Assessing the Impact of Microfinance Programmes: Learning through a Collaborative Project in Pakistan and Zimbabwe

Joana da Silva Afonso

The thesis is submitted in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy of the University of Portsmouth.

July 2018
Declaration

Whilst registered as a candidate for the above degree, I have not been registered for any other research award. The results and conclusions embodied in this thesis are the work of the named candidate and have not been submitted for any other academic award.

Word Count: 84,000
Abstract

The demand for accountability of publicly funded social programmes has increased substantially during the past two decades and, thus, so did the need to evaluate the outcomes of these programmes. Evaluating impact is, however, a complex process and different perspectives have emerged regarding evaluation goals and methodologies. The term ‘impact’ has, therefore, been associated with different concepts, which translate into different approaches in terms of rigour and usefulness, going from formal impact assessments, which privilege methodological rigour, to internal evaluations, which are associated with a learning/improvement perspective. The Lendwithcare (LWC) impact project positioned in between these two approaches and it can be seen as a collaborative project between the project leader (LWC), the academic consultant (University of Portsmouth) and the field partners (Akhuwat Islamic Microfinance and THRIVE Microfinance).

The research conducted within the PhD resulted from my participation in the project as academic consultant. It is based on a case study methodology implemented in two settings (Pakistan and Zimbabwe), using qualitative methodologies, including participation-observation. The analysis of the case studies was based on the utilisation-focused evaluation model developed by Michael Patton, with the choice of the model being rooted in its pragmatic approach and its fit with the conditions of the LWC impact project.

The PhD thesis explores two main research questions. The first deals with the lessons learnt from the implementation of the individual evaluations. The main field challenges and respective solutions were identified, giving attention to the resulting compromises in terms of rigour and the advantages associated with involving the field partners. The second explores the common project approach adopted by LWC in the two settings, identifying its main elements and how these have influenced the perception of the partners regarding their participation in the evaluations. The intention of all partners to continue the project as well as the interest of other microfinance institutions to implement similar evaluations seems to indicate that LWC and its partners have perceived relevant advantages in this approach.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>4</td>
</tr>
<tr>
<td>List of Tables</td>
<td>8</td>
</tr>
<tr>
<td>List of Figures</td>
<td>8</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>9</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>10</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>11</td>
</tr>
<tr>
<td>1.1 Research aims and context</td>
<td>15</td>
</tr>
<tr>
<td>1.2 Microfinance: a brief introduction</td>
<td>19</td>
</tr>
<tr>
<td>1.3 Microfinance today: the accountability drive and evaluation</td>
<td>21</td>
</tr>
<tr>
<td>1.4 Research design</td>
<td>26</td>
</tr>
<tr>
<td>1.5 Research main contributions and originality</td>
<td>27</td>
</tr>
<tr>
<td>1.6 Organization of the thesis</td>
<td>30</td>
</tr>
<tr>
<td>2. Outcomes and Impact Evaluation</td>
<td>32</td>
</tr>
<tr>
<td>2.1 Programme evaluation: clarifying concepts and perspectives</td>
<td>32</td>
</tr>
<tr>
<td>2.1.1 Outputs and Outreach</td>
<td>33</td>
</tr>
<tr>
<td>2.1.2 Outcomes and Impact</td>
<td>34</td>
</tr>
<tr>
<td>2.2 Overview of evaluation theories</td>
<td>38</td>
</tr>
<tr>
<td>2.2.1 An evaluation theory tree</td>
<td>38</td>
</tr>
<tr>
<td>2.2.2 Evaluation theories: the methods branch</td>
<td>39</td>
</tr>
<tr>
<td>2.2.3 Evaluation theories: the value branch</td>
<td>41</td>
</tr>
<tr>
<td>2.2.4 Evaluation theories: the use branch</td>
<td>43</td>
</tr>
<tr>
<td>2.3 Evaluation models and approaches</td>
<td>45</td>
</tr>
<tr>
<td>2.3.1 Context, Input, Process, Product model (CIPP)</td>
<td>45</td>
</tr>
<tr>
<td>2.3.2 Utilisation-focused model (UFE)</td>
<td>47</td>
</tr>
<tr>
<td>2.3.3 Constructivist model</td>
<td>49</td>
</tr>
<tr>
<td>2.3.4 Other evaluation models</td>
<td>49</td>
</tr>
<tr>
<td>3. Evaluation in Microfinance</td>
<td>51</td>
</tr>
<tr>
<td>3.1 Outreach, outcomes and impact of microcredit: a theoretical approach</td>
<td>51</td>
</tr>
<tr>
<td>3.1.1 Microcredit expected outcomes</td>
<td>52</td>
</tr>
<tr>
<td>3.2 Outreach, outcomes and impact of microcredit: the empirical research</td>
<td>54</td>
</tr>
<tr>
<td>3.2.1 Systematic reviews and reports</td>
<td>56</td>
</tr>
<tr>
<td>3.2.2 Methodological choices</td>
<td>59</td>
</tr>
<tr>
<td>3.2.3 Mixed results</td>
<td>61</td>
</tr>
<tr>
<td>3.3 Challenges in the implementation of evaluations in microfinance</td>
<td>66</td>
</tr>
</tbody>
</table>
Poverty Scorecards in Microfinance Evaluation

4.1 Poverty measurement in microfinance
4.2 Poverty scorecards in microfinance
4.3 Poverty Probability Index (PPI)
  4.3.1 Methodological issues and implementation challenges
  4.3.2 Empirical evidence

Methodological Issues

5.1 Case Study
5.2 Data Collection: the central role of participant observation
5.3 Data Analysis: a utilization-focused framework
  5.3.1 The choice of evaluation model
  5.3.2 The UFE-based framework of analysis
5.4 The LWC assessment project: an overview
  5.4.1 Lendwithcare: the crowdfunding platform
  5.4.2 Lendwithcare Impact Assessment Project

Case Study: Akhuwat Islamic Microfinance, Pakistan

6.1 Macro-Context
6.2 Micro-context
  6.2.1 AIM and Islamic Microfinance
  6.2.2 Microcredit programme: Family Enterprise Loan
6.3 LWC Evaluation Project
  6.3.1 Introduction
  6.3.2 Design of the Evaluation
    6.3.2.1 Primary Intended Users: motivations
    6.3.2.2 Role of Evaluator
    6.3.2.3 Household Survey: preparation
  6.3.3 Implementation Process
    6.3.3.1 Primary Intended Users: behaviours
    6.3.3.2 Role of Evaluator
    6.3.3.3 Household Survey: implementation and analysis
  6.3.4 Interpretation and Use of Findings
    6.3.4.1 Primary Intended Users: perceptions and intentions
    6.3.4.2 Role of Evaluator
    6.3.4.3 Household Survey: results and reports
6.4 Discussion
9.2 The parallel with utilisation-focused evaluations
9.3 Contributions towards transparency and credibility
9.4 Strengths and weaknesses of the research
9.5 Final considerations

References

Annexes
  List of Annexes
List of Tables

5.1 Interviews: Perceptions on the Evaluation Project 95
5.2 UFE-Based Framework of Analysis 106
5.3 Lendwithcare Portfolio 109
6.1 Pakistan Financial Inclusion Data 119
6.2 Pakistan Microfinance Sector (2015-2017) 120
6.3 AIM Main Indicators 122
6.4 LWC/AIM Survey Respondents by Geographical Area 134
6.5 Client Average Monthly Income and Expenses (AIM Survey) 142
6.6 Quantile Regression Results – AIM 143
6.7 Mann Whitney Tests for Gender and Educational Level 144
7.1 Zimbabwe Financial Inclusion Data 155
7.2 LWC/THRIVE Survey Respondents by Geographical Area 172
7.3 Average Monthly Income and Expenses (THRIVE Survey) 175
7.4 Quantile Regression Results – THRIVE 176

List of Figures

2.1 Programme Evaluation Chain 33
2.2 Evaluation Theory Tree 39
3.1 The Virtuous Cycle of Credit 52
4.1 PPI Form for Pakistan (English version) 78
5.1 Lendwithcare Crowdfunding Model 110
6.2 LWC/AIM Project Milestones 125
6.3 LWC/AIM Evaluation Stakeholders Map 126
7.1 THRIVE Lending Process 160
7.2 LWC/THRIVE Project Milestones 162
7.3 LWC/THRIVE Evaluation Stakeholders Map 163
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIM</td>
<td>Akhuwat Islamic Microfinance</td>
</tr>
<tr>
<td>AIMS</td>
<td>Assessing Impact of Microenterprise Programmes</td>
</tr>
<tr>
<td>ANDC</td>
<td>Associação Nacional de Direito ao Crédito</td>
</tr>
<tr>
<td>BRAC</td>
<td>Bangladesh Rural Advancement Committee</td>
</tr>
<tr>
<td>BRDB</td>
<td>Bangladesh Rural Development Board</td>
</tr>
<tr>
<td>CFI</td>
<td>Centre for Financial Inclusion</td>
</tr>
<tr>
<td>CIPP</td>
<td>Context, Input, Process, Product</td>
</tr>
<tr>
<td>DFID</td>
<td>Department for International Development</td>
</tr>
<tr>
<td>e-MFP</td>
<td>European Microfinance Platform</td>
</tr>
<tr>
<td>CGAP</td>
<td>Consultative Group to Assist the Poor</td>
</tr>
<tr>
<td>GDP</td>
<td>Gross Domestic Product</td>
</tr>
<tr>
<td>GRO</td>
<td>Group Relations Officer</td>
</tr>
<tr>
<td>IFAD</td>
<td>International Fund for Agricultural Development</td>
</tr>
<tr>
<td>IFPRI</td>
<td>International Food Policy Research Institute</td>
</tr>
<tr>
<td>ILO</td>
<td>International Labour Organisation</td>
</tr>
<tr>
<td>IPA</td>
<td>Innovations for Poverty Action</td>
</tr>
<tr>
<td>IRDB</td>
<td>Inter-American Development Bank</td>
</tr>
<tr>
<td>JCSEE</td>
<td>Joint Committee on Standards for Educational Evaluation</td>
</tr>
<tr>
<td>J-PAL</td>
<td>Abdul Latif Jameel Poverty Action Lab</td>
</tr>
<tr>
<td>LIC</td>
<td>Low Income Countries</td>
</tr>
<tr>
<td>LMIC</td>
<td>Lower Middle-Income Countries</td>
</tr>
<tr>
<td>LWC</td>
<td>Lendwithcare</td>
</tr>
<tr>
<td>MFI</td>
<td>Microfinance Institution</td>
</tr>
<tr>
<td>MIS</td>
<td>Management and Information System</td>
</tr>
<tr>
<td>NBFI</td>
<td>Non-Banking Financial Institution</td>
</tr>
<tr>
<td>NGO</td>
<td>Non-Governmental Organisation</td>
</tr>
<tr>
<td>ODA</td>
<td>Overseas Development Administrator</td>
</tr>
<tr>
<td>OECD</td>
<td>Organisation for Economic Co-operation and Development</td>
</tr>
<tr>
<td>PAK</td>
<td>Pakistan</td>
</tr>
<tr>
<td>PAT</td>
<td>Poverty Assessment Tool</td>
</tr>
<tr>
<td>p.c.</td>
<td>Personal Communication(s)</td>
</tr>
<tr>
<td>PICES</td>
<td>Poverty, Income, Consumption and Expenditure Survey</td>
</tr>
<tr>
<td>PIU</td>
<td>Primary Intended Users</td>
</tr>
<tr>
<td>PKR</td>
<td>Pakistan Rupee</td>
</tr>
<tr>
<td>PMN</td>
<td>Pakistan Microfinance Network</td>
</tr>
</tbody>
</table>
Acknowledgements:

This thesis is the direct result of the past three and a half years at the University of Portsmouth, but it started being prepared many years before when in 2005 almost by accident I started working at a small Portuguese NGO (ANDC) as a microcredit officer. Since then many people contributed to my path in the microfinance sector, and although it is difficult to name all, there are a few persons who I would like to thank for helping me taking decisions and giving me strength to stick to them, namely, José Centeio, Fátima Belo, Marta Mucha, Isabelle Guérin, Solène Morvant-Roux, Fei-Ling Tan, Eliana Linares-Gonzalez, Yi-Ling Lai and Andreea Molnar. I would also like to thank to my supervisors, Joe Cox and Andy Thorpe for their support during the PhD project and to the great teams I had the opportunity to work with at Lendwithcare, Akhuwat Islamic Microfinance and THRIVE Microfinance. A very special thanks to Michael O’Connor, with whom I had the pleasure to organise the 5th European Research Conference on Microfinance and whose support was fundamental in the most challenging moments of the PhD, and to the two examiners for their valuable comments. Finally, I would like to dedicate the thesis to my parents, Rosa and Ernesto, thanking them for all their patience and support.
1. Introduction

Microfinance and financial inclusion, as other areas in the development sector, share a common overall goal of improving the lives of poor and vulnerable populations in developing countries. The promoters of these interventions share also an increasing need to evaluate their work in order to provide evidence of the results attained through the programmes implemented, particularly those affecting the target populations (Gertler, Martínez, Premand, Rawlings & Vermeersch, 2016; Liket, Rey-Garcia & Maas, 2014). This accountability drive is particularly important in the field of microfinance considering the substantial funding which has been pouring into the sector in the form of donations or subsidised credit from donors and lenders, which include national aid agencies, the World Bank and other development banks, social investors, individual donors and lenders (Morduch, 2000).

The growing attention given to the assessment of microfinance impact was boosted in the past decade by the emergence of delinquency crises in several microfinance markets across the world. These crises, often linked to client over-indebtedness, triggered a reaction from the main actors in the sector and seem at the moment to have been surmounted (Chen, Rasmussen & Reille, 2010; Guérin, Labie & Servet, 2015). Nevertheless, they have casted doubts on the merits and the robustness of the sector and motivated donors, investors and the microfinance institutions (MFIs) themselves to give more attention to the evaluation of the microfinance programmes.

The assessment of the outcomes and impact of microfinance programmes is a complex task, with different actors adopting different perspectives regarding the meaning, the processes and the methodologies applied. Rajbanshi, Huang and Wydick (2015) explored the different results communicated by practitioners and academics in regard to impact assessments of microfinance programmes employing experimental methodologies. Sinha (2017a) referred to the various uses given to the term impact and its distinction from outcomes and outreach, calling attention to how microfinance institutions and social investors have often based their claims regarding impact on outreach statistics and anecdotal evidence. Equally, Haase (2013) distinguished the different approaches of academics and practitioners towards the impact of microfinance, suggesting that the latter usually do not have the time and resources required to produce rigorous research, hence, practitioner literature on the topic should not be assessed to “the same standards as the academic works”. Nonetheless, the author considered important to analyse the implicit values and assumptions associated with practitioner literature as they allow for identifying the “fundamental dilemmas of assessing impact” (p.4).
The suggestion of a perspective gap between academics and practitioners is not limited to the microfinance and financial inclusion sectors. In the same direction, Christie (2003) and House (2003) referred to distinctions between evaluation theorists and practitioners, and Bartunek and Rynes (2014) addressed the paradoxes of the relationships between academics and practitioners in the management field. These authors identified a trend in the sector literature to lay emphasis on the differences between the two sides and the feasibility and relevance of ‘bridging’ this gap. Their own perspective was that this might not be the most useful question to address. Rather than trying to bring together the two fields, they suggested that the focus should be on how the underlying tensions between academics and practitioners can benefit academic research.

Independently of the field, it should be noted that both academics and practitioners are not homogeneous categories. There are different kinds of academics (adopting different epistemological approaches) and different kinds of practitioners (more or less reflective regarding the results of their programmes). Therefore, the utilisation of the terms ‘academics’ and ‘practitioners’ should consider this caveat that not all academics and not all practitioners fit into the dichotomy presented.

Nonetheless, looking into the potential differences and tensions between academics and practitioners from an evaluation perspective is relevant in the context of increased accountability of the social programmes funded by external sources (public or private). It is expected that these tensions may translate into, on the one hand, practitioners not always being motivated to participate in research projects or sophisticated evaluations and looking suspiciously to research results, and on the other hand, academics maintaining the distances from the field actors and their active participation in the evaluations. Although it is important to develop both academic research and purely internal evaluations, exercises which produce different types of evidence, there is also space for collaborative projects as these projects, conciliating different mindsets, objectives and languages, have the potential to improve science and practice in the sector.

One of the objectives of this PhD thesis is, thus, to describe and analyse the implementation of a collaborative project which aimed to evaluate the social impact of the microfinance programmes funded by an external social investor. Beyond providing insights into the implementation processes of the two evaluation projects conducted, contributing to promoting transparency and accountability in the sector, the thesis gives also a unique perspective into the different actors learning process throughout the impact project. From a broader perspective, the research aims to contribute to the sector knowledge on impact
evaluation in the expectation that this may help improving the evaluations implemented, the programmes and the institutions, ultimately benefiting the intended target populations.

The focus of the thesis on the challenges associated with the field implementation of the evaluations derived, in part, from the less frequent attention given to this issue and its consequences on the impact of the microfinance programmes in the sector dominant literature, exceptions including Adams and Vogel (2013), Bolnick and Nelson (1990) and Gaile and Foster (1996). The discussion of the merits and limitations of different impact methodologies in the microfinance literature have been mainly centred on the ability of the different approaches to deal with self-selection and placement bias, complex issues given the characteristics of the microfinance programmes, and to cope with the fungibility of money, which makes it difficult to trace the effective use of the financial services provided by the MFIs (Duvendack et al., 2011; Khandker, 2013; Odell, 2010; Vaessen et al., 2014; Von Pischke & Adams, 1980). The overall conclusion is that experimental methodologies are preferred to deal with these methodological problems, but they are not applicable to all situations, with alternatives such as quasi-experimental methodologies, qualitative methodologies and mixed methods being advocated by several authors (Balkenhol, 2012; Copestake, 2012; Copestake, Dawson, Fanning, McKay and Wright-Revolledo, 2005; ILO, 2015; Odell, 2015).

The broader approach to impact and social performance assessment in the thesis is not limited to the different methodological perspectives found in the microfinance literature, with attention being given to theoretical ideas and models of evaluation theorists, many of them developed within the educational sector but applicable to other social sectors and programmes. Further discussion of the different perspectives on evaluation and impact assessment can be found in chapters two and three.

The PhD research was developed based on my participation in the evaluation project promoted by Lendwithcare (LWC), a UK-based crowdfunding platform for microfinance launched by CARE International UK. The primary objective of the LWC project was assessing the social ‘impact’ of the microcredit programmes implemented by the LWC funded MFIs. The project marked a new stage in LWC work with its partners, in which the implementation of social performance assessment at the MFIs level was deemed as important as obtaining credible impact data on the microcredit programmes.

Giving the learning process associated with the impact project, LWC chose to kick-start the process with partners which demonstrated strong interest in participating but had none or incipient experience in social performance assessment; that showed capacity to provide the necessary logistic support to the implementation of the household surveys and presented
local evaluation teams able to communicate easily (English speaking countries). The selection of the field partners included also a diversity criterion, exploring two very different institutions in two highly complex and distinct contexts. The challenges faced during the design and implementation of the impact project in the two settings, and the respective solutions, were to be the basis for LWC to decide on the merits and limitations of the project and the possibilities of its extension to other field partners in different parts of the world.

The first partner selected was Akhuwat Islamic Microfinance (AIM). AIM is a well-established MFI in Pakistan, where it has been providing *qard-hasan* (interest-free) loans since 2001. The growth of the institution was, however, particularly strong during the implementation of the project with AIM becoming the largest microfinance institution in Pakistan in terms of number of active borrowers, and one of the largest Islamic microfinance institutions worldwide. The second MFI selected was THRIVE Microfinance in Zimbabwe. In contrast with AIM, THRIVE is a small and young institution, founded by foreign social investors in 2012 when the hyperinflation crisis, which seriously affected the microfinance sector, had been overcome.

An important distinction between the two institutions respects their relationship with LWC, the commissioner of the impact project. Funding from LWC had little significance in the funding sources of AIM, representing less than 2% of the gross loan portfolio of the institution, which relied primarily on donations from local governments and private individuals and organisations in Pakistan. Conversely, THRIVE funding depended significantly on external investors, particularly the two crowdfunding platforms, KIVA and LWC. This distinction is relevant in the context of the implementation of the impact project, reflecting both on the motivations for participating in the impact project and in the power relations between the partners involved in the project. The two MFIs are presented in detail in chapters six and seven.

As a final introductory note, it should be stressed that for the PhD research, the participation in the LWC impact project implied the adaptation of the objectives, research questions and methods to the conditions of the project and to what was feasible to accomplish in the timeline of the PhD.

In the first conversations between LWC and the University of Portsmouth (UoP), the perspective of LWC was to implement the evaluation in one country and one MFI, which had experience in the use of the poverty assessment tool PPI (Poverty Probability Index). Therefore, the PhD initial research proposal focused on the field challenges of the evaluation
in one setting and the validity of PPI as evaluation tool for smaller MFIs, exploring the identified gap in the literature regarding independent assessments of the use of PPI.

However, LWC initial plan for the project changed and the assessment project was implemented in two different MFIs, neither of them with experience implementing PPI (in one of the cases, there was no PPI developed for the country). There was, thus, the need to re-think the research questions and re-shape the research design and the methodologies applied. This was a challenge for me as researcher, since I had no or little experience applying some of these methodologies and had also little knowledge about the two countries where the project was implemented, but it represented as well an excellent learning opportunity.

In this introductory chapter, the first section (1.1) presents the main aims and the research questions addressed in the thesis. The following sections (1.2 and 1.3) introduce the microfinance sector, the relevance of accountability in the sector and the different perspectives regarding evaluation in the sector, setting up the context for the research implemented within the PhD project. Section 1.4 describes the research design and the methodologies employed and summarises their advantages and limitations. Section 1.5 highlights the main contributions of the thesis and the original elements of the research. Finally, section 1.6 presents the structure of the thesis.

1.1 Research aims and context

The main aims of the research are to better understand the challenges associated with the implementation of impact evaluations in the microfinance sector, particularly for socially focused MFIs, as well as to assess how the adoption of a common project approach in the implementation of evaluations in different contexts can contribute to the engagement of the microfinance institutions and influence the perceived value of the evaluations conducted in the sector. These broad questions are particularly relevant where the financial resources required to implement more sophisticated impact assessment methodologies, such as randomised control trials (RCTs), are not available.

The topic of evaluation has been widely explored from a theoretical and empirical perspective, and literature on evaluation theories and models can be easily found. However, given its origins, much of this literature addresses evaluation in education (e.g. Cook, 2002; Madaus, Scriven and Stufflebeam, 1983; Stufflebeam, 1971). Much less academic research deals with the application of different evaluation models to microfinance. The impact evaluation literature in the microfinance sector, including academic papers, practitioner reports and literature reviews (with very few exceptions) do not make explicit (or implicit)
reference to the evaluation theories and models used, focusing instead on the design of the impact studies and the methodologies used in data analysis.

These are issues presented as fundamental to ensure the internal validity of the findings and the credibility of the research developed in this area (Khandker et al., 2010). However, they are not the only factors to take into account. Among the main conclusions of the systematic review on the impact of microfinance conducted by Duvendack et al. (2011, p.3) is that most studies are based on weak data and that these deficiencies cannot always be corrected by sophisticated econometric techniques. Equally, Liket et al. (2014) include poor data collection among the potential causes of implementation failures in the evaluation of development programmes. If data quality is fundamental, then the choice of data collection methods and its implementation in the field should not be neglected as they strongly influence the results. However, these concerns are frequently absent from much of the microfinance literature.

