loan and 14% had repayment problems.
- 43% of clients reported large unforeseen expense (main reason: accident, illness).

Box 1: Results of the decent work diagnostic in 2008, source: MF4DW: Results of Diagnostic Client Survey, ILO, 2009

The results of the diagnostic also supplied indications why MFIs could be interested in addressing decent work challenges of their clients:

- While following the institution's social mission, MFIs can improve their social performance which in turn might attract increased investment from social investors.
- MFIs have an interest to avoid negative publicity of being associated with, for example, child labour in their clients' businesses or over-indebtedness. The reputational effect can lead to criticism from regulators or donors withdrawing funding. Inversely, reports about socially beneficial initiatives make good public relations and please policymakers and donors.
- Over-indebting clients results in loss of income for the institution itself and cannot be in its own interest. Similarly, if clients suffer from work-related accidents they temporarily cannot work, which impairs their ability to repay loans. Therefore, an increased credit risk has a negative direct impact on the MFI.

Hence, the motivation for MFIs to improve the decent work status of their clients goes beyond their social mission and can be firmly anchored in the institution's business model.

3.3 The microfinance innovations

For the MF4DW, "innovation" was to mean a new practice in its interactions with clients, new or modified products and services, including both financial and non-financial services, changing combinations between products and the manner they are delivered. In the communications with the MFIs about "innovations" the ILO was confronted with a dilemma: on the one hand, it wished to put MFIs' concerns at ease of not being able to satisfy the requirements of an "innovation". To that end the ILO provided MFIs with a list of generic innovations (targeting, marketing, product design and diversification, collateralization, combinations of products, delivery systems and external partnerships). This was to guide and structure the search for innovations. However, several MFIs misunderstood this guidance as an invitation to identify a host of innovations.[8] Allowing a whole bundle of innovations at the same time would have meant an unmanageable research de-

8 One MFI submitted 17 ideas that it wanted to implement simultaneously.

sign as changes in client welfare would no longer be attributable to a single innovation. The ILO therefore had to encourage the MFIs to prioritize and focus on one or two innovations. What made practical sense for the MFI, was not feasible for the research design.

Following a first round of feedback from the MF4DW team, the partner MFIs and ILO technical specialists convened in Geneva in February 2009 to bring out the succinct feature of each decent work innovation. The subsequent fine-tuning of the innovations[9] took from five to 12 months. From late 2009 onwards, the MFIs started implementing their innovations in six areas as shown in *Table 1* below. *Boxes 2* and *3* illustrate two of these innovations in more detail.

Microfinance Institution	Country	Innovation
		Child Labour
Lift Above Poverty Organization (LAPO)	Nigeria	Awareness campaign against child labour; loan for school associated expenses
National Rural Support Programme (NRSP)	Pakistan	Extension of health insurance coverage
Nyésigiso	Mali	Client training on entrepreneurship, financial management, and child labour
		Risk Management & Over-indebtedness
PRIDE Microfinance Limited	Uganda	Leasing product
Negros Women for Tomorrow Foundation (NWTF)	Philippines	Entrepreneurship training for clients
Negros Women for Tomorrow Foundation (NWTF)	Philippines	Emergency savings product
Vision Fund	Cambodia	Financial education for clients
Angkor Mikroheranhvatho (Kampuchea) Co. Ltd (AMK)	Cambodia	Financial education for clients
Tao Yeu May (TYM)	Viet Nam	Client training on risk management and micro-insurance
Financiera Confianza	Peru	Multi-risk microinsurance and client training
Banco Popular	Honduras	Health microinsurance and client training

9 It involved, from the MFI, the project leader who coordinated discussions within the MFI including board, senior management and staff that would be implementing the innovation, and, from the ILO, one social finance specialist plus one topic specialist e.g. on entrepreneurship training or child labour.