This omission is not completely surprising considering that it represents a ‘dive’ into the complexity of the programmes and their implementation, which from the researchers’ viewpoint, especially if adopting a positivist perspective, is hardly desirable. It also implies a close collaboration between practitioners and researchers in the development of the evaluation, which in the case of this PhD was made possible through the active participation of the researcher in the assessment of two MFIs funded by LENWITHCARE, Akhuwat Islamic Microfinance (AIM) in Pakistan and THRIVE Microfinance (THRIVE) in Zimbabwe. However, this is not always the case in the evaluations conducted in microfinance and other development sectors.

Beyond an in-depth look into the field implementation challenges of evaluation in microfinance, the PhD addresses another gap in the microfinance literature regarding the use of poverty scorecards in the evaluation process. As further described below, the research on the application of poverty scorecards in MFIs, especially in the last decade, has been largely conducted by the promoters of the Poverty Probability Index (PPI), a poverty assessment tool initially developed by the Grameen Foundation. At the time of the research, few independent studies have addressed both the implementation challenges and the results obtained from the use of poverty scorecards in the sector.

The conditions of implementation of the two evaluations conducted, in which the Pakistani MFI was able to use PPI, but the tool was not available to the Zimbabwean MFI (leading the institution to develop their own poverty scorecard), allowed for exploring the topic from an innovative perspective. In doing so, the research has focused on the (opposite) motivations, behaviours and perceptions of the managers of the two MFIs towards this type of instrument.
In this way, the first research question, and sub-set of questions, addressed by the PhD project are:

1) **What are the main challenges associated with outcomes and impact evaluation in socially focused MFIs with limited resources?**

   1.1. **What are the motivations of different stakeholders regarding the design and implementation of an evaluation process?**

   1.2. **What are the main challenges in the field implementation of an evaluation in microfinance?**

   1.3. **Can poverty scorecards, such as the Poverty Probability Index (PPI), be a useful evaluation tool for MFIs?**

   1.4. **How do different stakeholders perceive and use the evaluation results?**

These questions are explored independently in the two case studies developed in Pakistan and Zimbabwe using as reference the utilisation-focused evaluation (UFE) model developed by Michael Patton (Patton, 2008). The main features of the model are described in chapter two, with the reasons for its selection in the thesis, mostly associated with its pragmatic roots, being presented in chapter five (methodologies).

Although the findings related to this first research question are interesting in their own right, especially considering that the evaluation conducted in the Pakistani MFI is the first employing a longitudinal approach into an Islamic microfinance programme, the fact that the LWC evaluations have included a set of common elements to the two settings opened up new opportunities for the research. Therefore, the second set of questions explored in the thesis is:

2) **How does a common project approach to evaluation influence the MFIs in the adoption of evaluation as a component of a social performance management system, aiming to improve the programmes developed and generate credible evidence of their results?**

   2.1. **To which extent is it feasible to implement a common project approach to evaluation considering the diversity and complexity of the contexts in which microcredit programmes are implemented?**

   2.2. **What are the advantages and limitations of adopting a common project in the implementation of evaluations in different contexts?**
2.3. *How can the application of a common project approach to evaluation contribute to the valorisation of the evaluations and the research conducted in the microfinance sector regarding its impact?*

It is important to emphasize that a common project approach to evaluation means that similar, not equal, evaluation design and methodologies were applied in the two MFIs, with the necessary adaptations to accommodate the differences between the institutions and the contexts in which they operate. Therefore, while similarities and differences associated with the processes are appraised in the thesis, there is no pretension to generalise the evaluation findings obtained for each MFI. The analysis and discussion of the common project approach to evaluation is not centred in the results, which are only partially referred in the thesis as starting point to explore the research questions.

By providing in-depth analysis of two evaluations conducted in different contexts, the research aims to offer an additional contribution to the knowledge on a heterogeneous and changing sector, in line with the plural view of the microfinance sector advocated by Copestake et al. (2016). This plurality, however, should not be seen as an obstacle to the development of a common culture regarding evaluation, social performance management and accountability, which values and encourages the efforts of all institutions involved in the sector.

The answers to this second research question and associated sub-questions are relevant to a wide range of stakeholders of the evaluation project, internal and external, being of particular interest to microfinance funders and investors with diversified portfolios, and for the sector networks working with a diversity of institutions in distinct settings.

In addition to exploring these two main research questions, which can be associated with an empirical view of evaluation in microfinance, the research has a secondary goal linked to closing the identified gap between the empirical microfinance impact literature and the more theoretical evaluation literature. These two bodies of literature seem to be disconnected, despite the existence of synergies between them.

For this purpose, the analysis of the two case studies was based on the principles and concepts of a theoretical evaluation model, in this case the Utilisation-Focused Evaluation (UFE) model. Thus, a third research question addressed in the thesis can be expressed as:

3) *How pertinent is the use of evaluation theories/models, and particularly of the UFE model, to develop the analysis of evaluation in the microfinance sector?*
In this introductory chapter, and before presenting the research design and methodologies used in the PhD, it is important to contextualise the research, briefly describing the origins and main features of the microfinance sector and of the evaluation conducted in the sector.

1.2 Microfinance: a brief introduction

Microfinance can be defined as the provision of financial products and services to low income populations otherwise excluded from the access to formal financial services (Center for Financial Inclusion, 2015). The sector has grown from small credit experiences developed in a few developing countries in the 1970s to a complex sector characterised by a diversity of actors, missions, products and clients (Armendariz & Morduch, 2010; Ledgerwood & Gibson, 2013).

By the end of 2013, the 3,098 microfinance institutions identified in the last Microcredit Summit Campaign Report, claim to have reached 211 million clients, with a significant part of these clients being microcredit borrowers taking loans to start or develop a business (Reed, 2015). According to the report, 54% of the borrowers belonged to the 40% poorest segment of the respective country’s population at the time of the first loan, and they were mostly women (82.6% of the poorer). The majority of these clients were concentrated in the Asia and the Pacific (79%), where India and Bangladesh were (are) the largest microfinance markets.

These figures can be considered as impressive, but they still display a significant gap relatively to the financial inclusion statistics provided by the World Bank through the Global Findex Survey. In the 2017 survey, the number of adults without access to formal financial services, including bank accounts, was of 1.7 billion ((Demirgüç-Kunt, Klapper, Singer, Ansar & Hess, 2018).

The positive association between financial inclusion and economic development has been accepted by most researchers and policy makers and it justifies the growth and continuous attraction of local and international funds for the microfinance sector (Beck, Demirgüç-Kunt & Levine, 2007a; Karlan & Morduch, 2010; Littlefield, Morduch & Hashemi, 2003). This support continues despite the delinquency crises associated with diverse microcredit markets such as Bosnia-Herzegovina, Morocco, Nicaragua and Andhra Pradesh (Chen et al.

---

1. Experiences in Bangladesh (Grameen Bank and BRAC), Indonesia (Bank Rakyat), Bolivia (FINCA) and Brazil (ACCION) are identified as pioneers in the sector (Armendariz & Morduch, 2010). However, it is important to note that the concept origins are much older. Hollis and Sweetman (1998) identify as predecessors of modern microfinance the experiences of English lending charities in the 18th century, and the Raiffeisen credit cooperatives developed in Germany and replicated in other European countries in the 19th century.
2010; Guérin et al., 2015; Rozas, 2013) and the academic debate surrounding the (lack of) impact of these programmes on poverty reduction (Banerjee, Karlan & Zinman, 2015; Chowdhuri, 2009; Duvendack et al., 2011; Morduch, 1998).

There also seems to be little doubt that poor people need access to financial services that are adequate, flexible and reliable, helping them to cope with the “triple whammy” presented by Collins, Morduch, Rutherford and Ruthven (2009): low and irregular incomes, and the lack of adapted financial instruments. This reasoning is associated with a recognition of the inability and general unwillingness of commercial banks to deal with these potential clients as well as the belief that informal finance, widespread in many developing countries (Guérin, Morvant-Roux & Servet, 2011), is not in most cases the best solution for the clients (Armendáriz & Morduch, 2010; Karlan & Murdoch, 2010).

The microfinance movement has allowed low income populations to access formal financial services by minimizing information asymmetries and cost problems through a set of innovative mechanisms such as group lending with joint liability and dynamic incentives (Mersland & Strøm, 2011; Morduch, 1999). For the first decades, the evolution of the sector was based on a principle: poor people when provided with productive loans that allow them to create or develop their own businesses can increase their income, improve their household well-being and repay their debt (Hulme, 2000a; Sengupta & Aubuchon, 2008). Thus, the growth of microcredit worldwide has been grounded on the virtuous cycle of credit (Hulme & Mosley, 1996a), and has associated an inherent ethical validation based on the public targeted, the poor; the focus on credit, linked to responsibility and dignity; and the financial margins, lower interest rates when compared with informal financial providers (Hudon, 2011). Credit to small businesses is still today the core product and main asset of most MFIs, including the two institutions studied in this thesis, and it is also the focus of many international funders/donors as Lendwithcare (LWC) exemplifies. It is, however, no longer the only microfinance product offered, and several studies have questioned the role of productive credit as the most important financial service for the poor (Collins et al., 2009; Dichter, 1996; Robinson, 2001; Rutherford, Collins & Johnson, 2013; Vanroose, 2007).

The recognition of the value of other financial services for low income populations translated into a shift from microcredit to microfinance that goes beyond purely semantic. It entails a broader concept that includes business loans but also other loans, savings, insurance, remittances and mobile money (Mader & Sabrow, 2015). The scope of products/services offered by MFIs is gradually enlarging by incorporating new products and innovative processes and technologies. This expansion relates not merely to financial services but
includes non-financial services such as training, consultancy, education and health services in what is named as ‘microfinance plus’ (Maes & Foose, 2006).

Heterogeneity in the sector is not confined to the mix of products/services. There are also different publics (rural/urban, women, poor entrepreneurs, poor with children, low income populations, self-employed, etc.); missions (financial inclusion, enhance entrepreneurship, improve household income, poverty reduction, etc.); organizational structures (NGOs, for-profit companies, non-regulated financial institutions, Banks, etc.); and funding sources (donations, subsidized loans, commercial loans, equity; external, internal) (Ledgerwood & Gibson, 2013). Diversity makes it difficult to generalize conclusions for the sector, and it adds to the uncertainty associated with the outcomes of development interventions, often unforeseen and not necessarily positive (Lewis & Mosse, 2006).

1.3 Microfinance today: the accountability drive and evaluation

Research, particularly on savings, has highlighted that access and use of formal financial services are distinct concepts and have different consequences on the financial inclusion of poor populations (Allen, Demirgüç-Kunt, Klapper & Martinez-Peria, 2012; Beck, Demirgüç-Kunt & Martinez-Peria, 2007b). The two concepts are closely linked to the shift from microfinance - ensuring access to formal financial services, to financial inclusion, which entails guaranteeing that the access to the services translates into effective use and positive change in the lives of the clients. Nonetheless, some authors question if this change in name corresponds to a real transformation or a ‘cosmetic’ operation to stem criticisms to the sector driven by the delinquency crises (Mader & Sabrow, 2015).

Studies on microcredit delinquency crises and client over-indebtedness, including Chen et al. (2010), Gonzalez (2008), Guérin et al. (2015), Guérin, Morvant-Roux and Villareal (2014) and Schicks (2013), have shown that the interaction of demand, supply and environment factors, many of these out of the control of the institutions, can lead to social outcomes opposite to those expected. Morvant-Roux, Afonso, Forcella and Guérin (2015, p.93), studying the efficiency of over-indebtedness prevention strategies in the Dominican Republic, suggested that this might be true even when MFIs adopt the recommended practices advocated by international organisations such as the CGAP and the Smart Campaign. Therefore, there is a recognized need to continuously monitor at field level how the programmes are being implemented and appropriated by all stakeholders, identifying the real (and not the intended) outcomes and assessing both the financial and social performance of the programmes (Murisa & Chikweche, 2013, p.5).
The need for more legitimacy and, hence, the pressure for increased transparency and accountability, comes from all sides (the sector itself, donors, investors, national and local governments, and public in general), but this is a relatively recent issue in the microfinance sector. For a long time, measurement concerns were not present at practitioner level. There was an assumption that MFIs knew very well the communities where they worked, and targeted specific populations. From their own experience, they presumed that objectives were being attained. Adding to this, donor pressure was almost inexistent if the MFIs covered their costs and there were anecdotal cases to illustrate the success of the programme (Morduch, 1999).

The ‘change of heart’ regarding the outcomes and impact of microfinance, and the consequent pressing towards accountability and transparency, were triggered by two main factors. First, the above mentioned delinquency crises (Chen et al., 2010; Rozas, 2013), which beyond the immediate negative effects on clients and institutions (Schicks, 2011; Vogelgesang, 2003), had also a significant impact on the reputation of microcredit as a key tool to fight poverty.

Second, different academic papers challenge the impact of microfinance on poverty reduction (Bateman, 2010; Chowdhury, 2009; Coleman, 1999; Karnani, 2007; Marr, 2003; Morduch, 1998). Reviews on impact studies (Duvendack et al., 2011; Gaile & Foster, 1996; Goldberg, 2005; Odell, 2015; 2010; Van Rooyen et al., 2012) introduce methodological doubts regarding reference works, such as the studies of Khandker (2005) and Pitt and Khandker (1998); and more recent analyses, based on randomised control trials, achieved ambiguous results, in most cases showing no or limited statistically significant impact of microcredit programmes on poverty reduction (Banerjee et al., 2015).

Beyond theoretical or anecdotal evidence, projects and programmes have to provide empirical evidence of targeting the intended publics and achieving the goals that they proclaim to pursue, having in this way a positive impact in the lives of people (Gertler et al., 2016). As in Hulme and Mosley (1996a, p.86), “the ultimate test of any institution is not whether it exists and sustains itself, but whether it manages to do something useful?”

Therefore, impact evaluation, as further developed in the chapter two of the thesis, is associated not only with the identification of the social outcomes (positive or negative) of the interventions, but also with the attribution in some degree of the changes observed at different levels (clients, households, communities) to the participation in the microfinance programme. Assessing the role of the programme in the production of the outcomes
represents a great challenge for most institutions, especially for socially focused institutions with limited resources.

Not-for-profit institutions (normally with the legal status of NGOs) with strong social missions were in the genesis of the microfinance movement (Karlan, 2014; Mersland & Strøm, 2011; Robinson, 2001). Bhatt and Tang (2001) have put forward four reasons for the leading role of NGOs in microfinance: first, the institutions have close links with civil society in the communities they work in; second, they have been entrepreneurial and innovative in the provision of non-financial services complementing the financial products; third, principles of “trust, generosity and ideology” guide the institutions activities which make them more reliable to work with the poor and, finally, some NGOs have adopted business-like practices aiming to improve the efficiency of their programmes.

While for-profit companies have entered into large segments of the microfinance sector, especially in the offer of credit services, Karlan (2014) suggested that NGOs still have a fundamental role to play in the provision of financial services to the poor since they are willing to serve the ‘unprofitable’ (too poor, too remote, too young); they are generally trusted by the poor which allows people to gain confidence in the financial services provided; and they can promote innovation in the sector (Karlan, 2014). Dichter (1996) referred to NGOs innovation profile arguing that these institutions are expected to be in a prominent position to take chances, experiment new solutions and bring about innovation to the sector from a social perspective.

The relevance of NGOs in the microfinance sector is suggested also by empirical studies such as the ones conducted by Bos and Millone (2015) and Microfinanza Rating (2012). The first concluded that MFIs targeting lower-end groups are more efficient than other MFIs serving not so poor clients. This higher efficiency of socially oriented MFIs contrasts with the findings of previous studies which indicate that fighting poverty and being financially sustainable - microfinance dual mission (Ledgerwood & Gibson, 2013; Morduch, 1999), are two conflicting goals (Hermes, Lensink & Meesters, 2011; Mersland & Strøm, 2010).

The latter is a study carried out by Microfinanza Rating based on a sample of 65 MFIs from 30 countries. Measuring the household poverty level using the Progress Out of Poverty Index (PPI) or consumption data, the authors found significant differences on poverty outreach between MFIs, with NGOs performing better in this matter than Cooperatives, Non-banking financial institutions (NBFI) and Banks.

Nonetheless, it is necessary to be cautious when generalizing. Being a for-profit organisation does not preclude having a socially focused mission. In many occasions, the decision to adopt
a for-profit legal status results from the country regulations for the microfinance sector. This is the case of THRIVE Microfinance (Zimbabwe), a registered NBFI whose mission is “provide training and credit to women excluded from the mainstream financial sector in a manner that is both socially responsible and financially sustainable”. It is also the case of Akhuwat Islamic Microfinance (AIM) since July 2017 when the microcredit operation of Akhuwat (NGO) was transformed into AIM, a NBFI.

Irrespectively of this diversity, and for practical reasons, in this thesis the term socially focused MFIs will be used broadly to include all these NGOs and for-profit institutions that privilege social objectives, and often have limitations accessing resources to develop their microfinance programmes and to evaluate their results.

Considering these limitations, donors and investors can play an important role by promoting transparency in the report of poverty outreach and outcomes of the programmes and supporting research on impact (Simanowitz, 2003). However, funding from donors to carry out impact evaluations of microfinance programmes is normally appropriated by larger MFIs. They have recognizable lobbying competencies as well as the financial and technical capability to prepare stronger applications to these funds; and they more easily can satisfy the scientific criteria associated with the use of large samples. Furthermore, when smaller institutions access these funds through a project led by an external institution, there is usually a time limit; after the project ends, in most cases, there is no continuity on the application of the methodologies or tools since the MFI will not be able to support the cost (Boucher, 2014).

In this context, it assumes particular relevance the development of evaluation projects focusing on socially focused institutions, like the LWC impact assessment project, in which the objectives go beyond the judgment of merit of the microfinance programmes. These projects aim also to sensitize the managers of the MFIs to the advantages of evaluation, and to build internal capacity on the process, with the ultimate goal being the development within the institutions of their own evaluation and social performance management systems.

In this respect, a significant initiative conducted in the sector was the Imp-Act project, an action-research project funded by the Ford Foundation, which involved a group of UK academics and the participation of 30 microfinance institutions. The project started from the concern of the social investor with obtaining “better-quality information about the poverty reduction and developmental outcomes of its own investments in microfinance”, aiming at improving the quality of the services provided and their impact on the lives of the microfinance clients. In order to achieve this objective, the project leaders set to distinguish
themselves from the evaluation models prevalent at the time, which were externally led and privileged external accountability objectives (Copestake and Simanowitz, 2005a, pp. 5-6). In this sense, the project adopted a learning and improvement perspective, focusing on the development of internal mechanisms to measure and manage the social outcomes of the programmes.

The institutionalisation of social performance assessment and the advocacy of its integration in the decision-making processes of the microfinance institutions has also been defended by a group of practitioners through the work developed by the Social Performance Task Force (SPTF). One of the working groups created within the network focuses on outcomes, having recently published two reports on the topic. The first report (Sinha, 2017a) made the case for outcomes management in MFIs. It starts distinguishing impact (attribution of change) from outcomes (observed changes) and acknowledging a misuse of the term impact in this sense. The focus is, hence, put on the outcomes and the advantages of adopting a systematic approach to the use of outcomes data beyond its simple measurement. In this way, MFIs can be more accountable, able to manage their mission and improve the social performance of their programmes (Sinha, 2017a). This line of thought follows the recommendation by Copestake and Simanowitz (2005b, p.214) of considering social performance assessment in the broader framework of social performance management, in which internal motivations within the institutions should be the driving force of the process.

The second SPTF report revisited the difference between outcomes and impact and provided guidelines on the implementation of outcomes management in the organisations, including a list of potential indicators to be collected and studied. The last step in the proposed guidelines relates to the use of findings, stressing the need for the MFI managers to allocate time to analyse and act on the results (Sinha, 2017b). There is an inherently pragmatic approach in the work developed by the SPTF, with many points of contact with the use-focused evaluation models (which are introduced in chapter two), including an emphasis on improving the programmes (as in the Context, Input, Process, Product model), and the importance of preparing the use of findings from the initial stages of the project (similar to the UFE model used in the thesis).

The attention in the two mentioned reports (Sinha, 2017a; Sinha, 2017b) to the different use given to the word ‘impact’ by many practitioners and many academics illustrates the differences of language regarding evaluation. In part, these differences result from different

---

2 SPTF is an international network of institutions and individuals linked to the microfinance sector, which has been developing and promoting standards and good practices for social performance management since 2005 (Social Performance Task Force, n.d.).
mindsets and interests. They derive also from different capacities to read and interpret certain types of data, especially when sophisticated analytical techniques are employed in the evaluation. In many occasions, the language and level of detail employed by academics in reports and papers of interest to broader audiences can be “befuddling to most readers” (Gaile & Foster, 1996, p.7).

The differences between ‘academics’, who in general privilege methodological rigour, and ‘practitioners’, who are more focused on the evaluation results and its usefulness, reflect in many aspects of the evaluation projects. Bolnick and Nelson (1990) provided a concrete example in their study of the KIK/KMKP credit programme in Indonesia. The researchers used recall information instead of panel data due to cost and time constraints, but also to “official insistence to minimize intrusions” (p.304). The degree of independence and freedom of the evaluators/researchers in the design and implementation of the evaluation can vary, but it is difficult to imagine a project in which there is total independence and control by the evaluator. In this way, and to some extent, the work of the evaluator and the decisions taken are influenced by and negotiated with other evaluation stakeholders, particularly the sponsors of the evaluation and the managers of the programme being evaluated.

1.4 Research design

The research developed within the thesis with the objective of answering the research questions presented in section 1.1 was based on qualitative methodologies, using a multiple case study approach. As mentioned above, the LWC assessment project was implemented in two different settings (Pakistan and Zimbabwe), allowing for the construction of two individual case studies, which were the basis for discussion of the first research question, as well as the comparison of the cases findings in order to address the second research question. The option for a case study methodology derived mainly from the nature of the questions studied and the conditions of implementation of the LWC project, in which the researcher participated actively but in a consultant position, not having full control over the project and the evaluations conducted.

The motivations to opt for a case study design, as well as the advantages and limitations associated with the methodology are further described in chapter five, with the data collection and data analysis techniques used in the research also being detailed in the same chapter. In this description, a parallel was established with the doctoral thesis developed by Gravesteijn (2014) who followed a similar approach.

The active participation of the researcher in the project has led to the use of participant-observation as the main data collection technique, in line with the recognition by Yin (2018)
of this method being advantageous in the collection of case study data. Other methods used in the research included formal interviews with the members of the evaluation teams, informal conversations with the team members, clients and field staff of the MFIs, and the review of literature and internal documentation of the partners.

In the analysis of the data collected, two of the general strategies suggested by Yin (2018, p.165) were employed, namely, developing the cases description and relying on theoretical propositions. In the construction of each evaluation description, there was a concern to provide “thick” descriptions as advocated by Lincoln and Guba (2000, p.40), which entails providing sufficient information on the case and its context to allow for the reader to understand the findings.

In regard to the theoretical framework, the analysis was based on the utilisation-focused evaluation (UFE) model developed by evaluation theorist Michael Patton. In chapter two, alternative evaluation theories and models are introduced and described. The justification of the choice of the UFE model to examine the LWC impact project and the specific framework of analysis used in the thesis are presented in chapter five.

Given the importance attributed to the selection of the evaluation methods within the UFE model, in the presentation of the LWC impact project in section 5.4.2, the evaluation tools applied in Pakistan and Zimbabwe are highlighted. These included a purposely built questionnaire and a poverty assessment tool administered through a household survey to a sample of clients and non-clients. In chapters six and seven, particular attention is given to the preparation and implementation of the household surveys and to their specificities in the two evaluation settings.

1.5 Research main contributions and originality

The PhD research resulted from the participation of the researcher in the LWC assessment project. Given the objectives of LWC and the conditions of implementation of the project, there was from the beginning an intention to involve all the partners (LWC, UoP and MFIs), in the evaluations conducted, and a strong motivation and commitment from the field partners to use the findings and the lessons learnt from the evaluations.

The project can, thus, be seen as a collaborative project, which is analysed in the thesis establishing a parallel with the utilisation-focused evaluations developed by Michael Patton (Patton, 2008). From the implementation of the LWC impact project, described in detail in the chapters five to seven, a first message and contribution from the research is that evaluation projects implying more collaborative approaches, such as UFE, when seriously
implemented, should be valued, even if field limitations and the involvement of stakeholders in the process imply not achieving the same quality standards associated with academic research on impact. Not treating these limitations lightly, they should not be the only conclusion drawn from the projects, since different evaluation projects and partners may pursue distinct objectives and uses of the results, which require different criteria in the design and the interpretation of the evaluation findings. Independently of this, it may be expected that the information provided in the description of the evaluations will be useful for different stakeholders.

There is scope (and need) for different types of evaluations, with the synergies between them potentially benefiting all actors involved in the microfinance sector. An expected consequence of an increased valorisation of these evaluations would be the encouragement to share the field challenges faced in the implementation of the evaluations and respective solutions. This will contribute to increase transparency and allow for a more informed analysis of the evaluation findings. Field challenges are often neglected, especially in academic publications, with the thesis offering a direct contribution to minimise the existing gap at this level by providing comprehensive descriptions of the main field challenges encountered in the implementation of the evaluations in AIM and THRIVE.

A second main message and contribution of the research relates to the adoption of a common project approach in the LWC impact project which translated into a set of common elements in the implementation of the evaluations in the two (very different) contexts. These included similar overall research questions, similar methodologies with respect to the application of longitudinal surveys, and the sampling and data collection techniques, same structure of the evaluation teams and stages of the evaluation.

The positive perception of the LWC team and the MFI managers regarding the project seems to indicate that the implementation by donors, social investors, or other key actors in the sector of evaluation projects in multiple settings, with similar methodological approaches, and focusing on microfinance institutions with social missions but no (or incipient) social performance assessment systems in place, can play an important role in triggering the process within the MFIs.

It is of utmost importance to stress at this point that the adoption of a common project approach does not contradict in any way the widely accepted principle of customisation of the evaluation design and methodologies to the contextual factors of the programmes and the project (Chen, 1990; House, 2003; Patton, 2008), which happened as well in the LWC project.
It is also fundamental to understand that this common project approach only makes sense under certain conditions, which are developed in the discussion on the topic in chapter eight. Among these relevant conditions are the definition of generic objectives for the evaluation, with the LWC impact project showing that this approach worked for MFIs unfamiliar with evaluation processes, giving the first steps in regard to social performance assessment and that were looking for an evaluation of the broad outcomes of their microcredit programme. It seems also crucial to ensure an active involvement of the promoter/commissioner of the evaluation, or a common team leader capable of guaranteeing the permanent engagement of the field partners and the exchange of experiences between them. Finally, it is important that the partners share capacity building objectives, privileging a learning approach in the evaluation.

In the LWC impact project, underlying the arguments just presented is the decision to adopt a pragmatic approach, closer to the mindset of practitioners, which facilitated the collaboration between the MFIs, the project leader and the evaluator/researcher. The characteristics and the preferences in terms of the epistemological perspective of the researcher(s) are, thus, also important.

The presentation of these main contributions of the thesis allows for the identification of its original elements, which are listed following without any particular order:

- The focus on a pragmatic approach as a facilitator of the working relationship between the different actors (academics and practitioners) in regard to evaluation.
- The link of the microfinance sector literature with broader evaluation theories and the use of a specific evaluation model in the analysis of the microcredit evaluations. The contributions at this level can be of interest not only to the microfinance literature but also to the literature on evaluation (theoretical and empirical).
- The analysis of the common project approach to evaluation implemented by LWC. The positive perception regarding the evaluation from the LWC partners involved directly in the first phase of the project, and the interest of other partners in be part of the project in the near future, seems to indicate that the approach has been perceived as having significant advantages.
- The promotion of transparency in the communication of the findings and lessons learnt during the evaluations, which is not original in itself, as the attention given to meta-evaluation by evaluation theorists indicates (e.g. Stufflebeam and Coryn, 2014, p.632). The variation in the thesis comes from the advocacy of
increased valorisation of collaborative evaluations for those projects demonstrating serious purposes and privileging the best possible methodological options, given the field conditions. As mentioned above, this valorisation can encourage the different partners to be transparent regarding evaluation findings and processes, which can enhance credibility and, consequently, the use of evaluations (Patton, 2008).

- The consideration of the LWC impact project as part of the initiatives conducted in the sector to develop shared understanding and a common language regarding evaluation, with special attention given to potential links with the work of SPTF.

Finally, it should be also acknowledged the analysis conducted around the application of the Poverty Probability Index (PPI), or rather said, of PPI in Pakistan and an alternative poverty scorecard in Zimbabwe. This analysis was not as developed as initially intended due to the particularities of the two countries with respect to this instrument and time constraints.

Being outdated in Pakistan and non-existent in Zimbabwe, the research questions around the utilisation of PPI as an evaluation tool had to be adapted focusing on the perceptions and behaviours of the MFI managers considering these conditions. This is one of the areas that can be further explored if the LWC impact project continues and MFIs with different experiences in the application of PPI are to be included in the project. Besides its utility measuring outreach and tracking changes in the poverty levels of the clients, the expectation is being an international recognised poverty assessment tool, its application will contribute to increase the credibility of the evaluations.

1.6 Organization of the thesis

The thesis is organised as follows. Chapters two, three and four introduce topics of interest to contextualise the research questions and analysis developed within the thesis and review the relevant literature. Chapter two identifies the main concepts associated with evaluation of outcomes and impact followed by the presentation of the evaluation theories and models related to programme evaluation. Chapter three reviews the literature on evaluation in the microfinance sector, focusing on specific issues considered of interest in the discussion of the research questions, including the challenges associated with the field implementation of impact evaluations. Chapter four addresses the use of poverty scorecards in the microfinance sector, specifically the Poverty Probability Index (PPI), contextualising its application within the LWC evaluation project.
Chapter five describes the methodologies employed in the thesis and introduces the LWC assessment project, providing an overall description of the project, its objectives, participants and methodologies. Chapters six and seven are dedicated to the description of the implementation of the evaluation in the two settings applying the analysis framework developed in the previous chapter. In the final sections of these chapters, the first research question and associated sub-questions are discussed. Chapter eight starts with the identification of the relevant similarities and differences between the two case study settings followed by the development of the arguments (for and against) the adoption of a common project approach to evaluation such as the one adopted in the LWC impact project (second research question). Finally, chapter nine presents the conclusions, summarising the main findings and messages as well as the limitations of the research and identifying areas for future research.
2 Outcomes and Impact Evaluation

There is an extensive literature addressing the evaluation of outcomes and impact of microfinance and other development programmes, both from a theoretical and an empirical perspective. This literature includes systematic reviews of microfinance impact studies (e.g. Duvendack et al., 2011; Odell, 2010; Odell, 2015; Gaile & Foster, 1996; Goldberg, 2005), comprehensive technical guides both on evaluation (e.g. Gertler et al., 2016; Khandker, Koolwal & Samad, 2010; White, 2009a; Ravallion, 2008; Baker, 2000) and on poverty assessment (Henry, Sharma, Lapenu and Zeller, 2003), and academic reflections on microfinance impact evaluation (e.g. Adams & Vogel, 2013; Armendariz & Morduch, 2010; Fouillet, Hudon, Hariss-White & Copestake, 2013; Ledgerwood, 2013). Equally prolific has been the production of social impact assessment guidelines by social investors networks (e.g. Social Impact Investment Taskforce, 2014; Kazimirski and Pritchard, 2014; Hornsby and Blumberg, 2013; Hehenberger et al., 2013; Bill and Melinda Gates Foundation, 2010).

Looking in particular to the microfinance sector, the many evaluations and impact studies developed so far do not provide conclusive findings. Neither there is consensus regarding the most adequate methodologies (Adams & Vogel, 2013).

The research developed within this PhD explores different aspects related to the implementation of the LWC evaluation project. This chapter identifies relevant concepts related to programme evaluation and contextualises the research within the evaluation literature, describing selected theories and evaluation models and justifying their choice.

2.1 Programme evaluation: clarifying concepts and perspectives

Evaluation is “the assessment of something’s worth or merit” (Stufflebeam & Coryn, 2014, p.111), or alternatively, “the determination (using a defined methodology) of the results (desired and undesired) attained by some activity designed to accomplish some valued goal or objective” (Suchman, 1967, p.32). The activity, which is object of the evaluation, is called evaluand and it can broadly refer to programmes, projects, personnel, products, organizations, policies and evaluation systems (Stufflebeam & Coryn, 2014).

A possible introduction to programme evaluation is the presentation of the evaluation chain as depicted in Figure 2.1. This approach, although deemed too simple in comparison to more recent complexity based approaches, is still frequently used, particularly in literature targeting practitioner audiences (e.g. Rogers, 2014).

Figure 2.1 suggests that the first steps in the implementation of a development programme imply gathering the inputs (monetary and non-monetary resources) required to develop the
activities which are expected to prompt change in the lives of the target populations (Rogers, 2014). In the case of microcredit, the activities relate to the credit process and the utilisation of the loan by the client.

Figure 2.1. Programme Evaluation Chain

Source: Adapted from Rogers (2014)

2.1.1 Outputs and Outreach

Outputs are the tangible results from the activities that can be delivered immediately (Sandhu-Rojon, 2010). In the evaluation of a microfinance program, the outputs are usually associated with the participation in the programme, hence, with the provision of access to formal financial products to the intended population. In this sense, a close concept commonly used in the sector is outreach.

A seminal paper by Navajas, Schreiner, Meyer, Gonzalez-Vega & Rodriguez-Meza (2000, p.335), further developed by Schreiner (2002), establishes a theoretical framework for defining and analysing outreach in the microfinance context. The authors, drawing on the theory of social welfare, expressed outreach as the “social value of the output of a microfinance institution” and defined six dimensions of outreach: breadth, depth, scope, length, cost and worth to users.

This outreach framework remains a reference in microfinance research, which is here applied to a microcredit programme (Schreiner, 2002):

- Breadth corresponds to the number of active borrowers;
- Depth refers to the value that society attributes to the net gain that a borrower gets by using the microcredit loan;³

³ The poorest clients are expected to have marginal gains from the access to microfinance that are significantly higher than those who are better off or not-poor (Mel, McKenzie & Woodruff, 2008). Thus, the depth of outreach is related to the client’s poverty level (‘how poor?’).
▪ Scope is the number of types of financial contracts offered by the MFI (diversity of products/services);
▪ Length relates to the time frame of the supply of microcredit (related to the operational sustainability of the institution);
▪ Cost captures the sum of the price and transaction costs associated with the loan;
▪ Worth to users refers to the clients’ willingness to pay for the loan.

Breadth and depth are conceptually more relevant; and they are at the centre of one of the liveliest debates in the microfinance sector – the commercial or institutionalist approach vs the poverty reduction or welfarist approach (Armendáriz & Morduch, 2010; Hermes & Lensink, 2011; Ledgerwood & Gibson, 2013; Robinson, 2001). The first focuses on the breadth of outreach based on the idea that promoting access to finance, by enlarging the number of clients, enhances financial inclusion and guarantees the sustainability of the institutions. The latter, aiming to reduce structural poverty, puts the emphasis on the clients’ level of poverty (depth of outreach), selecting the poorest segments of the population as the desired target of microfinance programmes.

Given the focus of the LWC assessment project on MFIs with explicit social missions, measuring the depth of outreach of the institutions became a priority and influenced the selection of the evaluation tools employed in the project, particularly the decision to use poverty scorecards.

### 2.1.2 Outcomes and Impact

Ensuring outreach (access to microcredit) is a necessary but not sufficient condition to produce the desired outcomes and impact (Bill and Melinda Gates Foundation, 2011; Khandker, 1998, p.14). Outcomes are the changes, or effects, on individuals or the environment that follow from the delivery of specific products or services (Social Impact Investment Taskforce, 2014). Therefore, assessing the outcomes of microfinance programmes means evaluating the changes occurring in certain variables defined as goals, for which may have been set specific targets to be achieved through the implementation of the programme. In chapter three, the expected outcomes of microcredit identified in the literature, as well as the related empirical evidence, are summarised.

An impact evaluation implies going a step further in the analysis (the last step in figure 2.1.), not only identifying the changes verified in the analysed factors (positive or negative), but also assessing if these changes resulted from the participation in the programme (Gertler et
Gertler et al. (2016) presented impact assessment as one type of evaluation, wherein there is a focus on cause-and-effect questions. Impact is, in this perspective, associated with causality and attribution.

Liket et al. (2014) referred, however, to some confusion around these concepts in the development field. Different perspectives co-exist with practitioners often linking the two terms to the time-horizon of the results. Outcomes are associated with the short-term and the immediate results of the programmes, while impact is applied for the long-term results, including the changes at community level (Ledgerwood, 2013; Rogers, 2014), as depicted above in Figure 2.1.

White (2009b) concluded that there should be no confusion as long as it is clear that there are two different definitions of impact, one linked to the attribution of change and the other to the long-term outcomes. Neither of these definitions is right or wrong and both are potentially useful for the policy makers. The author also drew attention to the misunderstanding of attribution as ‘sole attribution’, considering that an impact evaluation allows for identifying how much the participation in the programme has contributed to the overall change.

The distinct meanings attributed to impact are reflected in Copestake (2012) that distinguished between formal impact assessment, informal impact assessment and broader social science research. A formal impact assessment refers to studies conducted by external evaluators who purposely deal with the problem of attribution, often through statistical inference, i.e. the “quantification of selected variables for a sufficient number of cases to permit statistical analysis of correlations between them” (p.10). This is the perspective prevailing in the microfinance sector, adopted particularly by positivist researchers, but it is not, however, the only perspective. The author highlighted the QUIP (Qualitative Individual Impact Protocol) methodology as an alternative to the experimental approaches based on RCTs, suggesting that both approaches, although inherently different, have advantages and limitations, thus, being important to be clear about the objectives to achieve with the impact assessment.

Adding to this debate is the contribution of White and Phillips (2012). The focus of their paper was on different evaluation approaches adequate to small n interventions, including in this group not only cases with small sample sizes but also those in which strong sample

---

4 Gertler et al. (2016) present two other types of questions that can be addressed through an evaluation: descriptive questions (what is happening regarding processes, conditions, and stakeholders’ perspectives) and normative questions (comparison of actual inputs, activities and outputs/outcomes with what was established and targeted).
heterogeneity occurs. In these cases, it is often not possible to construct a comparison group with the same characteristics as the treatment group. The qualitative evaluation approaches described include realist evaluation (Pawson and Tilley, 1997), general elimination methodology (put forward by Scriven), Mayne’s contribution analysis, outcome mapping, most significant change, and the success case method.

Academics following a positivist approach aim to ‘proving formal impact’, which implies comparing what have happened to the individual participating in the programme, regarding the variables of interest, with what would have happened if he/her had not participated (Boruch, 1997, p.1). In this sense, an impact assessment implies a missing data problem since participation and non-participation are two mutually exclusive events. This problem cannot be solved but it can be minimized by finding a good counterfactual, i.e. a comparison group which is usually referred as a control group. There are different methodologies to address this issue, including randomised control trials (RCTs), matching methods, double-difference methods and instrumental variables methods (Khandker et al., 2010; Ravallion, 2008).

The debate on the validity and merits of the different impact assessment methodologies is an unresolved one and goes far beyond the microfinance sector. Experimental designs, particularly the RCTs, were (are) presented by some researchers as the ‘gold standard’ in impact evaluation (Banerjee & Duflo, 2011; Banerjee et al., 2015; Boruch, 1997; Cook, 2002; Duflo, 2004; Duflo & Kremer, 2005; Karlan, Goldberg & Copestake, 2009).

The advocacy of RCTs as the most rigorous scientific methodology to identify what works in development is “showcased” using the microfinance sector in the 2011 Banerjee and Duflo’s book “Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty”. In the book, RCTs were presented as the only credible approach (Ravallion, 2012), an approach that has been extremely influential within public institutions, international agencies and NGOs (Reddy, 2012).

Quasi-experimental methodologies, including matching and regression discontinuity, have also been adopted in many impact studies in microfinance, including reference works by Chemin (2008), Khandker (2005) and Pitt and Khandker (1998). However, non-randomised empirical research is largely affected by classic endogeneity problems: client self-selection and MFI strategies are likely to produce correlations between credit access, other outcomes and unobservable variables (such as attitude towards risk and entrepreneurial aptitude), restraining the possibilities of drawing causal inferences regarding impact from the data (Karlan & Zinman, 2011; 2010).
For Ravallion (2012, p.104), one of the attractions of RCTs is their simplicity as they promise to identify the causal effects of an intervention in a reliable way while implying few assumptions. Randomization (when properly conducted) is expected to solve some of the traditional methodological problems associated with other impact methodologies, including economists long-standing obsession with selection bias based on unobserved variables.\(^5\)

The methodology has, however, detractors who put forward several limitations. Ravallion (2012), Scriven (2008) and Stufflebeam (2001), among others, raise serious ethical concerns, considering problematic that people are seen as the objects of the experiments, and subsequently a mean to some end. In addition, Ravallion (2012) suggest that RCTs by conditioning who has access or not to the products (treatment), independently of the desire and needs of the potential clients, alter what would have been the normal functioning of the programmes they are supposed to evaluate.

Some authors (Bédécarrats, Guérin & Roubaud, 2015; Ravallion, 2012; Scriven, 2008) question the randomization process (how random it is effectively?), hence the validity of the superiority claim associated with the RCTs. These doubts are empirically illustrated by the application of the methodology in the International Labour Organisation (ILO) project ‘Microfinance for Decent Work’. The project was carried out between 2009 and 2012, involving 16 MFIs. In several of these institutions deviations from a pure randomization led the researchers to adopt complementary methodologies (differences in differences, propensity score matching, instrumental variables, regression discontinuity design, pipeline matching and client monitoring tools) to deal with differences between the treatment and the control groups (Balkenhol, 2012).

Furthermore, RCTs are costly and time consuming (Adams & Vogel, 2013; Khandker et al., 2010). If the analysis period is shortened due to imperatives related to the project timeline, it is most likely that results will suffer, meaning impact will be more difficult to probe (Ravallion & Chen, 2005). They also require a strong commitment from the MFIs in the implementation. Altogether, cost-effectiveness should be considered and compared with alternative methodologies (Marr & Awaworyi, 2012).

Critics of the methodology raise also questions regarding the external validity of RCTs, and the difficulties associated with the replicability in different contexts (Deaton, 2010; Karlan et

\(^5\) When randomisation is applied within sufficiently large samples, biases based both on observable and unobservable variables will be equally distributed between the treated and control groups. Outcome differences verified between the two groups can then be attributed to the intervention (Stufflebeam & Coryn, 2014, p.148).
Hughes and Hutchings (2011), although agreeing with the superiority of RCTs in large $n$ interventions, deem it as inappropriate and unrealistic for the evaluation of small $n$ interventions.

Finally, Adams and Vogel (2013) and Ravallion (2012) point out that a standard RCT does not provide information on the distribution of impacts (only a mean impact for those treated in the experimental population) and does not offer potential explanations for the results.

Deaton (2010, p.424) defends that in the development context, experimental designs do not necessarily produce more reliable knowledge compared with other methods, and that actual experiments often must deal with practical constraints that undermine any claim to “statistical or epistemic superiority”. RCTs are not a gold standard because “there is no gold standard” (Cartwright, 2007, p.11). Randomised, quasi-randomised and non-randomised have both strengths and weaknesses and probably the most adequate solution entails combining methodologies (Balkenhol, 2012; Hulme, 2000b; Reichardt, 2011). Focusing on the evaluation of client outcomes in microfinance programmes, Copestake (2012) suggests that there is a role for qualitative methodologies such as in-depth interviews, particularly if these methodologies can be combined with periodic surveys and poverty monitoring.

2.2 Overview of evaluation theories

2.2.1 An evaluation theory tree

Evaluation theories and practices have developed steadily since the 1960s boosted by several factors, including the economic growth and the interventionist role of the Government in the United States after World War II, an increase in social science graduates interested in public policy and field experiments, and previous work in the areas of education, social psychology and sociology (Shadish, Cook & Leviton, 1991, p.21).

Alkin and Christie (2013, p.11; 2004, p.12) identify three roots of evaluation theory, which support the trunk of the “evaluation theory tree” (Figure 2.2.): social accountability, systematic social inquiry and epistemology. Social accountability is considered by the authors in a broad sense, including a perspective of improvement of the programmes. These three roots are the basis for the development of the tree branches: use, methods and value. In their book, the authors summarise the work of the most influential evaluation theorists and position them within one of the branches. Their classification is used below to highlight the main issues associated with each of the branches.
2.2.2 Evaluation theories: the methods branch

The theorists included in the methods branch are mainly concerned with the recognition of evaluation as research (Alkin & Christie, 2004). Suchman (1967, p.12) illustrates this motivation by affirming that “evaluative research is, first and foremost, research and as such must adhere as closely as possible to currently accepted standards for research methodology”.

The first author referred by Alkin and Christie (2004) in the methods branch is Ralph Tyler. His work in the 1940s is considered precursor to many other theorists. Tyler (1942) put forward an objectives-oriented approach to the evaluation of educational programmes, in which the central elements are the specification of objectives to be achieved through the implementation of the programmes and the measurement of its outcomes.

The most influential author in the methods branch is, however, Donald Campbell. His work with Julian Stanley, “Experimental and quasi-experimental designs for research”, published in 1966, while not addressing specifically programme evaluation, became a reference in terms of evaluation methodologies (Alkin & Christie, 2004).

In their book, Campbell and Stanley (1966) establish a framework for the internal and external validity of findings. Internal validity is related to the results of the research being considered robust and replicable (Schram, 2005). Campbell and Stanley defined it as “the basic minimum without which any experiment is uninterpretable”, while associating external validity with the (never fully answered) “question of generalizability” (Campbell & Stanley, 1966, p.5).
The tension between the two concepts translates into the discussion among academics from different fields regarding the hierarchy of the concepts. Experimental economists tend to privilege internal validity when the objective of the experiment is testing a theory, but external validity gains relative importance in the development of new theories (Schram, 2005, p.234).