Microfinance Institution	Country	Innovation
		Formalization
Fédération des caisses populaires du Burkina (FCPB)	Burkina Faso	Sensitization to benefits of formalization, client training on enterprise management and incentives to formalize
Evangelical Social Action Forum (ESAF)	India	Awareness raising on formalization and Business Development Services
		Occupational Safety and Health
Tamweelcom	Jordan	Awareness raising on occupational safety and health, training on work improvement for small enterprises
Bharatiya Samruddhi Finance (BASIX)	India	Participatory Safety Education
		Job Creation and Women's Empowerment
International Microloan Fund (IMON)	Tajikistan	Entrepreneurship training for female clients
BaiTushum & Partners	Kyrgyzstan	Reorganization of micro and small enterprise finance

Table 1: Innovations implemented through the MF4DW

Addressing Occupational Safety and Health for Improved Safety Performance, Competitiveness and Profit for Client Microenterprises

Tamweelcom, an MFI operating in Jordan, identified productivity and occupational safety and health concerns among its clients as the most pressing challenges through the MF4DW diagnostic. Tamweelcom decided to address these challenges by:

i. Creating an awareness raising campaign for its clients,
ii. Adapting the ILO WISE methodology (Work Improvement for Small Enterprises) to the Jordanian and microfinance context,
iii. Training and capacity building of staff on how to improve working conditions and productivity through Kafâa (training on efficient methodologies),
iv. Training of clients by loan officers, and
v. Implementing a results-based incentive system for staff and clients.

Tamweelcom's trained loan officers built on local practices, encouraged exchanges of experience among clients and promoted their involvement. Tamweelcom expects that the awareness campaign combined with Kafâa training will reduce the total number of occupational hazards, improve productivity, and result in an improved relationship between client and MFI in the long term reflected in rising numbers of clients, in growth of outstanding portfolio and in increased loyalty.

Box 2: MF4DW: example of an innovation that addresses occupational safety and health

School Support Initiative: Access to Soft Loans for Associated School Fees and Anti-Child Labour Awareness Campaign to Reduce Child Labour

LAPO is an MFI in Nigeria that identified child labour as a major challenge for its clients. Therefore, LAPO decided to tackle the issue of child labour with a two-pronged approach:

i. Educating clients on the disadvantages of sending their children to work instead, or in combination with school, and
ii. Offering a financial product (a soft loan for school fees at 1% per month) to help struggling parents to spread the lump sum school payment over a monthly repayment schedule, and to cover other school-related costs, such as uniforms and books. Therefore, it was hoped that the impetus to put children to work to assist in covering the costs of their own school fees would decline with the offer of an education loan product, as well as from exposure to the awareness campaign.

Box 3: MF4DW example of an innovation tackling child labour

These adjustments and modifications are in most instances not dramatic. They do not constitute a radical departure from what the MFIs had been doing over the years. What mattered more from the angle of research design was the plausibility of the link between the innovation and the expected improvement in a given decent work aspect. For that purpose, the ILO and the partner MFI developed a results chain for each innovation. *Table 2* shows an example for a financial education innovation. If these logic chains were persuasively demonstrated, then one could be confident that the MFI would continue with the experimentations even after the end of ILO assistance and cooperation.

3.4 Challenges encountered

Several challenges emerged when putting in place innovations in the context of an action research project, both internal and external to the MFI.

The external environment was not always conducive to action research. Natural disasters like floods in Honduras and Pakistan made it impossible to pursue follow-up surveys with the same clients. Political unrest in Kyrgyzstan meant the elimination of an entire survey area. Regulatory approval required for a new financial product was slow in coming, and this proved especially challenging when the MFI needed to change its legal status.

A challenge internal to an MFI was to reach consensus between board, management and staff on where the action research fit in its institutional strategy. Not everyone understood the longer-term benefits to the MFI. Without a strong agreement and appreciation of these benefits MFI management also hesitated to assign the staff needed to steer the action research within the MFI. The introduction of financial product innovations or modifications to existing products generally progressed smoothly. Innovations took more time if they involved an external partner like an insurance company. This also applied to the leasing innovation, which entailed lengthy negotiations with equipment suppliers. Also, in some instances product uptake proved slower than predicted, for example in microinsurance innovations, which created challenges for data collection and minimum sample sizes. The project was also confronted with one financial product being rolled out to the whole institution which invalidated the research design because it eliminated the differentiation between treatment and control groups.

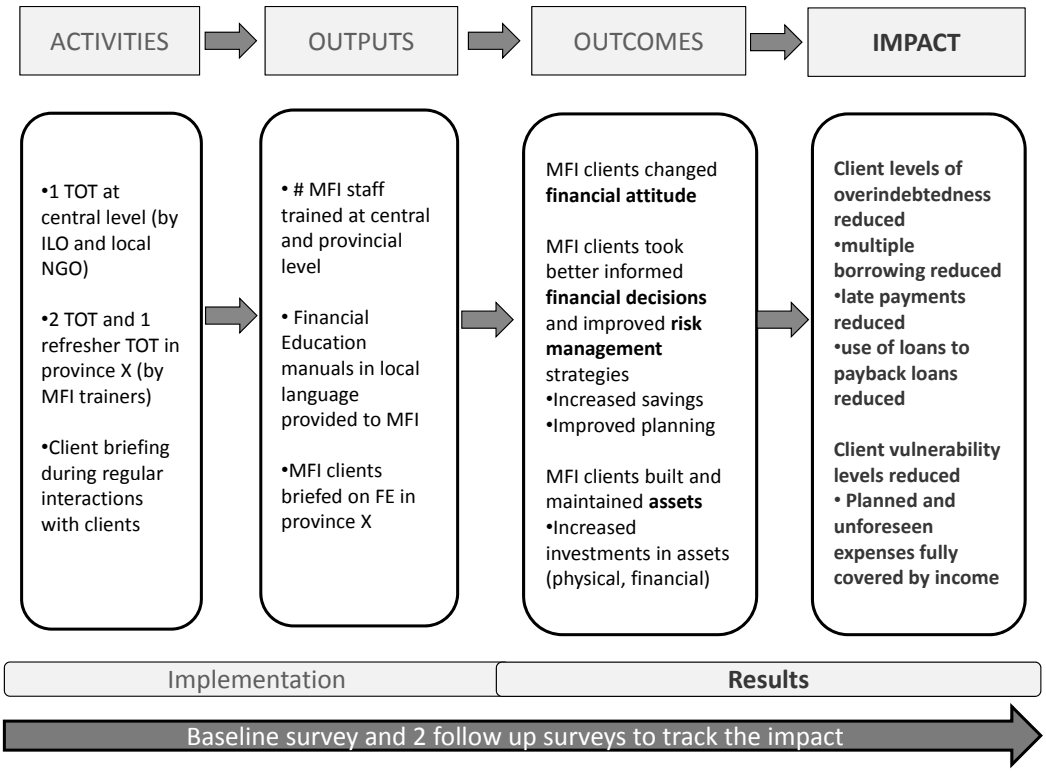

Table 2: Results Chain for a Financial Education innovation (Source: based on ILO and University of Mannheim (2012): Microfinance and risk management. An impact assessment of a financial education programme, AMK Cambodia)

Some innovations are staff intensive, and this posed problems. Training and other non-financial services require qualified staff. Once the decision was taken to either employ MFI staff (training department, loan officers) or external support, it turned out to be delicate to identify the individuals most suitable for the training. In a number of MFIs, high staff turnover or staff rotation eroded the pool of trained staff and re-training was necessary. While loan officers have a close bond to their clients, additional training tasks increased their workload at times. Individual training innovations often suffered from difficulties to attract clients to participate, which sometimes undermined reaching the minimum sample size.

Transmitting new concepts to staff and clients proved difficult for decent work topics such as formalization and child labour – several trainings and re-trainings of staff were needed, as well as exercises with the clients to reach a common understanding.

4. Impact evaluations of microfinance innovations

4.1 The choice of evaluation design

In recent years there has been a rising interest in quantitative impact evaluations, and the MF4DW action research was accompanied by such impact evaluations. In a nutshell, the MF4DW used an experimental research design and implemented innovations with a subgroup of clients (treatment group) while another, comparable group of clients did not have access to the same (control group). Ideally, clients of each group were selected randomly. Before the introduction of the innovations, all clients of the target and control groups were interviewed to establish a baseline against which changes over time could be compared. The design planned for regular follow-up client surveys; data collection was completed in February 2012.

While quantitative impact evaluations impose an additional administrative and financial burden on MFIs, which have to keep treatment and control group separate and collect comprehensive data on clients, there are good reasons to do so. Quantitative impact evaluation creates knowledge about which innovations work for which purpose, and it provides insights about the unintended side effects of innovations. Most importantly, while qualitative evaluations might only provide a rough idea of whether innovations are working, quantitative impact evaluations estimate the magnitude of programme impact. That is, they allow statements on whether there is impact and on how large or small the impact is. Furthermore, given that innovations are typically costly, a well-designed evaluation study involving impact evaluators from the beginning also enables one to compare the costs and benefits of the innovation.