Cronbach focus his attention on external validity, suggesting that in social sciences, if local contexts are taken into account, the generalisation of evaluation results is hardly achievable and should be perceived as a working hypothesis rather than a conclusion (Cronbach, 1975, p.125).

Different authors have been influenced by and expanded the work of Campbell and Stanley, either emphasizing the methodological superiority of experimental designs, as in the case of Robert Boruch (1997) who focused on randomised control experiments, or further developing the alternative quasi-experimental designs like Thomas Cook (Alkin & Christie, 2004). Quasi-experimental designs result from the introduction of experimental elements into natural settings. Despite the researcher not being able to fully control the experiment, quasi-experimental designs were considered by Campbell and Stanley (1966, p. 34) as a worthy alternative when better designs (i.e., true experiments) are not feasible.

In the methods branch, Alkin and Christie (2004) include also Peter Rossi, Huey Chen and Carol Weiss. The work of these three authors was at the genesis of programme theory and its application to evaluation. Recognising the lack of utilisation of evaluation findings, Weiss suggested broadening the scope of evaluation to include the theoretical principles of the programmes (Chen, 1990, p.32; Weiss, 1972, p.321). Chen, building on the work developed jointly with Rossi (Chen & Rossi, 1983), acknowledged the dominance of experimental approaches in evaluation but suggested that these approaches should be complemented with the development of theoretical models for the interventions. These underlying theoretical models can provide guidelines for determining the priority issues in an evaluation and the most adequate methods to address these issues. They can also clarify the connections between the operations of a programme and its effects (Chen, 1990).

To the advocates of theory-driven evaluations, the programme’s theory should be derived from existing social science research, not being dependent on the subjective, and potentially biased, values of the stakeholders. These theorists do not preclude the participation of stakeholders but assign the central role in the process to the evaluator (Christie & Alkin, 2003). Indeed, Chen (1990, pp.77) proposed that theory-driven evaluators see themselves as
one of the evaluation stakeholders whose main objective is to achieve a useful and high-quality evaluation.

### 2.2.3 Evaluation theories: the value branch

Value is the “very root of the term evaluation” (Guba & Lincoln, 1989, p.62). For the theorists included in this branch, evaluators must put value on their findings since the main aim of programme evaluation is making a value judgement about the programme (Alkin & Christie, 2004).

Michael Scriven is the main reference among value theorists. He describes evaluation as “the science of valuing”, in which it is not necessary to understand why a programme works to determine its value (Alkin & Christie, 2004). To him is attributed the distinction between a summative and a formative evaluation (Stufflebeam & Coryn, 2014, p.183). The first is carried out after the completion of the programme and is normally destined to an external audience or decision-maker. The latter is associated with the gathering of data (during a specific period of time) with the objective of improving the implementation of the programme, anticipating problems and making sure that the programme short term results are in accordance with the accomplishment of the intended outcomes (Patton, 2008).

Contrasting with Tyler’s tradition, Scriven (1991, p.56) defended that identifying and measuring the programme specific goals is unnecessary and can potentially be a “contaminating step” in the process. He suggested goal-free evaluations in which the evaluator does not consider the defined programme objectives, focusing the analysis on the expected outcomes associated with the fulfilment of the needs of the programme stakeholders, and identifying unexpected outcomes. This approach is sceptically seen by several authors. Alkin considered Scriven’s suggestion as only partially goal free (‘free’ from the goals defined by managers), since the evaluator ends up formulating what he/she considers to be the desirable goals of the programme. Stake (1983a) raised the question of the faith put by Scriven on the ability of evaluators to identify the real needs of the programme stakeholders.

In the value branch, two other important references are Robert Stake and Ernest House (Alkin & Christie, 2004). Stake put forward the responsive evaluation model. In this approach, evaluators investigate, negotiate and select a number of issues to address within the study, focusing on the concerns of the stakeholders rather than specific programme theories or

---

stated goals. In this process, evaluators are willing to sacrifice some accuracy in measurement in order to increase the usefulness of the findings for those with interest in the programme (Stake, 1983a; Stake, 2003). Stake is also one of the main advocates of the use of case studies in evaluation (Stake, 1983b; Stake, 1994).

House developed the deliberative democratic evaluation, an approach associated with the advocacy of social justice in evaluation. It is based on three main principles: inclusion of the interests of all stakeholders, dialogue between all participants (evaluators should not assume what are the interests of the stakeholders), and deliberation (House & Howe, 2000). House (2003, p.54) reinforced the advocacy for the participation of all stakeholders in the evaluation, although admitting different degrees of involvement. The author worried about the power balance between stakeholders, and how the poorest or less vocal may find themselves excluded from the study, when most probably they are the final beneficiaries of the programme.

Finally, Guba and Lincoln (1989) in their book “Fourth Generation Evaluation” introduced an alternative approach to evaluation based on the constructivist paradigm. The authors described three previous generations of evaluation focusing on the measurement of outcomes, the description of the programme and the judgment of its value. However, they identified a number of problems associated with these evaluations, including a tendency towards managerialism, a failure to accommodate multiplicity of values and an over-commitment to the scientific paradigm.

The constructivist paradigm denies the existence of an objective and unique reality. Guba and Lincoln (1989) support the idea of multiple realities, which represent different constructions of the mind of those involved in the programmes. Furthermore, social interventions are as likely to influence as they are to be affected by the context in which they are implemented. In this sense, the possibility of generalisation of evaluation findings is minimal, if not, impossible.

As the findings of the evaluation emerge from the interaction between observer and observed, evaluators become subjective partners with the programme stakeholders in the creation of data. Beyond the traditional evaluator roles (technician, describer, judge), evaluators are facilitators of a negotiation process between all stakeholders in the attempt to achieve a consensus regarding the value of the programme. It is the evaluator’s responsibility to make a judgment, taking in consideration the claims, concerns and relevant issues for the stakeholders (Guba & Lincoln, 1989, p.71).
Stake’s responsive evaluation, House’s deliberative democratic evaluation and Guba and Lincoln’s constructivist evaluation share a strong commitment to social justice through evaluation, advocating the engagement of all stakeholders in the process. The three models share limitations as well, which are mainly associated with the feasibility of their implementation and the challenge of encouraging different stakeholders to be engaged throughout the project and receptive to more open and exploratory findings (Stufflebeam & Coryn, 2014, p.207).

2.2.4 Evaluation theories: the use branch

The focus on evaluation use emerged in the 1970s as evaluators saw how infrequently decision makers used the findings of evaluations (Weiss, 1972, p.10; Chelimsky, 1978). Frustration with this lack of utilisation led to a new form of inquiry into the factors that influence and contribute to use (Patton, 2008, p.24), aiming to guarantee that evaluation results have an impact on the decision process (Christie, 2003, p.8).

This frustration was also the basis for the development of programme theory. In line with the main advocates of theory-driven evaluations, Wholey (1987) established the definition of the theoretical basis of a programme as a pre-requisite for the programme to be evaluated. It is important to note, however, that the need to develop a programme theory prior to the evaluation is not acknowledged by all authors. Christie (2003, p.33) suggested that an evaluation theory is not a requisite to practice. In her study of the practices of 138 evaluators of health programmes in the United States, she found evaluators frequently adopted specific parts of a theory according to their practical needs. Equally, House (2003, p.56) referred to the gap between theorists and practitioners, and how the latter are not expected to embrace a particular theory when implementing an evaluation.

The use branch has an underlying pragmatic basis: the value of an evaluation resides in how it is used and the consequences of that usage. To make an evaluation useful, the relationship between the evaluator and the programme stakeholders can take any form, as long as it serves the purpose of the evaluation and complies with ethical considerations (Wilson & Mertens, 2012, p. 90).

The promotion of use requires the evaluator to be active, play multiple roles (Alkin & Christie, 2004; Patton, 2008, p.213), and find different and creative ways to engage diverse stakeholders (Stufflebeam & Coryn, 2014). The involvement of stakeholders in the research is expected to increase the possibilities of producing useful findings as they will more “authentically reflect the social context of the participants” (Carey & Asbury, 2012, p.38).
An important question related to the role of the evaluator refers to the implications of cultural differences in the evaluation. Patton (1985) suggests that every evaluation becomes a cross-cultural encounter. Evaluators must be culturally competent, which entails being actively aware and appreciative of the cultural contexts (Sengupta, Hopson & Thompson-Robinson, 2004) and responsive in a “dialogue across difference and distance” (Wadsworth, 2001, p.48). In international settings such as the LWC assessment project, the differences can relate to beliefs and values, language, styles of interaction and communication, sense of time and expectations regarding infrastructures (Merryfield, 1985). This does not mean the evaluator needs to abandon his/her own cultural background and view of the world, or neglect training and skills, but makes it essential that the evaluator constantly self-examines the values, assumptions and cultural context of the evaluation (Sengupta et al., 2004).

The emphasis given to the role of evaluators in encouraging the use of evaluation findings is partially curbed by Contandriopoulos and Brousselle (2012). The authors recognise that the choice of a use-based evaluation model and the competence of the evaluator can contribute to increase use, however, they consider that evaluation context is also a determinant factor. They give particular attention to the cost-sharing structure associated with the evaluation, suggesting that whenever the users are willing to support the costs of the process, the probability of use significantly increases.

Wilson and Mertens (2012, p.91) propose that, similar to the relationship between evaluator and stakeholders, different evaluation methods can be chosen insofar as they match the goals of the study, which often means the adoption of mixed methods, combining quantitative and qualitative methodologies. Chen (1990, p.82), while defining the principles for planning a theory-driven evaluation, sets out the method contingency principle, which argues that there may not be an evaluation method or design which universally applies to every situation. The method should fit the evaluation context.

Finally, it is important to emphasize that the use of an evaluation may not refer exclusively to its results; process use can be equally important (Fetterman, 2003; King, 2007; Patton, 2008, p.109; Podems, 2007). In this sense, the lessons learnt during the implementation of the project contribute to build evaluation capacity within the institutions (King, 2007), allowing those involved in the evaluation to ‘learn by doing’ how to think and act critically (Bradley Cousins & Earl, 1992, p.400). In the same direction, Patton (2008, p.153) suggests evaluative thinking can bring permanent value to the organization, outlasting the findings of the evaluation project.
Process use is associated with participatory evaluation, defined by Bradley Cousins and Earl (1992, p.399) as “applied social science research that involves a partnership between trained evaluation personnel and practice-based decision makers, organisation members with programme responsibility, or people with a vital interest in the program”. This partnership adds complexity to the evaluation since questions relating to the control of the process and the scope and depth of participation need to be negotiated (Patton, 2008), but then it also potentially enhances the buy-in of the results and recommendations of the evaluation (Christie, 2003, p.29; Fetterman, 2003, p.49).

2.3 Evaluation models and approaches

There are different approaches to implement an evaluation, and there have been several attempts to classify and assess these different evaluation approaches and models, including Madaus et al. (1983), Scriven (2003), Stufflebeam (2001) and Stufflebeam and Coryn (2014). Stufflebeam and Coryn (2014, p.111) summarised 23 evaluation approaches/models and comparatively analysed nine of these models using the five categories of quality standards (utility, feasibility, propriety, accuracy and evaluation accountability) established by the Joint Committee on Standards for Educational Evaluation (JCSEE, n.d.). Although the quality criteria have been established within the Education field, the concepts were defined broadly and, thus, applicable to other contexts.

The nine models were selected by the authors based on their own experience, taking into account that these models are widely used in programme evaluation and represent different approaches to evaluation. From the analysis, the context, input, process and product (CIPP) model, the utilisation-focused (UFE) model and the constructivist model were classified as very good approaches, while the other six approaches, including the experimental and quasi experimental, and the case study approach were rated as good (Stufflebeam & Coryn, 2014, p.230-232). Considering this classification, the three best-rated approaches are briefly described below. The choice of the model used in the research, among these three, and its justification are presented in chapter five.

2.3.1 Context, Input, Process, Product model (CIPP)

The CIPP model was initially developed by Daniel Stufflebeam in the end of the 1960s and offers a comprehensive approach to evaluations. The purpose of a CIPP evaluation is to “provide timely information in a systematic way for decision making”, which implies an approach focused both on proving the merits of the programme and identifying
improvement areas and recommended actions (Stufflebeam, 1971, p.2; Stufflebeam, 1983, p.118; Stufflebeam & Coryn, 2014).

As conveyed by its designation, a CIPP evaluation can focus on four types of elements – context, inputs, process and product. Context evaluation is associated with planning decisions and translates into the identification of unmet needs, opportunities and problems that can be addressed through the programme in the specific setting. Input evaluation relates to the assessment of alternative strategies to achieve the programme planned objectives in regard to the different types of resources employed (structuring decisions). Process evaluation aims to describe, monitor and assess the activities being implemented, allowing for a comparison with the planned operations (implementation decisions). Finally, product evaluation assesses if the defined objectives of a programme are being achieved and whether the implemented programme should continue, be modified or terminated (recycling decisions) (Stufflebeam, 1971).

Considering the nature of the LWC assessment project, a parallel can be established to a CIPP product evaluation. Stufflebeam and Coryn (2014, p.313) consider four possible components of product evaluations: reach to the intended beneficiaries (outreach), effectiveness (achievement of targeted outcomes), sustainability (sustainable and affordable over time) and transferability (strategies and procedures adaptable to other contexts). In conducting a product evaluation, the evaluator should take in consideration both intended and unintended effects, as well as short-term and long-term outcomes (Stufflebeam, 1983).

An important feature of the CIPP model is the involvement of all stakeholders, or their representatives, in the evaluation process. Evaluators should pay particular attention to those more vulnerable and with less influence over the program, encouraging them to share information and opinions. The involvement of all stakeholders is not only seen as ethically responsible but also as sensitive, facilitating the process of acceptance, valorisation and action on the evaluation reports (Stufflebeam & Coryn, 2014, p. 316).

Stufflebeam (1971) suggests different data collection methods for each of the four types of evaluations. The preference for combining methods is also seen as a way to obtain a broader view of the results and cross-validate the evaluation findings (Stufflebeam, 1983).

Stufflebeam and Coryn highlight the importance of ethical principles in the model. CIPP evaluators should actively try to minimise bias and conflicts of interest while conducting assessments and interpreting results; triangulate information from multiple sources; report findings fairly and cautiously to all the interested audiences; and identify areas for further investigation. Essential to the model is that evaluations should not be regarded as “one-shot”
investigations which are only used as accountability tools to please external funders (Stufflebeam & Coryn, 2014; Stufflebeam, 1983, p.118).

2.3.2 Utilisation-focused model (UFE)

The UFE model was proposed by Michael Patton, who describes the approach in detail in the four editions (1978, 1986, 1997 and 2008) of his book “Utilisation-focused evaluation”. UFE evaluations are undertaken to provide information about the program, and in this way, inform decisions, clarify options and identify potential improvements. They are conducted with specific and intended uses and are, thus, a basis for action (Patton, 2008, p.37).

In the context of UFE evaluations, the identification of the evaluation stakeholders is of utmost importance. These are people who can affect (or be affected) by the evaluation process or its findings. Greene (2006) categorizes four groups of stakeholders: those who have decision-making authority over the programme (funders, political decision makers); those who have direct responsibility for the programme (managers); those who are the intended beneficiaries of the program; and those disadvantaged by the programme (e.g. people who lost funding opportunities).

In the definition presented above, UFE evaluations seem very similar to CIPP evaluations. There are, however, significant differences between the two models, mainly in what refers to the involvement of the programme stakeholders in the evaluation. The essence of the UFE model is its focus on a selected group of stakeholders – the intended primary users - and what Patton (2008, p. 66-67) designates as the “personal factor” in the evaluation (Alkin & Christie, 2004).

The primary intended users are the specific stakeholders who will be able to make changes in the programmes and are, therefore, chosen to work with the evaluator and participate in all aspects of the evaluation, making UFE evaluations inherently collaborative efforts. The depth and form of the participation depend on the context, the purpose of the evaluation, the skills of those involved and what has been agreed between the evaluator and users (Patton, 2008, p.177). There is no explicit request for the involvement of all stakeholders as in the CIPP model.

In UFE, the role of the evaluator is more of a consultant than an external, objective evaluator. Owen and Lambert (1998, pp.363-364) suggest that the differences between evaluators and organizational development consultants have become blurred. For the authors, evaluators with skills on organizational change are better positioned to promote an informed use of the evaluation findings.
The presence of a recognisable group of people within the programme who personally care about the evaluation and its findings is the “personal factor” required to enhance use (Patton, 2008, p.66). In all stages of the process, including the production of reports, the needs and values of the primary intended users, and not those of the evaluator, are to be addressed (Stufflebeam & Coryn, 2014, p.406).

Patton (2008, p.213) proposes a “active-reactive-adaptive” approach by the evaluator. UFE evaluators should be active in the identification of the primary intended users and design with them the evaluation project; they should be reactive, focusing on the users’ perspective and responding to their ideas; and they should be flexible to adapt to changes. The users’ reactions and way of thinking are expected to change along the evaluation, so the evaluator needs to accommodate different options and approaches to the project.

Negotiation is vital during an evaluation. Neither evaluator or users should dogmatically impose their views, with negotiation seen as a learning process and a mechanism to increase project ownership for all members of the evaluation team. In UFE, the evaluation is more responsive and interactive, and less independent. “It calls for problem solving and a creative process of adapting evaluation procedures to meet the local and specific evaluation needs as they emerge” (Stufflebeam & Coryn, 2014, p.411).

An important feature of the UFE model is, thus, the consideration of the evaluation context, which means that there is no single way, and no unique design and methods in the implementation of an evaluation. This argument fits in with Chen’s method contingency principle enunciated above and with House’s recognition that evaluation practices are strongly limited by contextual and structural factors (House, 2003, p. 53).

The UFE model does not prescribe specific evaluation methodologies, rather leaving to the evaluation team the choice of the most suitable methods to collect the best possible data, given the available time and resources, and considering the needs of the intended users. In this sense, the chosen evaluation approach should be the best fit to the program, which configures a mixture of people, history, politics, resources, constraints, values, needs, interests and chance (Patton, 2008, p.199).

In summary, the UFE is a process of matching intended uses and intended users whose critical elements are “situational responsiveness, methodological flexibility, multiple evaluator roles, political sophistication and substantial doses of creativity” (Patton, 2008, p.28).

Patton (2008, p. 36) refers to an increase in evaluation use since the UFE model started to be widely implemented, stressing the importance of focusing and properly preparing this use. Weiss (2004, p.161) also seems to have overcome some of the frustration expressed in her
1972 work, recognising that along her career she oftentimes found decision makers who have changed their priorities and questioned the programme’s assumptions following the implementation of an evaluation.

2.3.3 Constructivist model

The constructivist model is based on the work of Guba and Lincoln (1989). This evaluation approach is strongly paradigm driven, relying in the belief that there are multiple realities associated with the different perceptions and interpretations of those involved in the programme (Stufflebeam & Coryn, 2014). The individuals and the context in which they find themselves are fundamental to define social reality. Changing one or both also changes the reality (Guba & Lincoln, 2013, p.39).

Consequently, the evaluator should inform and consult stakeholders in all aspects related to the evaluation, from the definition of the evaluation questions to be addressed to its findings. In the attempt to achieve a consensus regarding the merit of the programme, the varying and many times conflicting values of the stakeholders must be considered (Stufflebeam & Coryn, 2014). Fourth-generation evaluations “do not report on an objective truth but rather represent the result of negotiations” (Scriven, 1993, p.64).

The constructivist approach is based on the use of qualitative methodologies. By employing exploratory and participatory approaches, which result in extensive and time-consuming evaluations, it becomes more difficult to plan and budget the evaluation project. This type of approach requires a constant involvement of a wide range of stakeholders as well as the acceptance by the project clients that the reported findings will not provide a clear judgement of the programme’s merit, but instead “competing, perspectivist answers” (Stufflebeam & Coryn, 2014, p.201-202).

2.3.4 Other evaluation models

Despite not developing the other evaluation models and approaches analysed by Stufflebeam & Coryn (2014), a note on the experimental and quasi-experimental approach is seen as relevant for its link with the previous discussion about RCTs. Stufflebeam (2001) presented some of the advantages of controlled experiments, an evaluation approach which had gained widespread acceptance after successful applications in medicine and agriculture. These included the ability to provide unequivocal causal inference regarding the programme impact, and the focus on objective results rather than intentions or judgments. However, the author also identified serious drawbacks associated with these methods, including serious ethical objections to experimenting on human subjects and the narrow scope of use of
experimental methods, not allowing for the investigation of specificities of the process or the needs of the target population. In this sense, experiments and quasi-experiments were considered insufficient to address the full range of questions required to assess the merit and worth of a programme.

It is also worthwhile mentioning the attention given by Stufflebeam (2001) and Stufflebeam & Coryn (2014) to the use of case studies as evaluation methodologies, in line with Stake (1983b). Stufflebeam (2001) presented supporting arguments to this idea, referring to the apparent easier application of the methodology (compared to experimental methodologies), how it can be applied to programs at different stages of development and how it may allow for a holistic perspective of the evaluand, looking at the programs in-depth and analysing contextual influences. Furthermore, case studies can be tailored to address its final users’ questions and the power of its conclusions can be enhanced if multiple cases are conducted within the same area of study.
3 Evaluation in Microfinance

3.1 Outreach, outcomes and impact of microcredit: a theoretical approach

Independently of the evaluation approach adopted and the motivations of the different programme stakeholders, the evaluation of a microcredit programme usually begins by assessing the outreach of the programme (particularly in its breadth and depth dimensions) and follows with the identification of changes in the lives of the participants during the selected period.

The definition of the target outreach by MFI managers is linked to the institution’s mission, and its position in the debate between the institutionalist and the welfarist approaches (Ledgerwood & Gibson, 2013; Robinson, 2001). Taking into consideration that the focus of the thesis is on institutions privileging social objectives (hence within the welfarist campus), targeting translates into the efforts to reach the poorest segments of the population or, alternatively, the poorest of the economically active (Mathie, 2002, p.17).

Measuring the outreach of MFIs is relevant as it allows MFI managers to acknowledge, seek the causes and correct the two types of targeting errors identified by Hoddinot (1999, p. 5-7). There might be errors of inclusion or leakage (reaching individuals not intended to be the beneficiaries/clients), and errors of exclusion or under-coverage which translate into excluding intended beneficiaries. The author suggests that these unintended results may result from inadequate selection criteria, insufficient dissemination of the programme to the intended public, or wrong alignment of staff incentives with target objectives, among other causes.

Targeting the poorest segments of the population often entails the definition of specific strategies to identify, reach and attract the poor and discourage the participation of the non-poor. These decisions imply costs; hence, targeting strategies must be carefully considered (Mathie, 2002, p.18).

Achieving the intended targeting is a necessary condition for MFIs to accomplish their social mission, but it is not sufficient. Measuring targeting efficiency and assessing the outcomes and impact of the programmes at client level are two complementary, but distinct, objectives (Khandker, 1998).
3.1.1 Microcredit expected outcomes

Microcredit programmes are expected to generate outcomes at different levels. Access to credit, and the consequent changes in social relations within and beyond the household, are likely to generate effects (positive and negative) on all aspects of the lives of the borrowers. As such, from a theoretical perspective, links can be established between credit and a wide range of socio-economic and well-being indicators (Duvendack et al., 2011). Khandker (1998, p.16) suggests that the best way to evaluate a microcredit programme is measuring its “impact on the poor in terms of employment, income, consumption, assets, net worth, nutrition, contraceptive use, fertility and children’s schooling”.

Productive loans (when effectively applied in the business), by enabling an increase in working capital, fixed capital or human resources, are expected to generate higher business profits and, subsequently, higher household incomes. These additional funds allow for increasing and diversifying consumption and/or increasing savings, granting the client and her/his household a better financial situation and a better bargaining position when accessing other financial products. Figure 3.1 represents a simplified version of this virtuous cycle of credit described by several authors, including Duvendack et al. (2011) and Hulme and Mosley (1996a).

![Figure 3.1 – The ‘Virtuous’ Cycle of Credit](image)

Source: Own construction

Other potential positive effects of microfinance described in the literature include: diversification of household income sources; business and household assets accumulation; and improvements in risk management by smoothing consumption and reducing vulnerability to external shocks, such as illness and climate disasters (Banerjee & Duflo, 2011; Collins et al., 2009; Hermes & Lensink, 2011; Karlan, 2014; Morduch, 1998; Vial & Hanoteau, 2015).
Access to finance may also have “social” and “empowerment” outcomes, including improvements in education, health and housing conditions, and in the social and economic situation of women when they are the borrowers (Duvendack et al., 2011, p.13). In the non-financial outcomes, Van Rooyen et al. (2012) add also nutrition, food-security and social cohesion, which links to Hermes and Lensink (2011) reference to potential spill-over effects at the local community level that go beyond the MFI client. These wider impacts which affect others beyond the client, including “family members, neighbours, employees, competitors, consumers and suppliers” were classified as indirect impacts in the context of the Imp-Act project (Copestake & Simanowitz, 2005a).

Moreover, microfinance, due to its expected capacity to serve poorer populations on a regular basis, often in remote places, can be an important platform for the development and provision of non-financial services (including education and health services) that are as important to the poor as financial services (Gray, Gash, Reeves & Crookston, 2011).

Microcredit outcomes are, however, not always positive (Bateman, 2010; Chen et al., 2010; Chowdhury, 2009; Fernando, 2006; Guérin et al., 2014; Hulme, 2000a; Karnani, 2007; Marr, 2003; Schicks, 2013). The emergence of delinquency crises in different microfinance markets across the world in the 2000s has brought these concerns into the mainstream. As a result, there has been an increase in academic research focused on the “dark side” of microcredit (Hulme, 2000a), and practitioner initiatives have been developed at national and international levels (e.g. the Smart Campaign) to sensitize microfinance providers to these negative effects, and to promote client protection mechanisms.

In practice, the virtuous cycle of credit does not always happen. External shocks such as job loss by a household member, health emergencies or natural disasters, as well as poor management decisions resulting from lack of skills and knowledge can result in lower business profits (Duvendack et al., 2011; Hulme, 2000a). In these cases, the capability to repay the loan decreases, and the borrowers often default on their loans. Non-repayment of microcredit loans can be a relevant sign of client over-indebtedness, but as Gonzalez (2008), Morvant-Roux et al. (2015) and Schicks (2013), among others have shown, it is neither a necessary nor a sufficient condition. Gonzalez (2008) and Schicks (2013) identified in Mexico and Ghana, respectively, clients who repay their loans by making sacrifices in other areas of their lives, including reducing consumption, selling assets, using savings, or working extra time. They can also resort to borrowing from informal lenders, increasing their indebtedness level beyond their repayment capacity. These situations are especially important in countries such as the Dominican Republic where maintaining creditworthiness is perceived as essential by the poor (Afonso, Morvant-Roux, Guérin & Forcella, 2017; Morvant-Roux et al., 2015).
Criticisms of microcredit in the literature have also been based on the public merit of the initiatives, given the type of businesses supported by microcredit and the scale of operations. Mayoux (2006) suggested that the rapid growth of microfinance in urban areas, such as Lusaka and Harare, has led to a concentration on labour intensive and low-profit activities, leading to market saturation in some of these activities. Several authors questioned the ability of businesses funded by microcredit to generate enough returns to make an impact on the poverty level of the households (Dichter, 1996; Karnani, 2007; Mel et al., 2008), and some advocated alternative policy measures aiming to promote employment and develop the small and medium enterprise segments (Bateman, 2010; Karnani, 2007). Finally, there have been authors calling into question the positive association between microcredit and women’s empowerment (Fernando, 2006; Garikipati, Johnson, Guérin & Szafarz, 2016; Guérin, Kumar & Agier, 2013; Johnson, 2005; Kabeer, 2001; Vaessen et al., 2014).

3.2 Outreach, outcomes and impact of microcredit: the empirical research

Morduch (1999) drew attention to the lack of a performance assessment culture within the microfinance sector. Institutions and programmes were (are) many times operating based on myths; which is in line with Michael Patton’s recognition that committing to test goal achievement and question beliefs is “neither natural nor widespread”. In this sense, programme evaluations focused on use are a form of reality testing as they allow for verifying if what is hoped for is, in fact, taking place (Patton, 2008, pp.43-44).

A first step in this reality testing is the measurement of poverty outreach, especially for socially-focused MFIs - is the programme reaching the target clients? In microfinance research, a common approach to this complex task is the utilisation of poverty proxies such as the average loan size; the percentage of female borrowers; the outreach in rural areas; and the percentage of group loans in the MFIs’ portfolio (Bos & Millone, 2015; Mersland & Strøm, 2010; Morduch, 2000; Schreiner, 2002). The rationale for the utilisation of these poverty proxies is described in section 4.1.

These proxies have been mostly employed in comparative studies on the performance of microfinance institutions using secondary data. Such an example is provided by Cull, Demingürç-Kunt and Morduch (2007) who compare the performance of 124 institutions based on the privileged lending methodology (group-lending, village banking or individual-lending).7 The authors conclude that group-based institutions performed better regarding

7 Village banking is a specific lending model which is also based in a group methodology.
depth of outreach, since their average loan size was significantly lower, and they served a larger percentage of female borrowers, when compared with individual-based institutions. It is expected that group-based microcredit programmes reach poorer clients since the group lending methodology was introduced to overcome the barriers faced by these segments of the population when trying to access commercial loans. Among these barriers are the lack of a credit history and personal assets to pledge as collateral, and the irregularity of income usually associated with the small and informal businesses of these potential clients. Group loans are, thus, more likely to appeal to the poor segments of the population, which is strengthened by specific targeting strategies employed by the MFIs.


Departing from the comparison of MFIs (which is not the topic of the thesis) to an analysis focused on the clients, Henry et al. (2003) identified three main methods to assess the poverty level of a household. The first is based on the analysis of household expenditure and the computation of an income poverty line. The poverty line identifies the level of income sufficient “to meet the food and other basic needs of all household members needed for a healthy and active life” (Henry et al., 2003, p.169). National poverty lines are constructed from data collected through nationally representative household surveys such as the Living Standard Measurement Survey; for purposes of international comparison, the World Bank publishes reference international income levels associated with poverty and extreme poverty ($2/day and $1.25/day, respectively). 8

An example of outreach measurement using a reference poverty line is given by Navajas et al. (2000). The authors investigated the depth of outreach of five Bolivian MFIs through a cross-sectional study. By means of a household survey administered to a sample of 622 clients in 1995, they concluded that the institutions were not reaching the poorest segments of the population but the better-off among the poor. They also found that institutions based on group lending methodologies performed better on depth of outreach compared with individual-based organizations.

---

8 The widely used international poverty income line of $1.25/day was established with reference to 2005 Purchasing Power Parity (PPP) rates for consumption. In 2015, the World Bank updated these values with reference to 2011 PPP, and the income line for extreme poverty was raised to $1.90/day (Ferreira et al., 2015).
The second method proposed by Henry et al. (2003) is the use of participatory appraisal methods. These methods involve the participation of the community in the identification of poor and vulnerable groups, and they are used by many development institutions as a targeting strategy. These are, however, assessments that best suit local projects as they have embedded a subjective analysis strongly dependent on the local context.

The third option is indicators-based. The process begins with the identification of a range of indicators that reflect different dimensions of poverty, and for which credible data can be collected rapidly and at a low cost. The data can then be aggregated into a simple index of poverty (Henry et al., 2003). This methodology is the basis for the development of the poverty scorecards further explored in the next chapter of the thesis.

Going a step further, there is an extensive and constantly growing academic literature on the empirical evaluation of outcomes and impact of microcredit programmes offering mixed results, particularly in regard to the impact of microfinance on poverty reduction.

### 3.2.1 Systematic reviews and reports

The literature review presented in this and the next sub-sections of the thesis does not aim to be exhaustive, highlighting a selection of key reference studies. It starts by chronologically appraising reviews and reports on the impact of microfinance, which is followed by a summary of the methodological choices and findings from the main studies. It concludes by looking at the papers from an evaluation process perspective. As the thesis lays emphasis on microcredit, the literature on other microfinance products and services, such as savings and insurance, is consciously overlooked.

Sebstad and Chen (1996) reviewed a comprehensive group of 32 evaluation reports of microcredit programmes (19 using quasi-experimental designs). The authors summarised the findings of these studies, distinguishing different levels of impact, and briefly presented concerns regarding the methodologies used. These respect to not rigorously address selection and placement bias, and not include the beneficiaries who have dropped-out from the programme in the study samples.

Gaile and Foster (1996) paid more attention to the methodological challenges, reviewing eleven impact studies on microenterprise programmes. Although these studies were selected due to being considered the most rigorous impact assessments at that time, the authors concluded that in all of them the control for selection bias was problematic due to the adoption of quasi-experimental designs. Both Gaile and Foster (1996) and Sebstad and Chen (1996) were prepared to support the design of evaluations of microfinance
programmes supported by the United States Aid Agency (USAID), within the AIMS (Assessing the Impact of Microenterprise Services) Project.

Brau and Woller (2004) provided a more comprehensive literature review on microfinance. The authors drew attention to the focus of the initial impact studies on Bangladesh, identifying ten studies on the country, including Pitt and Khandker (1998). They also highlighted the relevance given to women’s empowerment, an outcome evaluated in seven of these studies, including Goetz and Gupta (1996) and Hashemi, Schuler and Riley (1996).

In 2005, in the first review commissioned by the Grameen Foundation, Nathaniel Goldberg was one of the first to refer to the use of randomised control trials in the sector (at the time the first projects were starting), placing high hopes on the ability of this methodology to eliminate the problems associated with selection bias (Goldberg, 2005). Odell (2010) updated Goldberg’s report. The author contrasted the mixed results of the initial RCTs with the more positive findings of the studies using quasi-experimental methodologies, presenting the comparative advantages and disadvantages associated with experimental studies.

In 2005 was also published the book “Money with a Mission”, one of the outputs of the action research project Imp-Act funded by the Ford Foundation. The project had its roots in the aim of the social investor to gather evidence of the impact of the microfinance programmes it was financially supporting, but the approach to impact evaluation was different from the conventional at the time, privileging a learning/improvement perspective in which the participating MFIs were actively involved. The book summarises the overall conclusions of the individual projects developed in the 30 participating MFIs, both from the perspective of the different levels of impact of the microfinance programmes and the methodological challenges faced in the implementation of these projects, and discusses the advantages and challenges associated with the internalisation of social performance management by the MFIs (Copestake and Simanowitz, 2005a).

In 2011, a seminal paper by Hermes and Lensink offered a summary of reference empirical studies relating to two main questions: the impact of microfinance on the well-being of the poor and the sustainability of microfinance institutions, focusing on the trade-off between sustainability and outreach (Hermes & Lensink, 2011). In the same year, a systematic review by Duvendack and colleagues went further in the discussion of the methodological issues associated with impact evaluation. The authors considered most of the impact studies developed in the sector (until that moment) had employed weak methodologies and inadequate data. They put particular emphasis on the need to collect high-quality data as
some of the underlying problems of data may not be corrected by applying sophisticated econometric methodologies (Duvendack et al., 2011, p.3).

Vaessen et al. (2014) and Van Rooyen et al. (2012) presented systematic reviews of impact studies on microfinance covering specific topics. Van Rooyen et al. (2012) focused on studies in Sub-Saharan Africa. The authors identified 15 studies (11 on microcredit programmes) considered to be of sufficient quality, favouring the four RCTs included in the review for their methodological superiority. The authors emphasized the challenges of generalising conclusions due to the multiplicity of programmes and outcomes evaluated, as well as the diversity in results, with both positive and negative outcomes being identified in the studies. Vaessen et al. (2014) reviewed papers dealing with the impact of microcredit on women’s control over household spending. The authors included 29 quantitative-only studies, concluding that weak theoretical frameworks and methodologies call into question the positive results found in some of the studies.

In 2015, the third review in the Grameen Foundation series was authored again by Kathleen Odell. The author started by acknowledging significant transformations in the sector since 2010, with the shift from microfinance (providing financial products and services) to financial inclusion (ensure access and effective use of financial products in an overall ecosystem that includes payments systems and integrates a wider range of providers), and the subsequent implications for research in the sector (Odell, 2015).

In line with Duvendack et al. (2011), Odell (2015) focused on the RCTs being developed in the sector, giving notice that while the initial projects focused on microcredit as isolated interventions, there has been a progressive enlargement of the scope of the programmes studied. Notwithstanding, the author emphasized that RCTs are not suitable for all evaluation projects and pointed out alternatives, including qualitative methodologies.

The challenges associated with the operationalisation of RCTs were also a conclusion in the ILO final report from the action-research project ‘Microfinance for Decent Work’. The project aimed at testing “new ways of working by launching a pilot experiment and client level assessment in one specific decent work aspect” (Gravesteijn, 2014, p.7). Between 2009 and 2012, 16 MFIs implemented programme innovations associated with four thematic clusters: vulnerability, child labour, business performance and formalisation. ILO (2015) presented the project, its goals and results, and included a methodological chapter in which the individual projects implemented were classified in terms of data quality.

In the same year, Beck provided a critical review of literature on microfinance as a part of a large-scale evaluation project led by the World Bank in 15 countries (Beck, 2015); and
Gopalaswami and colleagues published a systematic review focused on impact studies in South Asia. The latter summarized the findings of 69 studies, most of them conducted in Bangladesh or India (Gopalaswami, Babu & Dash, 2015).

Finally, Peters, Lockwood, Munn, Moolu and Mishra (2016) reviewed qualitative studies focusing on client perception regarding the participation in microfinance programmes. In their systematic review, the authors included an in-depth analysis of the data collection methods, which seems to reflect a concern with the field implementation of these methods and its implications.

All these reviews and reports show the heterogeneity of the sector, its institutions and programmes, the multiplicity of methodological approaches employed in the evaluation of outcomes and impact, and the diversity of findings. Some of these reports referred to the importance of collecting high-quality data, drawing attention to issues associated with data collection. However, especially in academic publications, this concern was overshadowed by the focus on selection and placement bias and, consequently, on the evaluation design and data-analysis stages.

The proliferation of reviews and reports has motivated Duvendack and Mader (2017) to formalise an intention to conduct a systematic review of the reviews on the impact of microfinance in the developing countries.

3.2.2 Methodological choices

From a methodological perspective, evaluating impact in microfinance implies dealing with complex problems related to money fungibility, as well as selection and placement bias, which have not been totally resolved by any of the available methodologies (Adams & Vogel, 2013; Duvendack et al., 2011; Khandker, 1998, Von Pischke & Adams, 1980).

In Goldberg (2005), the studies conducted using methodologies developed in the AIMS project were highlighted. The author drew attention to the constitution of the control group in some of these studies using new clients who had not yet received the loan. In line with Karlan (2001) this procedure was not recommended. The most rigorous of the AIMS evaluations were considered to be the three studies implemented in MFIs in India (Chen & Snodgrass, 2001), Zimbabwe (Barnes, 2001) and Peru (Dunn & Arbuckle Jr, 2001) using large samples, longitudinal data and comparison groups of non-clients. Data analysis was based on a before and after comparison using different statistical methods.

The evaluation tools and the manual developed as a result of the AIMS project became instrumental in the sector. As described by Greeley (2005), they were the basis for the
methodological choices made by many of the participating MFIs in the Imp-Act project, with the (longitudinal) impact survey and the loan use and savings tools being the most applied. In fewer cases, the MFIs preferred to develop their own survey instruments with the support of partner universities.


The methodological choices are not confined to quantitative methods, which is probably linked to the recognition of the limitations associated with quantitative survey data in explaining the impact of the microfinance programmes (Sebstad & Chen, 1996, p.5). Hashemi et al. (1996) adopts a mixed methods approach combining ethnographic work in six Bangladeshi villages with a household survey to a large sample of borrowers. The cross-sectional data was analysed using logistic regression models. Mixed methods are also employed by other researchers. Hulme and Mosley (1996a; 1996b) describes research conducted in 13 MFIs in seven countries where data was collected from the MFIs’ database and through questionnaires and interviews with clients, MFI staff and other key informants. Copestake, Bhalotra and Johnson (2001) studied the impact of a microcredit programme in Zambia, using a questionnaire-based survey of clients, focus groups and interviews with key actors and non-clients, as well as secondary data. Copestake, Dawson, Fanning, McKay and Wright-Revolledo (2005) details an impact study carried out in Peru, in which a baseline household survey to clients of three MFIs was followed by in-depth interviews with sixty clients.

Although on a smaller scale, there are also a small number of influential studies based exclusively on qualitative approaches. Goetz and Gupta (1996) construct loan histories for a purposive sample of female borrowers from four microcredit programmes in Bangladesh. Todd in her book “Women at the Center: Grameen Bank Borrowers After One Decade” develops an ethnographic experiment in villages served by the Grameen Bank (Goldberg, 2005).
Despite the increasing interest on the application of mixed methods to evaluate outcomes and measure impact (Balkenhol, 2012; Copestake, 2012; Copestake et al., 2005; White, 2002; 2009b), experimental methodologies, and particularly randomised control trials, are considered by many academics the most credible methodologies (Duflo, 2004; Duvendack et al., 2011; Odell, 2015; Van Rooyen et al., 2012). RCTs have been increasingly applied in the sector in the last decade, with the first project being conducted in Hyderabad, India (Banerjee et al., 2009; Banerjee et al., 2015).


3.2.3 Mixed results

The heterogeneity of the microfinance sector (and the implemented programmes) is important to understand the diversity of impact evaluation methodologies applied and the large number of variables and indicators used in the various studies (Adams & Vogel, 2013; Brau & Woller, 2004; Hulme, 2000b), as illustrated by Van Rooyen et al. (2012).

There are, however, a number of recurrent outcomes which are common to many research projects, including variables related to business outcomes (profits, business assets, employment created, business income); personal and household income, expenditure and assets (personal and household total income, salaries, subsidies and other sources of income, total expenditure, expenditure on food, education, health, non-basic activities and luxury items, household assets variation); financial practices (borrowings, savings, insurance) and social outcomes (schooling, health, food security, women’s empowerment).

The first reference study in the sector was conducted by Hossain (1988) in Bangladesh and compared Grameen Bank members with eligible non-participants from villages with and without a Grameen presence. The results showed that members had an average household income 43% higher than non-participants in comparison villages, and 27% higher than non-
clients in Grameen villages. Members also spent more money on food, clothing and housing than non-members. The author acknowledged differences between the groups that may have overestimated the results, namely the fact that members were on average younger, better educated and more likely to own land. Goldberg (2005) emphasised that the results of Hossain study must be put into context as the Grameen Bank in 1985 (when the data was collected) was still an experiment and many of its clients were men.

Hashemi et al. (1996) found the Grameen Bank and Bangladesh Rural Advancement Committee (BRAC) microcredit programmes in Bangladesh to have positive impacts on eight dimensions of women’s empowerment (mobility, economic security, ability to make small purchases, ability to make larger purchases, involvement in major household decisions, relative freedom from domination within the family, political and legal awareness, and involvement in political campaigning and protests).

In 1998, Pitt and Khandker published what is viewed as one of the most influential studies in the sector. They analysed the impact of participation in three group-based microcredit programmes (Grameen Bank, BRAC and Bangladesh Rural Development Board - BRDB) on labour supply, schooling, household expenditure and assets. Their main finding was that there was a positive impact of the programme in poor households, and that this impact was larger when the participants were women.

Pitt and Khandker (1998) was the first impact study trying to seriously address selection bias and non-random programme placement (Goldberg, 2005). The authors used a quasi-experimental survey to control for unobserved differences at individual, household and village level. The survey was conducted during the period 1991-1992 and collected data for a sample of 1,798 households, including clients from the three programmes and non-clients. The three programmes shared one common target criteria: potential clients should own less than half an acre of land, meaning that those above this threshold should not be allowed to participate in the programmes. This rule was used by Pitt and Khandker to apply a sophisticated econometric regression model using instrumental variables (Pitt & Khandker, 1998).

Morduch (1998) contested the results of this study questioning the methodology applied. Morduch argued that between 20 and 30% of the members actually owned more than half an acre of land; thus, the comparison with the group of non-clients, selected to comply with the rule, was flawed. He also criticised the econometric model applied arguing that it did not correct programme placement bias. Instead, he applied a simpler model to the same data and found little evidence of impact on the different variables. He did, however, confirm Pitt
and Khandker’s finding that participation in the programmes contributed to consumption smoothing (meaning less variation in household consumption across different seasons), a result that held across the three microcredit programmes involved in the study.

Pitt and Khandker’s study, as well as Morduch’s response, has continued to gather the attention of academics over the years. In 2005, Khandker published an update of the study using longitudinal data, in which the 1991-1992 survey is complemented with a new survey conducted in 1998-1999. The panel data allowed for the application of an econometric model with fewer assumptions, generating more reliable estimates of impact (Goldberg, 2005). The author found a larger effect of microcredit on household expenditure for women than in the first paper, with diminishing returns to borrowing over time. There was no statistically significant evidence of impact for men. Khandker (2005) also compared poverty rates at the two moments in time and concluded that poverty had decreased in all villages, but that credit clients experienced sharper declines, especially those in extreme poverty. Furthermore, microfinance spill-over effects in the villages with the presence of MFIs were also partially responsible for the decline in poverty of non-participants in those villages.

The AIMS evaluation studies found, in general, positive outcomes resulting from microcredit programmes. For example, Barnes (2001) concludes that in Zimbabwe, Zambuko Trust clients were more likely to have acquired new household assets (particularly stoves and fridges), and continuing clients were more likely to have higher incomes compared with new clients and non-clients. Greeley (2005) concluded that for the MFIs participating in the Imp-Act project, the financial services provided had contributed to improvements in terms of income and assets at the household level in general, although this was not true for all clients and all branches of the MFIs. Other impact studies showing positive impacts of microcredit for poor clients include Dunford (2006), Imai et al. (2010), Khandker and Samad (2013) and Littlefield et al. (2003). Simanowitz and Walter (2002) found microfinance outcomes to be beneficial for most groups of clients, although the evaluated MFIs did not reach the poorest segments of the population.