The knowledge on innovations, the magnitude of their impact, their costs and benefits, and their unintended side effects provide solid guidance to decision makers (e.g. policymakers, donors, MFI managers) about how to improve innovative products, or whether to scale them up and replicate them elsewhere. The innovations conducted through the MF4DW project were often implemented in pilot branches rather than the whole branch network. Based on the results from the impact evaluation, MFIs decide whether to introduce these innovations in all branches. Plus, given that taxpayers' money is typically involved, quantitative impact evaluation also shows results for public spending.

The overarching aim of impact evaluations should be to provide international public knowledge on the strengths and weaknesses of certain products and services as well as of the details of their implementation. This may be directly beneficial for the implementing MFI to learn about the impact of its services. The acquired knowledge, however, should also be helpful to other MFIs. Knowledge about the impact of microfinance in one country may not be directly transferable to other countries with different contextual factors or different levels of development. Nevertheless, a variety of impact evaluations conducted in a number of countries will help build a knowledge base that can be used to predict how a product or service might possibly fare in other environments as well.

To increase validity, one would also conduct a large-scale study in several regions of the country, ideally combined with several variations of the innovations. This may often pose serious challenges to the MFIs with limited management and operational capacities. A small scale study with a single pilot branch, where the innovation is implemented, and a single control branch, where the innovation is not (yet) implemented, would provide evaluation results that might be very specific to this single branch and may not enable one to

draw wider conclusions. In particular, it would usually not be justified to predict from this small-scale study the impact a possible country-wide roll out might have. For the latter, one would need a larger study including MFI branches covering most regions of a country and in particular the diversity of the population. Such larger studies would further allow one to analyse whether and by how much impacts differ between rural, urban and semi-urban areas, how impacts vary with the local and regional financial development and industry structure, how impacts vary with household structure or by gender, and so on. Such detailed analyses of the heterogeneity of impacts would also help to better predict the possible impacts of such products if they were adopted by other MFI or in other countries.

While the advantages of quantitative impact evaluations are evident, they are not adopted in many development projects. One reason might be a fundamental trade-off between the practitioners' interest of having projects implemented as smoothly as possible (after all, impact evaluation is not the core purpose of an innovation) and the evaluators' interest in a rigorous evaluation design as a solid basis for credible inference about the impact of the innovation. As a rule of thumb it often turns out that the higher the administrative complexity of the evaluation design (possibly including smaller interventions into the operating business of the MFI), the more credible is the evidence from the quantitative impact evaluation (and vice versa). The typical approach to impact evaluation taken for MF4DW innovations represents a compromise in the sense that, on the one hand, it is administratively less complex than other evaluation designs (in terms of interventions in the operating businesses of MFIs) which, on the other hand, accompanies a reduction in the credibility of the findings.

The evaluation design that maximizes credibility of results is based on the concept of a Randomised Controlled Trial (RCT). In a simple design where a single new product (e.g. a new loan or insurance product) is introduced, the randomised design would foresee that some group of individuals receives access to the new product, while others do not have access to this product. The latter is the control group and its members should be denied access to this product usually for at least one year, better even longer. During the time the control group cannot access the new product, one can examine whether the treatment and control groups evolve differently. If the sample size is large enough, the differential developments between these two groups can be attributed to the impact of the new product. Importantly, though, the impact can only be attributed to the intervention if the treatment and the control groups do not differ systematically in any observed or unobserved characteristics. The latter can only be guaranteed if the assignment to the two groups had been random. That is a random number generator decides who receives access to the innovation and who does not.

In most studies on financial innovations, a cluster-randomization is being used. That means that the level of randomization is the cluster or the branch of the MFI; some branches introduce the innovations while the remaining branches serve as control group. Within a branch, usually all clients are treated equally, i.e. either all have access to the innovation or none.

Impact evaluations on the basis of randomization designs can come in many varieties. One could implement different varieties of the innovations or alternative bundles of innovations in different branches to learn the impact differences between these varieties or bundles. In such a design one could even imagine to proceed without any control group, although one would then only be able to estimate the differential impacts and not the ab-

solute impacts of the innovation. Furthermore, in the absence of full control over the implementing branches, one could alternatively implement various types of randomized encouragement designs, meaning that incentives or encouragements are randomized and the implementation of the product or action which shall be encouraged is voluntary. Evaluations based on randomized encouragement designs, however, are more limited in their interpretation and often lead to imprecise estimates.

The precision of estimates obtained from randomized controlled trials generally improves with the ex-ante balancing of the characteristics of the participating branches. Hence, although a simple assignment to pilot versus control groups would reap the credibility of the design, more complex matched-randomization routines are generally advisable to obtain more accurate estimates.

Sample size is another important consideration that is often underestimated. Particularly, when a cluster-randomization is used, one should not only have a reasonable number of observations within each cluster but also a large number of clusters. Most of the MF4DW innovations have been implemented in one pilot branch, with complementary data being collected in a control branch. Although the pilot branch has often been randomly selected, the total number of branches is only two, and therefore one cannot ensure that these branches are identical in terms of observed and unobserved characteristics. In addition, also after the introduction of the innovation independent shocks, such as unusual weather conditions or a natural disaster, could happen to one branch but not to the other, such that any impact of the innovation would be confound with the impacts of these shocks. A sufficiently large number of clusters would help to make it very likely that any random shocks occurring would be uncorrelated with the innovation status.

Randomisation designs offer yet another advantage over the other approaches discussed below: They can often be much cheaper by being less demanding with respect to data collection. Most of the other approaches require either the collection of panel data or at least the collection of representative household data before the introduction of the innovation as well as after (i.e. at least a baseline and one follow-up data collection are required). With a randomised design of sufficient scale, a full baseline data collection is not needed. Often baseline data from secondary sources is sufficient to compose matched-randomized pairs. Extensive household data collection is then only needed for the follow-up waves.

Randomised trials, on the other hand, often add substantial administrative complexity to the implementation of innovations and might even involve interventions in the operative business of the MFI. For these reasons, the experimental approach to impact evaluation might not always be operationally feasible, warranting the use of other methods. These methodologies include Difference-in-Difference, Instrumental variables, Regression Discontinuities, Propensity Score Matching and Pipeline approaches.[10] These methodologies all have their pros and cons as compared to each other, but they all have disadvantages as compared to RCTs (please see *Khandker et al.* (2009) for a very accessible discussion of these methodologies.)

Given these operational challenges, often a sequential chain of evaluation designs can be observed: A new product is being implemented and tested in a small pilot area with accompanying process evaluations leading to continuous adaptations of the product to weed

10 We should add that these methodologies (including RCTs) are not mutually exclusive. Combinations of methodologies are possible.

out obvious deficiencies in product details, implementation, marketing, and service. Once the new product is well defined, it is often implemented within a small-scale trial with usually few participating MFI branches and few control branches. This is the stage of most impact evaluations of the MF4DW examined here. The small scale limits the use of randomisation, and thus limits the generalizability of results. Nevertheless, if the pilot study delivers promising results, a larger scale study could follow to examine different varieties of the innovation within a wider array of branches in different regions and learn how much impacts depend on contextual factors. Learning about differences in impact is also important to help predict how the product might fare in other environments or other countries.

4.2 MF4DW and difference-in-difference evaluations

The evaluation of the impact of the MF4DW innovations is mainly conducted drawing on the difference-in-difference methodology. In a nutshell, we compare changes in the treatment group before and after the implementation of the innovation to changes in the control group before and after the implementation. Given that the innovation works and the experimental design is clean, we should see a change in outcomes in the treatment group (because there is an innovation) but no change in the control group (because there is no innovation).[11] The difference between the changes in the two groups is then the effect of the innovation on a particular decent work outcome.

The common characteristics of the quantitative impact evaluation of the MF4DW innovation are the following:

- A treatment and a control group,
- Panel data with a baseline collected before the innovation was launched and one to four follow-up surveys collected after the innovation was introduced,
- Two data sources, one being client interviews based on structured questionnaires and the other being administrative data from the MFI's management information system,
- The assignment of the innovation taking place at the branch level (depending on the innovation between 2 to 28 branches, however for most innovations there is only a single treatment and a single control branch),
- And the typical evaluation methodology being difference-in-differences (sometimes supported by Propensity Score Matching).

Irrespective of the methodology chosen for impact evaluation, the project faced a number of challenges in implementing the MF4DW evaluation designs and conducting the empirical analysis. To name a few:

- Staff turnover, especially project manager and research staff (enumerators, encoders),
- Client-specific data in the MIS are often not up to date,
- Client drop-outs, although this varies substantially between MFIs,

11 To be more precise: There may be changes in treatment *and* control branch. However, if they are unrelated to the innovation these changes must be the same for both groups (common trend assumption).

- The MFI's unfamiliarity with basic evaluation concepts:
  - Forming treatment and control groups, and maintaining these groups over the evaluation period,
  - Contamination of control group,
  - and the need to keep all factors constant except the innovation (over the evaluation period),
- Producing random samples based on insufficient MIS data,
- Low data quality, such as missing values and errors in data entry.