In their study on the impact of microcredit on women in Bangladesh, Goetz and Gupta (1996) found that 39% of the women in their sample had no or little control over the loan, implying that men were the final recipients of the credit. This finding is used by several authors to question the positive impact of microcredit on women’s empowerment but, within their paper, the authors do not explore the causes or the consequences of this situation (Goldberg, 2005). The appropriation of loans by men was reported in several other studies (Garikipati, 2008; Mayoux, 2006; Safavian & Haq, 2013; Todd, 1996).
Two other studies conducted in Pakistan question the impact of microcredit on women’s empowerment. Asim (2009) found no evidence of statistically significant differences between the clients of two MFIs and non-clients. Zulfiqar (2017) suggested that the increasing trend for commercialization of microfinance services has resulted not only in the programmes failing to achieve their empowerment objectives, but instead contributing to increased gender inequalities in the access to formal financial services.

Banerjee et al. (2015, p.1) summarised the results of the six pioneering RCTs in the microfinance sector, concluding that their authors consistently found “modestly positive, but no transformative, effects”. A significant finding common to the studies was the lack of impact of microcredit programmes on poverty reduction, one of the emblematic claims of the sector, and this finding was used to produce a number of negative media articles on microfinance (Odell, 2010). Banerjee et al. (2015), however, presented a more nuanced perspective on the results, suggesting that while null results in the econometric tests do not support the hypothesis of impact of microcredit on poverty, they also do not refute it, and that other positive effects (albeit modest) were found.

Among the limitations common to these RCTs is a focus on the average borrower (Adams & Vogel, 2013), not exploring the recognised heterogeneity among borrowers across the MFIs participating in the studies. The differences between microcredit clients, namely in regard to the poverty levels, have been acknowledge in several studies and translate into differences in the impact of the programmes (Hulme & Mosley, 1996a).

In addition to the average approach, the time-span of the studies is too short for any single intervention to produce transformative results in poverty (Chen et al., 2010). In some of the studies, most notably Crépon and colleagues (2011) in Morocco, the take-up of credit by the potential borrowers was initially very low, exposing an incapacity of the researchers to fully understand the context of intervention and the determinants for participation (Morvant-Roux, Guérin, Roesch & Moisseron, 2014). In this case, a subsequent qualitative study identified social and cultural reasons for the low take-up of the programme in rural areas, namely local perceptions on debt and creditors. The Morocco example gives credence to advocates of a pluralistic perspective and supports the adoption of mixed methods in evaluation (Morvant-Roux et al., 2014).

From the different impact studies implemented in the sector over the last 30 years, the overall conclusion is that the results from microcredit programmes are mixed and vary strongly with the context of the programme (Brau & Woller, 2004; Garikipati, 2017).
Subsequently, the generalisation of results across different studies/contexts cannot be easily accomplished (Hermes & Lensink, 2011).

To a large extent, the diversity in the results is explained by the difficulties associated with “documenting credit impact” (Adams & Vogel, 2013). These difficulties are associated with the nature of the loans (fungible) and of some of the expected loan benefits (qualitative), as well as the diversity of programmes and the multiplicity of methods used to deal with selection bias.

Adams and Vogel (2013) draw also attention to other factors, equally important, but seldom overlooked when an impact evaluation and its results are reported. Among these under- accounted factors are problems associated with inaccurate information provided by the interviewees (intentionally or not) and careless behaviours by interviewers; transaction costs for the MFI clients and non-clients related to the time spent on interviews; difficulties measuring community externalities of the programmes; and conflicts of interest arising from the involvement of programme stakeholders in the evaluation.

The quality of data used in microfinance impact assessment is, thus, linked to the mitigation of both fungibility and endogeneity (selection bias) problems, and of measurement errors resulting from the field challenges as identified by Adams and Vogel (2013). The quality standards and rigour for the evaluations are not consensual, being dependent on the objectives and intended uses of the assessment results, as well as on the understanding and epistemological preferences of the researchers.

An example of this was provided in the ILO final report of the ‘Microfinance and Decent Work’ project (ILO, 2015). As mentioned above, the report includes a methodological chapter in which the individual projects were classified as low, medium and high in terms of the overall credibility of the evaluation results. This classification was achieved by attributing to each project a positive (+) or negative (-) performance in six categories dealing with different methodological issues, including sample sizes, baseline and follow-up surveys, randomisation of the innovations and data quality. Although some examples of the criteria used are given (see e.g. p.28 on data quality) these are not enough to have a clear image of the process.

Partially because there is some subjectivity associated with an exercise like this, the authors of the report made a distinction between the “credibility of the evaluations and the significance of the results”, suggesting that there were valuable results across all the classification categories. In the cases with low overall quality, they suggested to not “discuss impact effects but, rather only descriptive indicators for the observed outcomes” (ILO, 2015, pp.26-28).
3.3 Challenges in the implementation of evaluations in microfinance

The empirical studies and reviews on microfinance impact and outcomes included in this literature review, setting aside the comparative studies based on secondary data, tend to identify research design and data analysis methodologies as the critical factors to guarantee the reliability and credibility of the studies. Comparatively less attention has been paid to the challenges and bias potentially emerging from the field implementation of the chosen data collection techniques.

In the papers employing quantitative methodologies, the authors emphasized the minimisation of response bias, both in the choice between experimental and quasi-experimental designs, and the selection of statistical and econometric methods used for data analysis. Navajas et al. (2000) and Pitt and Khandker (1998) are two examples. These papers presented the results of microcredit programmes evaluations funded by international organisations, including the Organisation for Economic Co-operation and Development (OECD), USAID and the World Bank. Although the assessed MFIs and their programmes are briefly presented in the papers, there is no information on why these institutions were selected, the motivations of their managers to participate in the studies, the challenges associated with the data collection, or how these may have affected the results.

Goetz and Gupta (1996) used qualitative methodologies, particularly in-depth interviews based on the life stories of the respondents. The authors were very thorough in describing their sampling process, contextualising and justifying the selected outcomes, and discussing the evaluation findings and their policy implications. However, there is not much information on field procedures (who conducted the interviews for example) and the challenges encountered. Hashemi et al. (1996) used mixed methods to analyse gender related questions, combining a structured household survey with ethnographic fieldwork. Their article gives more detail on the data collection process, including information on the local ethnographic team and interviewers, and describing and contextualising the choice of outcomes linked to women’s empowerment. There is also a justification for the selection of the Grameen Bank and BRAC, although, as with the other studies mentioned above, there is no reference to the motivations of the institutions to participate in the evaluation.

Gaile and Foster (1996) highlighted the clear definition of the research purpose by the authors of the eleven studies included in their review (to evaluate the impact of a specific microcredit programme), as well as the rich description of the methodologies selected, especially for data analysis. The authors considered, however, that most studies lacked clarity in the identification of the criteria used to assess the evaluation results; did not provide much
information on the questionnaire design (including issues such as its duration and the translation of the questions to the local language); and, equally, paid little attention to the field implementation of the household surveys, namely to the selection, training and supervision of the survey enumerators (essential to data accuracy and reliability).

An exception in this regard is the study of Bolnick and Nelson (1990) which highlights a number of questions related to the field implementation of the research. The authors shared their satisfaction regarding the results of using recall information, considering that they had expected this method to be problematic in terms of data quality. However, they refer as well to other not so successful field issues and the measures taken to minimise them. Among these issues were the specification of data requirements for some of the indicators, the careless implementation of the survey by the enumerators, and questions related to data cleaning and analysis. The authors concluded that the survey methodology overall “worked well”, and the data were “free of significant bias”, despite the identified issues (p.306).

In two systematic reviews on microfinance, Vaessen et al. (2014) and Van Rooyen et al. (2012) focused almost exclusively on the analysis of the methodologies and findings of the studies which made the cut in terms of methodological rigour. Brau and Woller (2004) concentrated on the results and the methodological debate surrounding impact, but they included a reference to a paper from 1999 presenting the conclusions of a survey on the evaluation practices of microcredit institutions based in the United States. The paper authored by Woller, Wheeler and Checketts (1999) suggested that the majority of the 73 respondent institutions evaluated their microcredit programmes regularly, although applying inexpensive and unscientific methods. It is important to note that most of these institutions were small sized (less than 100 active loans) and running domestic programmes (only 16 institutions worked in the developing countries). The authors found funding sources to be a “significant impetus for performing evaluations” (p.72), and simultaneously, the lack of funding as the main obstacle to effectively implement the evaluations. In the paper, they drew attention to the role of different stakeholders in the process of evaluation, including funding sources and programme beneficiaries, but did not develop the topic.

Duvendack et al. (2011), Goldberg (2005) and Odell (2015;2010), similar to Brau and Woller (2004), developed two main themes: the debate around research design and methodologies, and the findings of the impact studies included in the respective reports. It should be noted that the Duvendack report, being a systematic review, is more extensive in terms of number of studies reviewed and develops the methodological debate further, compared with the reports sponsored by the Grameen Foundation.
Duvendack et al. (2011, p.7, 31) drew attention to the quality of the data collected in the studies, suggesting that the poor (overall) quality derives mainly from deficiencies at the design stage of the research. The potential challenges associated with the field implementation (for example, attrition in longitudinal studies) were seen by the authors as a part of the relevant design considerations.

The focus on the research design and the applied statistical and econometric methodologies can be partially explained by the target academic audience of most of these papers, and the pursuit of some detachment to guarantee impartiality and objectivity in the evaluation judgment as prescribed by Pawson (2014). Many of these studies, which offer an academic perspective regarding the impact of microfinance, appear to have been conducted for the purpose of increasing the general knowledge on the topic, not giving particular attention to the use of the findings in the decision-making process of the programmes.

Considering what has just been described, the main researchers implementing RCTs in microfinance have been innovative in their advocacy for transparency regarding data and procedures, and this approach has probably been one of the reasons contributing to the growing support for these studies. Banerjee et al. (2015) noted the effort made by the authors involved in the six RCTs analysed in the paper to use similar ways to present the methods and results of their studies, making them more easily comparable. Adding to the communication strategy, the datasets of the studies are freely available. Odell (2015, p.19), however, drew attention to some voices from practitioners pointing out that this harmonization effort may have resulted in treating microcredit as equivalent interventions in the six countries, when in reality the MFIs, programmes and contexts were very different.

The idea of a common project approach to evaluation and impact assessment of microfinance programmes in different contexts is a difficult one to address given the heterogeneity of the sector, the relevance of the contextual factors and the nature of the programmes dealing with social issues. Looking through the microfinance literature, the topic is hardly mentioned with scattered references found in papers addressing other topics. Obviously, part of this apparent lack of interest results from the smaller number of cross-country research or evaluation projects documented.

In the systematic reviews and reports included in the thesis, there are two experiences in which the common methodological approach is in some way referred – one is related to the research implemented by Hulme and Mosley (1996a) with the support of the Overseas
Sebstad and Chen (1996, p.4) drew attention to the common methodological approach used in the studies conducted by Hulme and Mosley. The studies shared the “same basic framework, set of research questions, and, in most cases, impact variables”, with this “consistency in the methods” being considered useful in producing substantial evidence in relation to the research questions and allowing for addressing key policy issues.

It is interesting to note that in the main publications associated with the ODA research, Hulme and Mosley did not emphasize the methodological aspects of the research, not being explicit the intention of applying similar methodological approaches in the five countries included in the project. In fact, Hulme and Mosley (1996a), the first volume of the book “Finance Against Poverty”, and Mosley and Hulme (1998), an academic journal article, focus on the global results of the studies and their implications for the MFIs and the sector overall. In both cases, the design of the research and the methodologies employed are very briefly and broadly presented, without any detail regarding the different sites of implementation of the impact assessment. In the presentation of the findings, some of the field challenges are referred in explaining the limitations of data.

Similarly, the second volume of the above mentioned book (Hulme and Mosley, 1996b), which describes the country case studies, does not go much further in the methodological approach. From the five case-studies included, only the Kenya and Malawi cases include an appendix detailing the employed methodology. In the Indonesia, Bangladesh and Sri Lanka chapters, specific references to the methodology are scattered throughout the text and do not give much information. From what is possible to retain from the case studies, there was a common approach, as suggested by Sebstad and Chen (1996), at the level of the theoretical framework and overall research questions, but with the field strategies being adapted to each context. This is exemplified with the sampling strategies, different in each country. The Kenya, Malawi and Bangladesh cases had in common the inclusion in the study of a group of new borrowers who had not yet received the loan, but while this group was used as control group in Bangladesh, in Kenya and Malawi, the control group was formed by non-borrowers.

---

9 ODA, presently designated Department for International Development (DFID), is the UK governmental aid agency.

10 The case of Bangladesh includes general information on the sampling strategy, the enumerators recruited and the questionnaire translation and structure. In the cases of Indonesia and Sri Lanka, the information is much more incomplete (Hulme and Mosley, 1996b).
The impact assessments conducted in India, Zimbabwe and Peru within the AIMS Project seem to have employed similar types of studies, sample sizes, control groups and analysis methods as suggested in Goldberg (2005). The comparison of the methodological approaches followed in the studies is better perceived by consulting the individual project publications, which provide extensive detail and discussion of each case (Barnes, 2001; Chen & Snodgrass, 2001; Dunn & Arbuckle Jr, 2001).

Comparing with the ODA and AIMS impact projects, which focused on the evaluation of mature microcredit interventions (programmes who have been implemented for some time), the ‘Microfinance for Decent Work’ project developed by the ILO aimed to assess the implementation and impact of programme innovations addressing decent work aspects for the MFIs’ clients. The initial intention of the project promoters was to conduct RCTs in the 16 participating MFIs, but in the implementation of the project they have encountered multiple obstacles and ended up with a mix of methodologies being applied in the different research settings (Balkenhol, 2012; ILO, 2015).

The main methodological challenges were summarised in the final report. These included staff turnover at the MFI level; MFIs’ client databases not update; product (innovation) uptake slower than predicted, influencing sample sizes; client drop-outs; missing data, errors in data entry and data inconsistencies; issues in the experimental design (differences between treatment and control groups); and unfamiliarity with research concepts, particularly in regard to the control group (ILO, 2015, p.28).

The reservations in conceiving a common project approach to evaluation and impact assessments are not necessarily linked to research in multiple countries. The much more frequent evaluations of different institutions within the same country (geographical area) can also raise questions in this regard. The heterogeneity in these cases derives mostly from differences between the organisations and the programmes implemented. This was the case of Pitt and Khandker (1998), in which three microcredit programmes were evaluated (Grammeen Bank, BRAC and BRDB). Gaile and Foster (1996) reviewing the study considered that Pitt and Khandker treated the three programmes as similar (focusing on their common features, namely the group lending methodology) when they were different in many aspects, including the complementary non-financial services associated with the loans, which should had been included in the design of the research.

The focus so far has been on the systematic reviews and reports previously mentioned, which in most cases analysed externally led evaluations, privileging the independency of the processes and results. An exception was the ILO project, in which the partner MFIs
participated in the design and implementation of the innovations tested and in some cases were also involved in the collection of evaluation data. This fact, considered along other factors, resulted in different levels of rigour and robustness of the individual evaluations conducted (ILO, 2015).

The other exception in terms the involvement of the MFIs is the Imp-Act project. In this case, the active involvement of the field institutions and their internalisation of social performance assessment mechanisms were purposeful: “each was responsible for its own work” (p.8). The internalisation of social performance management by the institutions was seen as having multiple advantages, including an increase commitment of the MFIs’ managers towards accountability and transparency as well as capacity building on the evaluation process and social performance, with the integration of mechanisms facilitators of the MFIs’ decision-making process. In this way, the participation in the project was expected to contribute to the implementation of an improvement culture of the programmes and their impact (Copestake and Simanowitz, 2005a).

As mentioned in section 3.2.2, in the implementation of the individual projects, many of the participating MFIs based the methodological choices for their evaluations on the tools and the manual developed within the AIMS project. These tools were designed not only aiming to produce rigorous evidence, but also to fulfil criteria related to feasibility, cost-effectiveness and relevance, an approach which was consonant with the objectives of the Imp-Act project (Greeley, 2005).

From the description of the results of eight of the individual projects presented by Greeley (2005), it can be concluded that although the implementation of longitudinal household surveys, as prescribed in the AIMS manual, was common, the surveys as well as the complementary tools and methodologies applied varied across the different projects. The author identified broadly the overall implementation challenges: “sampling procedures, identifying and interviewing comparison groups, design of instruments, staff training, data analysis and report write-up were all problems of varying seriousness” (p.65).
4 Poverty Scorecards in Microfinance Evaluation

The increasing interest in the evaluation of outreach, outcomes and impact in the microfinance sector has led to a growing focus on the methodologies used and how they may allow for comparisons between programmes. This chapter focuses on the use of poverty scorecards in the microfinance sector, contextualising the application of this tool in the LWC evaluation project.

4.1 Poverty measurement in microfinance

As referred in the literature review on evaluation in microfinance presented in the previous chapter, average loan size, women, rural location and group lending methodologies are poverty proxies frequently used to compare the average poverty levels of the clients of different MFIs.

Bos and Millone (2015) made the case for the wide use of average loan size as a poverty proxy by showing that average loan size and depth and breadth of outreach were negatively correlated. In their study using data reported to the Mix Market by 1,146 MFIs, they have found that smaller loan amounts were significantly associated with higher levels of client poverty. In a similar direction, Morduch (1999) considered the average loan size to be a convenient proxy (in the absence of better data) under the assumption that poorer households cannot afford higher loans and will be the ones willing to take the smallest amounts. Its easy computation contributed also to its widespread use in the sector. Being one of the indicators reported by MFIs to the Mix Market, it became easily accessible for academics using secondary data in their research, such as the studies listed in section 3.2.

Average loan size is not, however, a perfect proxy (Simanowitz & Walter, 2002). Schreiner (2001) identified different dimensions of loan size, including term to maturity, dollars disbursed, average balance and time between installments. Thus, a first issue associated with loan size is its multidimensionality, implying different ways to calculate it according to the dimensions considered. Dunford (2002) suggested that Morduch’s (1999) assumption that only poorer clients apply for smaller loans was not empirically confirmed, which is in line with Lønborg and Rasmussen’s (2014) argument that loan size varies with the economic sector giving the example of agriculture loans which are traditionally larger, independently of the poverty level of the borrower.

The identification of lower average sizes with MFIs’ pro-poor approaches is not linear. Gonzalez and Rosenberg (2006) pointed out that loan size may be determined also by the MFIs’ lending policies, including cross-subsidization strategies and progressive lending.
mechanisms which are implemented with the objective of diversifying funding sources and risk, as well as retaining clients within the MFI as their businesses increase (Armendáriz & Szafarz, 2011).

Women are poorer than men (Armendáriz & Morduch, 2010; Duflo, 2012), and they frequently face stronger barriers to access formal financial services (Guérin, 2011). They are expected to make better use of the loans from a development perspective, prioritising family interests ahead of their own (Duflo, 2012; Pitt & Khandker, 1998). However, the “feminization of microfinance” (D’Espallier, Guérin & Mersland, 2011) is not only linked to the poorer status associated with women, but also to the belief that female clients perform better regarding the repayment of the loans, which is supported by studies such as Schmit and Marrez (2010) in Morocco. This positive relation between repayment performance and gender is not totally confirmed by the literature, which offers contradictory results in this matter (Hermes et al., 2011).

The demand for lower loans from women can also be explained by the smaller scale of their businesses and the type of activities they usually perform. In many societies, women are pushed by socio-cultural factors to more labour intensive and with lower returns activities (Fafchamps, McKenzie, Quinn & Woodruff, 2011; Mel et al., 2008). The fact that many donors and investors consider positively the maintenance of lower average loan sizes when assessing the social performance of the institutions, introduces perverse incentives towards privileging women and controlling the loan size (Dunford, 2002). The gender proxy has therefore to be analysed carefully as it may cover up gender bias causing differences in the loan requests even for similar poverty levels (Armendáriz & Szafarz, 2011).

The case for rural poverty is based on strong empirical evidence showing that rural populations are generally poorer than their urban counterparts (IFAD, 2016, p.17), while the rationale to consider the share of group loans on the MFIs portfolios as a proxy to the poverty levels of the clients, is related to the essence of the lending methodology.

Group lending was one of the innovations introduced by the microcredit movement to overcome the barriers faced by poorer populations when trying to access the traditional financial markets (Mersland & Strøm, 2011). Poorer clients do not have a previous credit history to support their willingness to repay the loans, they do not have assets to offer as

---

11 These inequalities may stem from lower rates of labour market participation; concentration in businesses with traditionally lower returns, lower growth opportunities and higher competition; limited ownership and access to assets (especially land); constrained spatial mobility; and lower educational levels restraining their ability to deal with bureaucratic formalities (Dichter, 1996; Guérin, 2011; Mel et al., 2008).
collateral and their incomes are usually low and irregular. Group loans with joint liability allow for self-selection of the clients and peer monitoring, reducing the operational costs of the institutions and allowing for the replacement of traditional forms of collateral (Armendariz & Morduch, 2010; Karlan & Morduch, 2010). Several empirical studies indicate that MFIs focusing on group lending had a stronger pro-poor approach than those privileging individual lending methodologies (Cull et al., 2007; Hermes et al., 2011).

These indicators are commonly used in the microfinance literature to compare MFIs. They are, however, not suitable for the evaluation of the poverty levels of the clients of a particular programme (and their evolution) and, given the described limitations, they can only be described as convenient proxies in the absence of better data as suggested by Morduch (1999).

An alternative to these poverty proxy measures is the use of income indicators. These indicators seem to be straightforward measures of the poverty levels but, in practice, they present several difficulties. First, data collection through surveys is costly, time-consuming, and measurement errors are frequent. Second, cultural reasons may result in reluctance sharing information about money which can create a significant volume of missing data or incorrect information. Third, information is not easily verified given the informality of the activities, and it is difficult to accurately value non-paid work, auto-production and non-monetary incentives. Finally, poor households have often irregular incomes, which can be associated with casual work and/or the seasonality of the activities. Therefore, data from a specific period will probably be insufficient for a rigorous image of the household income dynamics (Banerjee et al., 2015; Desiere, Vellema & D’Haese, 2015; Henry et al., 2003). Thus, the methodological choices addressing these issues should be taken in consideration in the analysis.

Another option is the calculation of consumption measures, which seems more attractive as “consumption covers a larger perimeter than income, in that it encompasses both market and non-market resources” and “compensates for the shortcomings of underreported income” (Vial and Hanoteau, 2015, p.143). Other advantages of consumption measures are the fact that most people share information on their consumption easily than on income, and that consumption is expected to be more constant through the year minimizing potential seasonality effects associated with the different sources of income. Moreover, more than having sufficient income to access products and services, it is their use what generates well-being (Deaton & Grosh, 2000, p.91-94).
Similar to income measurement, collecting data on consumption presents many challenges. One of the most significant is the need to define a reference consumption basket since it is not feasible to gather information for all the items consumed in the period. Even so, this basket should be as comprehensive as possible, implying that data collection will take longer and will have higher costs compared with the collection of income indicators (Deaton & Grosh, 2000). Henry et al. (2003, p.170) refer to the potential ambiguity associated with the goods to be included in the reference basket when the study involves different countries, especially for the non-food items. The authors mention also the influence of the recall period and registering method in the accuracy of the data collected, with more precise methods implying usually higher costs. Zeller, Sharma, Henry and Lapenu, (2006, p. 449) add the requirement of advanced skills in statistics associated with the analysis of consumption data.

The indicators presented in these brief paragraphs do not cover all possibilities regarding the measurement of poverty, which corresponds to a much vaster research field. Not included here are, for example, the multidimensional poverty index (MPI) (Alkire and Santos, 2014) and the measures associated with Amartya Sen’s capability theory (Sen, 1983). The focus on this section was, however, on the simpler measures which have been most frequently used in the microfinance literature.

The limitations associated with the above referred indicators as well as the high implementation costs of more complex and rigorous poverty measurement and evaluation methodologies, have prompted a search for credible and accessible evaluation tools in a context of co-existence of institutions with different business models and dimensions and operating in different environments (Zeller et al., 2006). Poverty scorecards such as the Poverty Probability Index (PPI) are an “example of indirect poverty measurement methodology”, which seems a close fit for the needs of different types of institutions in the measurement of outreach and outcomes (Boucher, 2014, p.9).

4.2 Poverty scorecards in microfinance

This section presents the efforts conducted within the microfinance sector to develop simpler poverty assessment tools capable of measuring the depth of outreach of the microfinance institutions. The poverty scorecards presented here preceded PPI, which is presently the most used poverty scorecard in the sector.

Boucher (2014) summarised the expected advantages of this tool for the MFIs. In countries for which PPI has already been developed, MFIs have access to an externally validated instrument, which is easy to understand and implement, and to extensive support materials (manuals, case studies) at no monetary cost to the users. Furthermore, PPI not only allows
for the identification of the poverty level of the clients (outreach), but also tracking changes over time and comparing the results with similar institutions in the same country (outcomes).

One of the efforts to develop a poverty assessment tool happened in the late 90s. The CGAP, in collaboration with the International Food Policy Research Institute (IFPRI), aimed to develop a tool to improve transparency in the assessment of the depth of outreach of the microfinance programmes. The method used involved identifying a number of indicators associated with poverty that could be easily collected; developing a questionnaire-based assessment to be applied to a random sample of 300 non-clients and 200 clients living in the operational area of the MFI; and constructing a “single summary index”, using principal component analysis, which combined the information of the selected indicators and allowed for the comparison of the poverty levels of the client and non-client households (Henry et al., 2003, p.4).

Simanowitz (2003) reviewed the application of the method in seven countries, considering it to be effective in evaluating the depth of outreach of the MFIs in its geographic area of operations, but failing to allow for comparisons at national and international levels. The author also questioned the potential for dissemination of the tool given the implementation costs of the household survey, and the fact that it was designed as a one-off application being associated with a specific period. As a conclusion, the author suggested that the CGAP tool could be of use for donors interested in an external evaluation of the MFI’s outreach, but only at one point in time.

Zeller et al. (2006) focused on the pilot application of the method in four MFIs in Nicaragua, Kenya, Madagascar and India in 1999, describing the tool and the detailed results for each MFI. They considered the method to fulfil CGAP and IPFRI requirements by providing a measure of relative poverty, which was low-cost, flexible and applicable to local and national contexts and to all types of development projects. Taking into account that these institutions were included in the review conducted by Simanowitz (2003), it appears that different criteria was used in the appraisal of the tool in terms of the implementation cost and in its utility at national level.

An alternative poverty scorecard was developed for USAID by the IRIS Centre at the University of Maryland. The project started in 2003 and led to the development of PAT (Poverty Assessment Tool). The methodology, as with CGAP tool, is based on the design of surveys of 10 to 20 questions, but in the case of PAT, these questions are national specific. The scorecards were developed for 37 countries and can be accessed by the MFIs without
As USAID targeted mostly MFIs receiving funds from the institution, the dissemination of the tool among other MFIs and other countries did not seem to be a priority. The creation of PATs for new countries was halted by USAID in 2012. Consequently, PPI, initially known as the Progress Out of Poverty Index, has become the most widely used poverty scorecard.

It is important to note that in parallel to these institutional approaches to the development of poverty scorecards in the sector, there were also a small number of research studies focusing specifically on the development of tailored poverty measurement tools. This was the case of Dinh and Zeller (2010) who developed and tested a local tool (F2 tool) in the Northern Vietnam and compared it with the national-based poverty scorecards PAT and PPI in terms of accuracy.

### 4.3 Poverty Probability Index (PPI)

PPI was commissioned in 2005 by the Grameen Foundation with the support of CGAP and the Ford Foundation. It was developed by a team led by Mark Schreiner who had been previously involved in developing a pilot poverty scorecard to PRIZMA, an MFI in Bosnia-Herzegovina (Schreiner, Matul, Pawlak & Kline, 2004). In 2016, the Grameen Foundation has formed with IPA (Innovations for Poverty Action) the PPI Alliance, a “collective governance and funding structure” with IPA assuming the coordination of the tool. The following year, the designation of the tool was changed to the Poverty Probability Index, maintaining the acronym PPI (IPA, n.d.)

PPI scorecards have been developed for 61 countries and have been applied by a wide variety of organisations across different sectors (IPA, n.d.). The growth in the utilisation of the tool, associated with the reputation of its promoters, has enhanced the quantity and credibility of data potentially available to measure poverty outreach and has allowed for the definition of national benchmarks at this level. In this way, the use of PPI appears to present advantages not only at the practitioner level, as suggested by Boucher (2014), but also for researchers. It would, thus, be expected an increase in the production of academic papers, both theoretical and empirical, using PPI and the data generated by its application.

---

12. PPI is presented as a “poverty assessment tool for organisations and businesses with a mission to serve the poor”. Further information is available in the dedicated website (IPA, n.d.).
The PPI corresponds to a set of 10 questions relating to household characteristics and asset ownership, which are selected specifically for each country. As an example, the Pakistani scorecard is reproduced below in Table 4.1 (Schreiner, 2010).

Figure 4.1 – PPI Form for Pakistan (English version)

<table>
<thead>
<tr>
<th>Table 4.1. Simple Poverty Scorecard for Pakistan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entity</td>
</tr>
<tr>
<td>Member:</td>
</tr>
<tr>
<td>Loan officer:</td>
</tr>
<tr>
<td>Branch:</td>
</tr>
<tr>
<td>Household size:</td>
</tr>
<tr>
<td>Indicator</td>
</tr>
<tr>
<td>In what province does the household live?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>How many household members are 13 years old or younger?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>How many children ages 5 to 13 attend school?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>How many household members work in elementary occupations (not senior officials, managers, professionals, technicians or associated professionals, clerks, salespeople, service or shop workers, skilled workers in agriculture or fishery, craft or trade workers, or plant/machinery operators)?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>What is the highest educational level completed by the female head/spouse?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>What is the main source of drinking water for the household?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>What type of toilet is used by your household?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Does the household own a refrigerator or freezer?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Does the household own a television?</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Does the household own a motorcycle, scooter, car, or other vehicle?</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Source: Schreiner (2010, p.328)
The complex statistical methodology, including extensive testing, used in building the Pakistani scorecard is described in Schreiner (2010). It is similar to the methodology applied in most countries (apart from China) using data from the respective national household surveys (Schreiner, n.d.). The answers to the 10 questions included in the form are scored to compute the likelihood of a household living below the selected poverty line. The materials provided for each national PPI include look-up tables for the national poverty line and other internationally recognised poverty lines (e.g. the $1.25/day 2005PPP). The total scores range between 0 and 100, and the look-up tables provide for each score the probability of the household being considered poor (Pakistani summary look-up table is included in Annex II as example).

In between the utilisation of simple poverty proxies and the development of sophisticated impact studies, poverty scorecards are an alternative gaining attention from practitioners and academics in microfinance. These instruments allow for assessing both outreach and outcomes and, from a theoretical perspective, the methodology employed seems to fulfil its objectives, i.e. to enable institutions, even those with limited resources, to evaluate their performance (Boucher, 2014; Schreiner et al., 2004). However, by the time this research was developed, there were still a small number of independent empirical studies on the topic.

Research concerning the microfinance sector has been largely promoted and conducted by the promoters of the tools themselves. In the case of PPI, Mark Schreiner has produced an extensive literature on the methodology. A large part of the papers produced refer to the individual countries scorecards. These documents entitled “Simple Poverty Scorecard” were created for all countries, except China (Schreiner, n.d.).13 They contextualise the development of the instrument for the country; provide information on the data used and the building process of the scorecard as well as guidelines for its utilisation. The document for each country has been updated whenever the respective scorecard was reviewed. For some countries, this information was also the basis for papers published in academic journals such as Schreiner (2010) regarding Pakistan.

Other documents include Schreiner et al. (2004) which illustrates the process of building the poverty scorecard for PRIZMA. Schreiner (2014a) lays the theoretical foundations for the development of the scorecards, Schreiner (2014b) provides a comparison between PPI and PAT; and Schreiner (2015) studies the influence of different interview methods on PPI results.

13 For China, the document is entitled “An Expert-Based Poverty Scorecard for Rural China” and the difference derives from a different methodology used in the construction of the scorecard based on expert judgment instead of statistical data. The links for the papers on the countries scorecards can be found in Schreiner (n.d.).
The Grameen Foundation has published a report on PPI implementation in ten Philippine MFIs (Chua, Sebastian & Silva, 2012) and a global report on the implementation of the tool (Grameen Foundation, 2014).

Beyond Mark Schreiner and the Grameen Foundation papers, at the time the PhD project started there was an assessment of the validity of PPI using national data from Rwanda (Desiere et al., 2015); reports on PPI implementation in Peru and Ecuador (Boucher, 2014) and in Asia (Awais, 2010a; Awais, 2010b; Bhat & Yamini, 2012), and a research paper based on field work in Cambodia (Lanzavecchia, 2012).

During the period of the project, it was published an impact study implemented by Oikocredit (Dutch social investor)\textsuperscript{14} on two Asian MFIs using PPI as dependent variable (Gravesteijn, Hoepner & Jain, 2015), and more recently, Diamond et al. (2016) compared the performance of a national simple poverty scorecard (PPI) with regression-based estimators in the calculation of the poverty rates of a defined population.

### 4.3.1 Methodological issues and implementation challenges

The statistical process used in the construction of the individual country PPIs is not described in detail since the focus of the thesis is on the challenges associated with its implementation (in the context of the LWC evaluation project). The methodological issues are, therefore, addressed from an empirical perspective. The statistical process is extensively presented and discussed in Boucher (2014), Desiere et al. (2015), Schreiner et al. (2004) and all the individual “Simple Poverty Scorecard” documents mentioned above.

Desiere et al. (2015) assessed the validity of the PPI developed for Rwanda, with reference to the SMART criteria (Specific, Measurable, Available cost effectively, Relevant and Timely Available), using data from national household surveys for two periods (2005/06 and 2010/11). The authors concluded that the tool, by design, complies with four of the criteria, but with its Relevance being dependent on the model of implementation.\textsuperscript{15} They also asserted that the tool is accurate in the identification of poor households, however, they questioned its usefulness to track changes in poverty over time. In the tests conducted, two of the 10 indicators included in the index contributed to 80% of the variation observed in the poverty scores during the period. Given that this was a period of strong economic growth

\textsuperscript{14} Oikocredit has been one of the most active promoters of PPI in the sector, encouraging its funded MFIs to adopt the tool and providing training to guarantee the rigorous application of the methodology (Grameen Foundation, 2007).

\textsuperscript{15} The authors follow the definition of SMART proposed by the European Evaluation Network for Rural Development of the European Commission (Desiere et al., 2015).
and changes in the poverty rates in Rwanda, PPI became outdated and in need to be reviewed to maintain its relevance over time.

Lanzavecchia (2012) illustrates the importance of how PPI is implemented and analysed. In the project implemented in a Cambodian MFI, PPI was applied to a randomly selected sample of existing borrowers. The questionnaire was applied once, through in-house interviews in which retrospective information was also collected relatively to the period six months before. The data collected was used to compute PPI scores at the two moments in time. The data showed a slight improvement of the average poverty score for the sample, but the change was not statistically significant after a robustness test was performed. Lanzavecchia (2012, p.164) concluded that the PPI methodology is “misleading” on the account that it is too slow tracking changes in poverty over time.

A number of methodological issues can be raised regarding this study, starting with the very short period of time (shorter than a loan cycle) used to track changes of poverty levels among clients, and the use of retrospective data. In this particular case, it seems it would have been more appropriate to focus on the poverty outreach of the institution at the moment of the interviews, and not attempt to measure changes in poverty levels. Even then, as the reports from Awais (2010a; 2010b) also suggest, the information would not be very useful if not complemented with analysis of other elements, including data collected by the MFI during the application process.

Bhat and Yamini (2012) is based on the MicroSave experience piloting the use of PPI in MFIs in India, Sri Lanka and Bangladesh. The authors did not explore the results obtained from the application of the tool, focusing instead on the implementation process. They mention two relevant questions in the utilisation of PPI - cost and property. While the access to the questionnaires and support materials is free of cost, PPI implementation has associated training and opportunity costs (or financial costs if the implementation is outsourced), that may not be insignificant, especially for smaller MFIs.

Equally important is the intellectual property of the tool. MFIs do not own the methodology - it was a Grameen Foundation, and now IPA, copyrighted tool. Thus, changes and updates can only be introduced by the promoters, making it more difficult to customize the tool to specific local contexts (Bhat & Yamini, 2012). If the MFIs have internalized the advantages of consistently using poverty scorecards, then there will be an incentive to develop internal competences on the scorecards and update them, ensuring that they are credibly used over time.
Boucher (2014) is a report conducted for the Inter-American Development Bank on the implementation of PPI in Peru and Ecuador. It includes theoretical considerations related to the design of the instrument and the description of the different implementation models adopted by the MFIs, discussing its advantages, limitations and challenges.

The author praises the theoretical robustness of the instrument and the transparency of the promoters of the tool regarding its limitations. Nonetheless, he raises concerns about the time consistency of PPI and the validity of the instrument to measure poverty for sub-groups of the population, e.g. urban and rural clients (Boucher, 2014). The accuracy of PPI is expected to decrease with time (Schreiner, 2014a) so updating the questionnaire becomes particularly relevant. However, updating the scorecard is not an automatic process; it is dependent on the availability of new national survey data and implies costs associated with the re-design and test of the questionnaire.

The challenge of dealing with sub-group differences is more complex. PPI scores are based on the relationship between poverty levels and a group of selected indicators, but these relationships may vary across sub-groups (e.g. across different districts or regions within a country). As the scores are computed based on national averages, this may lead to an overestimation or underestimation of the likelihood of being poor for a household belonging to a specific sub-group. Boucher (2014) gives the example of Peru to illustrate significant poverty level differences between regions, but this is also the case in Pakistan.

In the same direction, Diamond et al. (2016) found that a simple poverty scorecard was as accurate as regression-based estimation models when analysing national poverty rates and using nationally representative samples, but underperformed in more specific settings, including sub-groups of a population. In these cases, there is a trade-off between simplicity of use and accuracy. One of the reasons which may have contributed to this finding is the fact that the simple poverty scorecards only use the information obtained from the ten questions, even if the researcher has access to additional data.

Regarding the field implementation of PPI, Boucher (2014) highlighted two main challenges. The first, similar to Bhat and Yamini (2012), is the cost for the MFIs. The implementation of the questionnaires and the data analysis have associated costs that may differ according to the implementation model chosen. The MFI may decide to train their own staff to collect PPI data; they may integrate the questionnaire in their loan application process or apply it separately; and they may apply it to all their clients or to a selected sample of clients. In all these cases, training costs and time-related opportunity costs must be considered.
Alternatively, the institution may contract an external party to implement the questionnaire and analyse the data, which entails financial costs.

The second challenge identified by Boucher (2014) derives from the incentive to use PPI. There can be two main sources of motivation to apply PPI: internal and external. Internal motivation is associated with a desire to integrate PPI data in the MFI’s internal management process and use it to improve the social performance of the institution. It guarantees that the resources needed for implementation will be available and provides a commitment to the continued use of the instrument. The main challenge in these cases is to buy-in the staff involved in the process to guarantee rigour in the application since the PPI questions are normally incorporated in the loan approval process. The data obtained is used in management decisions.

External motivation is usually linked to reporting poverty outreach to outside stakeholders. In many cases the utilisation of PPI is pushed by these external parties, often international donors or investors such as USAID. In these situations, moral hazard may arise as there can be an incentive for the institutions to over-state the poverty levels of their clients to gain access to donations or cheaper credit. In addition, it may lead to less commitment in the application of the questionnaire, with consequences on the rigour of implementation and the quality of the poverty estimates (Boucher, 2014).

The five MFIs included in the report represent different approaches regarding the questions described above, representing different models of implementation of PPI (Boucher, 2014). They will be used as reference in the discussion of the case studies presented in chapters six and seven.

4.3.2 Empirical evidence

A common critique of the empirical studies using poverty scoring to measure outreach is that the MFIs studied are not reaching the poorest segments of the population, even when this has been the declared objective of the institution (Schreiner, 2014a). Examples include Coprestake et al. (2005), Ghalib (2013), Hulme and Mosley (1996a), Lønborg and Rasmussen (2014), Microfinanza Rating (2012) Ramanathan, Agaba and Hwang (2015) and Simanowitz and Walter (2002).

Ramanathan et al. (2015) describes the first application of PPI by a Ugandan MFI, and the subsequent management decisions to foster the use of PPI in the institution. In the pilot, a sample of 1,070 clients from two branches was found to have an average likelihood of being
below the national poverty line that was lower than the national poverty average, which came as a surprise for the MFI managers at that time.\textsuperscript{16}

Ghalib (2013) shows that for rural areas of the Lahore province in Pakistan, the depth of outreach was considerably lower than what was expected by the financial service providers participating in the study. Lønborg and Rasmussen (2014) provides similar evidence for a VSLA (Village Savings and Loans Association) programme in Northern Malawi where they found regressive targeting (i.e., participants in the microcredit schemes were less poor than the general population in the community). In their paper, the authors highlight the limitations and constraints associated with outreach measurement methodologies, namely the data collection timings and the metrics used.

Microfinanza Rating (2012) in a cross-country study found that only 19\% of the clients of the 65 MFIs in the sample were likely to be poor, although there were significant differences among regions (Africa and Asia with larger outreach than Latin America/Caribbean, and Europe/Central Asia), legal status of the institutions (NGOs and non-deposit taking institutions) and MFI mission (poor oriented).

Awais (2010a; 2010b) seem to be an exception. The results of the pilot implementations of PPI in four MFIs in Sindh, Pakistan and two in Nepal, show that the percentage of clients likely to belong to a poor household was close to the national poverty rate averages (using the respective national poverty lines). However, in the Pakistani report, while the MFIs’ poverty rates were relatively close to the national average, they were lower when compared to the figures for the Sindh province.

The lower (than expected) depth of outreach found in this range of studies can arise for a number of different reasons. Hermes and Lensink (2011) summarized the potential causes: self-exclusion of the very poor who lack confidence to approach the institutions and doubt their capacity to repay the loans; the perception of a higher risk of debt associated with the very poor which leads to them not being accepted as group members by other borrowers or being refused by the MFI staff; and the design of the programmes, particularly requirements such as saving before being granted the loan, or presenting an identification card can become insurmountable obstacles.

\textsuperscript{16} This is a case study published by the Grameen Foundation, which also published three other case studies and a regional report on MFIs using PPI to measure poverty outreach (Crowther, 2015; Grameen Foundation, 2011a; 2011b; 2011c); Additionally, the Grameen Foundation India carried out similar studies in four Indian states (Grameen Foundation India, 2015) and in the state of Karnataka (Grameen Foundation India, 2016).
Gravesteijn et al. (2015) adopted a different perspective since PPI is not used to measure poverty levels of the clients of the two MFIs involved in the research in India and the Philippines. The authors employ a fixed-effects regression model aiming to study the effect over time of microfinance in the poverty levels of the individual clients, in which PPI is used as the dependent variable. They conclude that the access to the loan had a small but significant positive effect, contributing to reduce poverty among the microcredit clients.

From the review of the empirical literature on PPI, a common conclusion can be drawn in respect to the risks of one-time applications of the tool. These risks help explaining the reluctance of some MFI managers to introduce the use of poverty scorecards as an outreach measure, especially if they were to claim the MFI is targeting poor clients. Therefore, the application of PPI should not be treated as an isolated event aimed to give an absolute ‘answer’, and it should be combined with other sources of information as suggested by Gravesteijn et al. (2015).

The preparation and consolidation over time of the use of PPI seems fundamental for the institution to understand the advantages and limitations of this type of instrument. In this sense, even if the main motivation for measuring outreach is in many cases external (Boucher, 2014), MFI managers need to be persuaded (and convince their staff) of the utility of the instrument and adopt the application model that best suits the organization. Only then, they will be able to use the poverty scorecards to improve the MFI’s social performance management rather than being a simple client poverty measurement tool (Schreiner, 2014b).
5 Methodological Issues

This chapter presents the methodologies used in the research project, presenting the rationale for the two case studies developed in Pakistan and Zimbabwe (section 5.1) as well as the data collection methods employed in the project (section 5.2). The third section (5.3) describes the framework of analysis based on Michael Patton’s utilisation-focused evaluation (UFE) model. Section 5.4 introduces the Lendwithcare (LWC) assessment project, presenting its objectives, participants and methodologies to contextualise the discussion on the evaluation challenges and lessons learnt by the partners.

5.1 Case Study

The research conducted aims to study two broad questions: what are the main challenges in the implementation of impact evaluation projects in socially focused MFIs with limited resources, and how does a common project approach to evaluation influence the institutions to engage in the process and integrate evaluation as a component of a social performance management system? These questions can be broken down into a number of more specific ‘why?’ and ‘how?’ type questions, e.g. Why did the different stakeholders decide to participate in the evaluation? How did they perform in the different stages of the project? How did they perceive the results (and intend to use them)?

The nature of the research questions was, therefore, one of the main reasons to opt for a case study research design. Another important factor was the fact that the researcher did not have control over the object of study. Although participating in the evaluation project, the role of the researcher was mainly one of consultancy and offering advice to the leader of the project (LWC), who ultimately took the decisions regarding the different stages of the evaluation.

In this way, the project fitted perfectly with the case study approach to research as proposed by Robert Yin: “A how or why question is being asked about a contemporary set of events over which the investigator has little or no control” (Yin, 2018, p.13).

One of the areas in which the researcher had no control was the choice of research settings. From the PhD research perspective, the goal was to work with institutions with an explicit social mission, limited resources, and preferably located within complex contexts. Having the possibility of studying challenging and distinct contexts was expected to put in evidence the influence of contextual differences in the implementation of the impact project and test the application of the common project approach, which will be further described and discussed in chapters five and eight.
The institutions working with LWC, including those chosen to participate in this first phase of the evaluation, all have strong social missions, and, to different degrees, also have limited resources.\textsuperscript{17}

Moreover, the two selected settings can be surely described as complex contexts. Pakistan is mired by security and political issues and registers very low levels of literacy and financial inclusion (Lieven, 2011; World Bank, 2018). Zimbabwe has long been affected by political instability and a precarious economic situation (Tyson, 2017), which has been reflected in the evaluations conducted in the microfinance sector. Goldberg (2005) emphasized the relevance of contextual factors, namely the hyperinflation experienced during the survey period and its effects, in the results of the AIMS study (Barnes, 2001). In the case of Hulme and Mosley (1996a), Zimbabwe was one of the five countries initially included in the research proposal presented to the ODA, but was excluded from the project since the “1992 drought made Zimbabwe a difficult environment for research” (p. xiv).

The implementation of the case studies in these two contextually different settings provides therefore a good illustration of arguments made by Pawson (2014) considering the circumstances in which an intervention occurs as an endless source of complexity.