Nevertheless, emerging results of the evaluations are promising, and certainly, important lessons were learnt about how to implement pilot-innovations which are to be accompanied by an evaluation of their impact.

5. Lessons learnt so far

Currently we are in the process of finalizing the MF4DW impact evaluation. A number of evaluation reports on individual innovations and a synthesis report will be available toward the end of 2012. The analysis done so far shows promising results for some innovations, but also some unexpected and inconclusive results for other innovations.

To give one example: Financial education has been shown to have a strong impact on financial attitude in one MFI, but mixed impact in another. This proves, on the one hand, that financial education may have very strong impact but, on the other hand, that the MFI's context and specific design and delivery of financial education strongly affects the innovation's success. These findings illustrate the usefulness of having a cluster of related or similar innovations. Comparing across these innovations will thus help to learn more about how the less successful innovation should be adopted and modified to improve its impact.

Several conclusions can be drawn from the evaluation process so far:

i. Allow time to mobilize support and provide the ground for an impact evaluation.
ii. Involve impact evaluators from the very beginning of the programme to get the best possible evaluation design given a MFI-specific context (e.g. for the experimental evaluation design, design of client sample, design of questionnaires).
iii. Encourage MFIs to regularly update their MIS, emphasizing that it is not only important for research collaborations but also a critical data source for the MFI's operations.
iv. Allow sufficient time to convey the basic evaluation concepts and point out the benefits for everybody involved.
v. The implementation of the innovation and its evaluation are medium- and long-term projects. Therefore it is extremely important to identify reliable coordinators in the MFIs to support the projects while they last.
vi. Consider a suitable incentive scheme for interviewed households given the particular context. Sometimes opportunity costs for households are very high (e.g. during the harvesting season). Then a monetary compensation of interviewed households will reduce the amount of missing data (i.e. because clients take more time to answer all questions).
vii. Consider outsourcing the data collection to local research institutes or universities as an option rather than have the MFIs do it. If loans officers have to take care of data

collection on top of their regular work activities, the quality of the interviews suffers (producing missing data). In one case clients felt uncomfortable with loan officers visiting them at home as they were afraid that neighbours might think that they were having repayment difficulties. However, this option needs to be weighed against the benefit of loan officers having intimate knowledge about clients including where to find them and speaking their language.

By sharing our experiences and findings we hope to encourage MFIs' management and staff, microfinance practitioners, and researchers to follow our example and engage in innovations and their evaluation while avoiding our mistakes, because it is a promising way to improve microfinance's impact to the better of all stakeholders involved.

References

Banerjee, A., et al. (2010): The miracle of microfinance? Evidence from a randomized evaluation, J-PAL and MIT.

ILO (2005): ILO policy statement: Microfinance for Decent Work, GB. 94/ESP/3, 294th Session.

ILO (2009): MF4DW: Results of Diagnostic Client Survey, Internal Report.

ILO (2012): MF4DW Closing Workshop, Jordan in June 2012, Conference Report.

Karlan, D./Zinman, J. (2009): Expanding microenterprise credit access: using randomized supply decisions to estimate the impacts in Manila, Working paper, Yale University.

Khandker, S./Koolwal, G./Samad, H. (2009): Handbook on Impact Evaluation: Quantitative Methods and Practices, Washington D.C.

Bernd Balkenhol, Prof. Dr., is Professor at the Faculty of Economic and Social Sciences at the University of Geneva, Switzerland, and former Director of the ILO Social Finance Program.

Markus Frölich, Prof. Dr., is Professor of Econometrics at the University of Mannheim, Germany, Visiting Professor at the University of St. Gallen, Switzerland, and Research Fellow at the Institute for the Study of Labor (IZA) and the Centre for European Economic Research (ZEW).

Niels Kemper, Dr., is Postdoctoral Researcher at the University of Mannheim, Germany.

Patricia Richter, Dr., is working as Technical Officer at the Social Finance Programme of the International Labour Organization in Geneva, Switzerland.

E-mails: bernd.balkenhol@unige.ch, froelich@uni-mannheim.de, niels.kemper@uni-mannheim.de, richter@ilo.org

Acknowledgments: A special note of thanks goes to Craig Churchill, Yousra Hamed and Hillery Midkiff who provided input to writing this article.