Also supporting the use of a case study approach is the time-span of the research of more than three years. This period allowed for the development of a relationship with the partners, which facilitated the preparation of rich descriptions of the implementation of the impact project in the two settings (Lincoln & Guba, 2000). Finally, piloting the evaluation project in two MFIs allowed for an initial comparative study and a subsequent analysis of the effects of applying a common project approach to evaluation in different settings (one of the intended purposes of LWC).

The case studies developed in the thesis can be seen as instrumental (Stake, 1994), given that they offer insights into a particular issue needing refinement - in this instance, the field challenges associated with the evaluation of microcredit programmes and, in parallel, the implementation of a common project approach to evaluation in different contexts. The expectation is that the research findings will either discourage the adoption of a common project approach by LWC or reinforce its intention to extend the project to other partners.

\textsuperscript{17} Akhuwat Islamic Microfinance in Pakistan has grown significantly during the period of the project, becoming one of the largest microfinance institutions in the country, and has attracted increasing inflows of funds (Khan, Ishaq, Afonso & Akram, 2017). However, these funds are in most cases donations constrained to specific uses.
Case studies naturally have limitations in the same way as any other methodology. Case studies have often been considered to be lacking in rigour and objectivity (Rowley, 2002), or to focus excessively on the description of the cases to the detriment of collecting sufficient judgmental information (Stufflebeam, 2001). These concerns have led some of the most prominent advocates of the methodology, namely Robert Yin, to be prescriptive in the proposition of strategies, methods and techniques to be used in the different stages of a case study in order to increase its rigour and, thus, strengthen the internal and external validity of its findings. These include developing the case study database, relying on theoretical propositions, examining plausible rival explanations, and using multiple sources of evidence and analysis techniques such as pattern matching logic, explanation building and cross-case synthesis (Stufflebeam & Coryn, 2014; Yin, 2018).

The major, and probably most discussed, challenge in case study research regards the external validity of conclusions, with opposing views on the subject being proposed by authors from different epistemological schools. From a positivist perspective, “generalizations are assertions of enduring value that are context-free” but, in reality, human activities are hardly context-free (Lincoln & Guba, 2000, p.27). Lincoln and Guba (2000), proponents of the constructivist evaluation model, consider that the differences in local conditions make it impossible to achieve generalisations in its classic form. They suggest a more modest approach, in line with Cronbach’s (1975) idea of generalisation as a working hypothesis. To these authors, the degree of transferability of the case study findings to other contexts depends on the degree of similarity between the settings, which requires the person making the judgment to have enough information on both contexts. It is, thus, fundamental that the researcher provides “thick descriptions” of the case, including “everything that a reader may need to know in order to understand the findings” (Lincoln & Guba, 2000, p.40).

Considering the focus of my research on the processes associated with the LWC impact assessment project (developed first in Pakistan and replicated in Zimbabwe), the research can be described as applying a multiple-case replication design, with each MFI corresponding to a unit of analysis (‘AIM’ in Pakistan and ‘THRIVE’ in Zimbabwe) (Yin, 2018, p.58). Each unit of analysis represented an individual case in itself and it was analysed independently in what concerns the thesis first research question. The cross-case analysis was used to test the theoretical hypothesis embedded in the second research question, i.e. the advantages of the project common approach (as defined in chapter eight, p. 185) in the encouragement of MFIs to implement/develop their own social performance management systems. In order to

---

18 Improving the potential for generalisation of case study findings has been addressed by several authors, including Gomm, Hammersley and Foster (2000) and Schofield (2000).
strengthen the results of the cross-case analysis, it would be important to replicate the approach in other contexts (e.g. other LWC field partners) as this would allow to corroborate (or not) the initial findings presented in the thesis in chapter eight.

The following chapters of the thesis describe and discuss the implementation of the impact assessment project in the two individual cases (chapters six and seven) and thereafter present the results of the cross-case analysis (chapter eight).

5.2 Data Collection: the central role of participant observation

The PhD research design is based on qualitative methodologies, which is reflected in the chosen data collection techniques, including participant-observation, formal semi-structured interviews, informal conversations and documentation. While all these techniques were important to ensure the use of multiple sources of evidence, participant-observation played a central role in the research.

Participant observation emerged from the field of ethnography. Atkinson and Hammersley (1994, p.248) define participant observation as “observations carried out when the researcher is playing an established participant role in the scene studied”. The researcher “wears two hats”: one related to group membership, with its associated obligations, and the other as a researcher studying the group from the inside.

While applications of the method in anthropology studies often imply an intensive immersion of the researcher in the study settings as a participant (e.g. Bourgois and Boelstorff works cited by DeWalt and DeWalt, 2011, p.14), the use of the method within this PhD project is more modest. The researcher was an active member of the evaluation team and a project stakeholder aiming to ensure a useful and high-quality evaluation as advocated by Chen (1990, p.77). However, given the resources available and the decision to involve the field partners in the implementation of the impact project, the fieldwork conducted by me as evaluator in Pakistan and Zimbabwe was limited to short periods of time (five to ten days in two field visits to each country). These field visits were complemented by meetings of the partners in the United Kingdom and an extensive exchange of email messages and Skype conversations throughout the period of the project. From the PhD research perspective, London and Portsmouth were as much stages of the LWC evaluation project as Lahore and Harare.

The approach followed in the PhD research is similar to the one adopted by Gravesteijn (2014) in his doctoral thesis, although the subjacent project, the research questions and the contexts explored in the two theses are significantly different. In both cases, the researchers
were “able to gain access to an appropriate setting” (Jorgensen, 1989, p. 13), with Gravesteijn working as a consultant in the previously referred ILO-led project ‘Microfinance for Decent Work’. In his PhD research, the author focused on two MFIs in Central Asia to study, in parallel with his own participation as consultant, how each MFI managed their participation in the project, and to what extent could an international organisation such as ILO have a long-term influence in the MFIs’ performance management (Gravesteijn, 2014).

The two theses have in common the attention given to the field partners, namely to how their managers perceived their participation in externally-led projects, how they have used the outcomes of these projects and how they have articulated with the external project leader (in both cases, the funders of the research being implemented). There is, however, an important distinctive factor between the two studies which relates to the time spent in the two countries in the implementation of the respective evaluation projects, much longer for Gravesteijn which gave him, comparatively, more in-depth knowledge of the two microfinance institutions involved and more opportunities to observe and interact with the MFIs managers, staff and clients. In the comparative discussion of the cases (chapter eight), I return to this research and the similarities and differences between the two studies.

Jorgensen (1989) suggested that participant observation as methodology is, in general, practiced as a form of case study, including detailed descriptions and in-depth analysis of the case. Yin (2018, p.124) recognises the “unusual opportunities for collecting case study data” created by participant observation. It provides access to events and people, the possibility to perceive reality from an insider viewpoint and the ability to manipulate minor events such as arranging meetings with persons of interest to the case. Indeed, all these advantages can be identified in the research experience reported in the thesis and detailed in the following chapters.

A complementary perspective is offered by DeWalt and DeWalt (2011, p.19). The authors, in their guide for fieldworkers, attribute the relevance of participant observation to its potential to enhance the quality of the data collected and the interpretation of these data, as well as the encouragement for a continuous re-examination of the initial research questions and hypothesis, eventually leading to the development of new research questions. In the same direction, Keiding (2011, p.108) stresses how the direct involvement of the observer “produces opportunities to test understandings and meaning attribution”.

It is important, however, to concede that this potential strength of the method (the access and quality of the data collected and its interpretation) can become its weakness, with several authors calling attention for the possibility of bias in both recording and interpreting
data (e.g. Murray and Lawrence, 2000; Yin, 2018). Participant observation is a complex and challenging method - while participation requires emotional involvement, observation is associated with detachment, and the tension between the two should be acknowledged by the researcher (DeWalt & DeWalt, 2011). In participant observation, the observer/researcher interacts with others, and in this process, brings his/her own background and experiences into the situations. In this sense, the observer is also a co-producer of the observed interaction (Keiding, 2011, p.108).

The successful use of the method is dependent on the ability of the researcher to critically reflect on his/her own mental model and agency regarding the research (DeWalt & DeWalt, 2011). Reflexivity is, thus, a central concept and practice in action-research and, in general, qualitative research (Atkinson & Hammersley, 1994; Gravensteijn, 2014; McGee, 2002; Reason & Mc Ardle, 2004), requiring that the researcher questions the multiple selves that she/he adopts in the conduction of the research project (Lincoln & Guba, 2000). This recognition of different ‘selves’ translated into the presentation of the PhD research and this thesis through a combination of use of the third person when referring to the partners or other internal or external stakeholders with the first person (I, me,) in the description and discussion of the methodology adopted in the research.

In the LWC assessment project, despite the evaluation team being relatively small, there was great diversity among its members regarding several characteristics, including age, gender, ethnicity, religion, language, educational and professional background and experience in the microfinance sector. Adding to this, the interactions between the members occurred in very different settings (UK, Pakistan, Zimbabwe). From the start, the project was a “cross-cultural encounter” (Patton, 1985), which required me to be aware of the cultural context, and to be self-reflective in regard to values and assumptions (Sengupta et al., 2004).

These two previous paragraphs refer to the two functions of reflexivity in participatory research put forward by McGee (2002, p.21). One is related to the “self-critical monitoring” of the application of the research methods, which relates to the tensions at personal level stemming from the dual role as participant and observer. The other respects to the development as researcher of “non-judgemental attitudes and non-hierarchical behaviours” towards other participants, particularly those based locally. In the description of the case studies, examples of some of these situations will be included as well as the strategies I have used to overcome or minimise these issues.

The difficult balance between the two roles played by the researcher, implying that the participant role may require too much attention, was one of the challenges of participant
observation suggested by Yin (2018, p.124). As a consequence, the amount of time dedicated to the observer role and, consequently, the possibilities to take notes and raise questions from different perspectives (data triangulation) may be reduced.

This was a manageable, but still significant, challenge in the PhD project. Although occupying a consultancy and advisory role, I have been involved in all stages of the process, which had implications in terms of time and focus. My inexperience in the implementation of an evaluation project of this nature led to needing more time than initially expected to prepare and implement the different tasks associated with the evaluations.

In this brief incursion into participant observation, the importance of explicit recording and analysis of the data collected should be emphasized. Observations should not be considered as data unless they are recorded in some form. For this purpose, accurate and detailed field notes are particularly relevant. As the researcher chooses what to include in the field notes and its level of detail, these can be seen as a construction of the researcher and part of the process of analysis (DeWalt & DeWalt, 2011).

During the participation in the LWC evaluation project, I have taken field notes in all events involving the evaluation partners. In the field visits, these were handwritten notes taken (as much as possible) during the several activities and at the end of each day. I have reviewed and summarised these notes into field visit reports, which included also some of my reflections on the issues explored in the notes. These summaries are stored in the University server, along with the interview transcripts and other documentation of the project. Within the stored data, the names of the evaluation team members were omitted and replaced by references to the institution for which they work. This procedure intends to protect (as far as possible) the identities of those involved in the project.

The field visits to Pakistan and Zimbabwe, while motivated by the LWC evaluation project (particularly the preparation and implementation of the household surveys), also allowed me to apply other data collection techniques, including pure observation, formal/informal conversations and review of internal documentation. These opportunities were facilitated by the fact that representatives from LWC participated in all of these visits in order to conduct their annual evaluation process, which includes visits to different branches and collection of client testimonies. The contacts with clients and staff, as well as the opportunity to observe daily routines in different branches of the institutions, were especially important for me to gain a better understanding of the culture and context of each MFI, and to identify similarities and differences between the two research settings.
The review of literature and internal documentation on the institutions, the informal conversations and exchange of email messages, and the formal partner meetings were critical sources of information on the perceptions of those involved regarding the project and the challenges faced during its implementation.

In regard to the literature review, it included the literature presented in the chapters two to four of the thesis, which starts with a more theoretical approach to evaluation (chapter two) and then focus on impact evaluation and the use of poverty scorecards in the microfinance sector (chapters three and four). It included also academic papers, technical reports and news/social media articles on the institutions and the countries context, which were included as references in chapters six and seven.

The literature review was conducted using academic search engines through the University of Portsmouth library portal, but also more generic search engines such as Google Scholar and sector/topic-related publications databases, including the Findev library, the World Bank publications and the 3ie resources. The initial search was based on generic terms such as ‘impact evaluation’ and ‘evaluation theory’, with more refined and sector-based search terms being used as the research advanced.

It should be noted that the review of papers on evaluation theories and, especially, on impact evaluation in the microfinance sector was selective. Taking into account the large number of publications related to evaluation and impact and time constraints, I have focused the analysis mainly on the significant number of systematic reviews and reports on the topic as these documents allow for an overview of the academic literature produced.

Given that most of these reviews and reports adopt a positivist approach towards impact evaluation, they lay emphasis on the methodological aspects related to causality and attribution of impact. This was a limitation of this strategy that I have tried to overcome with the inclusion of other perspectives on impact through documentation related to projects such as Imp-Act and the ILO ‘Microfinance for Decent Work Project’ (see pp. 57-58), and to the work developed at practitioner level by SPTF (see p.25).

Finally, structured interviews were conducted in the final stages of the project, with the objective of documenting the perspectives of the different evaluation team members regarding the evaluation process. Therefore, the members of the evaluation teams from the two MFIs (AIM1, AIM2, THR1 and THR2) and LWC (LWC1 and LWC2) were interviewed, with the exception of the temporary member of LWC (LWC3) who was no longer in the institution by the time of the interviews.
The transcripts of these interviews, which were validated by each of the respondents, are included in Annex I. The interviews with the LWC and the AIM evaluation teams were conducted face-to-face. This type of arrangement was not possible with THRIVE representatives, thus, the interviews were conducted through a combination of written responses and video-call. In all the interviews, the list of questions was sent to the participant in advance. During the interviews, I adopted a neutral position, not reacting or elaborating much on the answers. These strategies were deemed appropriate given the existing working relationship between me and the interviewees, as well as my active participation in the evaluation project. A combination of these factors could lead to reluctance to participate in the interview and would inevitably be a potential source of bias in the answers.

Despite these inherent limitations, the interviews were important to clarify the perceptions and understandings of the project partners on specific themes/topics related to the PhD research questions, complementing and, in most cases, corroborating the data collected from other sources. Although the question guide for each interview was different, considering the evaluation settings and the roles assumed by the interviewees, there were common themes and some common questions.

Table 5.1 summarises the main themes included in the interviews (coded using NVIVO) and relevant topics defined within the codes, identifying the main quotations and respective interview for each. The table gives a broad overview of the main arguments developed in the discussion in chapters six, seven and eight, where the quotations were integrated in the text as appropriate, being identified as personal communications (p.c.).

---

19 The interviews with the Pakistani representatives of AIM were conducted in March 2018 in London, taking advantage of their participation in the workshop organised by LWC for its partner MFIs.
<table>
<thead>
<tr>
<th>Theme (Code)</th>
<th>Topic</th>
<th>Arguments</th>
<th>Quotations (Interview)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motivations</strong></td>
<td>Confirmation</td>
<td>Obtain evidence to legitimate work, both externally and internally</td>
<td>“reveal the extent to which training and micro-credit were impacting livelihoods as an intervention” (THR1) “people need some evidence”; “we need this evidence, not only for the researchers and external parties, but for the staff also and for the board of directors - that we are here on a very strong social mission” (AIM1) “as practitioners, we often are more involved in the actual implementation, but we don’t step back enough and see whether the work that we are doing is having positive impact or not” (LWC1) “what is the impact of our loans on the poverty levels of the clients and their households” (LWC2)</td>
</tr>
<tr>
<td></td>
<td>Exploration</td>
<td>Learning process about the clients and poverty context Capacity building on the evaluation process</td>
<td>“Thrive is a learning organisation”; “understand what characterises poverty in the context of its operating catchment area and broadly Zimbabwe” (THR1) “we have this mission of providing technical assistance”; “helping them to integrate better social performance management procedures into their routine operations” (LWC1)</td>
</tr>
<tr>
<td><strong>Challenges</strong></td>
<td>Macro-environment</td>
<td>Influence of external factors in the lives of the clients and the implementation of the microcredit programmes</td>
<td>“absence of a standard poverty assessment tool” (THR1) “the macro-economic difficulties during the test period set a very low upper limit in terms of what any intervention could do to improve livelihoods” (THR1) “For dynamic environments such as those prevalent in Zimbabwe (...) The politics and as a result the economics are always changing, poor livelihoods respond sharply to these changes and so do outcomes” (THR1)</td>
</tr>
<tr>
<td></td>
<td>Partners</td>
<td>Time available for the project</td>
<td>“main challenge for us has been finding the time. Our partners finding time, and particularly us as a team within LWC finding the time” (LWC1) “it was time-consuming, it took a lot of time and I needed to review every aspect of the questionnaire” (AIM2)</td>
</tr>
<tr>
<td>Distance and cultural differences</td>
<td>Distance and cultural differences</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“During the training of the Interviewers I would dedicate 75% of my work time to the study. During the first week of the interviews 60% of my working time was also dedicated to the study” (THR2)</td>
<td>“During the training of the Interviewers I would dedicate 75% of my work time to the study. During the first week of the interviews 60% of my working time was also dedicated to the study” (THR2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“three different parties, all placed in different parts and different countries”; “try to put together those initial questionnaires that also kind of met the requirements of different people” (LWC2)</td>
<td>“three different parties, all placed in different parts and different countries”; “try to put together those initial questionnaires that also kind of met the requirements of different people” (LWC2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>“sometimes the MFIs find it difficult to communicate with researchers from outside, once they identify the researchers have particular assets or insights I think that facilitates the collaboration for them” (LWC1)</td>
<td>“sometimes the MFIs find it difficult to communicate with researchers from outside, once they identify the researchers have particular assets or insights I think that facilitates the collaboration for them” (LWC1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data entry &amp; validation</th>
<th>Measurement errors and inconsistencies in data entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>“there are some technical mistakes”; “we are not 100% sure they [non-customers] are giving true information”; “some interpretations were not as explanatory as they should be” (AIM2)</td>
<td>“there are some technical mistakes”; “we are not 100% sure they [non-customers] are giving true information”; “some interpretations were not as explanatory as they should be” (AIM2)</td>
</tr>
<tr>
<td>“interviewing of the same clients twice in some instances” (THR2)</td>
<td>“interviewing of the same clients twice in some instances” (THR2)</td>
</tr>
<tr>
<td>“inputting the data (…) we used like an agency which then was a bit of a challenge in terms of analysing the data” (LWC2)</td>
<td>“inputting the data (…) we used like an agency which then was a bit of a challenge in terms of analysing the data” (LWC2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control group &amp; attrition</th>
<th>Constitution of the control group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location and participation in the second wave of the surveys</td>
<td>Location and participation in the second wave of the surveys</td>
</tr>
<tr>
<td>for non-customers which was not as easy”; “most of the clients who were in default would not avail themselves for the interviews” (THR2)</td>
<td>for non-customers which was not as easy”; “most of the clients who were in default would not avail themselves for the interviews” (THR2)</td>
</tr>
<tr>
<td>“and we have to deal with the control group”; “the biggest problem will be where the non-customers are. We haven’t written their proper address” (AIM1)</td>
<td>“and we have to deal with the control group”; “the biggest problem will be where the non-customers are. We haven’t written their proper address” (AIM1)</td>
</tr>
<tr>
<td>“issue was only with those clients we have not given a second loan because they had not repaid their first loan quite well. So, they were the problem, they don’t want to answer because they were angry with our staff” (AIM2)</td>
<td>“issue was only with those clients we have not given a second loan because they had not repaid their first loan quite well. So, they were the problem, they don’t want to answer because they were angry with our staff” (AIM2)</td>
</tr>
<tr>
<td>“the control groups could have been in higher size, larger and we should have made more effort to chase them for the second round of interviews” (LWC1)</td>
<td>“the control groups could have been in higher size, larger and we should have made more effort to chase them for the second round of interviews” (LWC1)</td>
</tr>
<tr>
<td>“control group was really hard to find, to work out how to find them, where to find them, whether we pay for them to take part or we don’t pay for them, and then obviously in the subsequent round of interviews tracking them” (LWC2)</td>
<td>“control group was really hard to find, to work out how to find them, where to find them, whether we pay for them to take part or we don’t pay for them, and then obviously in the subsequent round of interviews tracking them” (LWC2)</td>
</tr>
<tr>
<td>Strengths</td>
<td>LWC relation with MFIs</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Buy-in from partners</td>
<td>Strong commitment from the beginning and during all process</td>
</tr>
<tr>
<td></td>
<td>Change in perception regarding evaluation process</td>
</tr>
<tr>
<td></td>
<td>very keen to take part”; “very quick to reply”; “the will of everybody who was taking part for it to work” (LWC2)</td>
</tr>
</tbody>
</table>
| Results | Replicability of the process | “there would be a lot of things that might be learnt (...) that we could probably take on and adopt in future evaluations” (LWC1)  
“definitely I think it is replicable” (LWC2) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>From Results</td>
<td>Dealing with negative results</td>
<td>if we are seeing potentially not positive changes happening, what does that mean for CARE as an investor in these microfinance institutions” (LWC2)</td>
</tr>
<tr>
<td>From Results</td>
<td>Further analysis of the results</td>
<td>“opportunity to kind of dig a bit deeper into some of the things that come out of the impact study (...) the opportunity to delve into that question of women’s economic empowerment a bit further” (LWC2)</td>
</tr>
</tbody>
</table>
| Capacity building | Development of skills internally on the evaluation process | “On a personal level it has widened my perspective on how a big study can be done in practice” (THR2)  
“have learned loads and I am actually going to do a monitoring and evaluation course” (LWC2)  
“What is possible, what isn't possible, what's feasible, what isn't feasible? What are the challenges, what are the weaknesses? And how can we incorporate this into our day to day routine work?” (LWC1) |
| Results | Perceptions on evaluation results | “the findings have been really interesting”; “has been really useful insight to see potentially not so positive changes and potential reasons why that might be happening”; “we were all pleasantly surprised and pleased by some of the changes we seem to be seeing in Pakistan over a relatively short period of time” (LWC2)  
“both sets of results are positive. I think they were not as positive in Zimbabwe, but that’s probably largely due to the macroeconomic and political environment that’s been afflicting that country for the last few years. And, in fact, you could probably say, despite everything, the results in Zimbabwe are not negative” (LWC1)  
“disappointed with the overall result, but we are not surprised” (THR1)  
“we are thinking that the results are quite satisfactory right now, but we want more improvement than that” (AIM1) |
<table>
<thead>
<tr>
<th>Uses</th>
<th>Sharing findings</th>
<th>Accountability purposes – evidence to supporters, funders and other external parties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“research report can then be shared as widely as possible and the report can refer to Thrive as openly as necessary in so far as it shouldn’t unnecessarily affect reputation and business prospects” (THR1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“our main objective was to get evidence and show to the people, so we would love to share as soon as possible with government officials (…), we will share these results of the impact assessment with international forum. We will share these results with academia, researchers, scholars, PhD students” (AIM1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“we disseminated first and foremost to our supporters.” “How do our partners use them? Again, they disseminate it to their investors and supporters” (LWC1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>“we’ve been writing, we’ve been presenting at conferences and that’s been happening in a whole range of levels” (LWC1)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Changes (effective and aspirational)</th>
<th>Development of social performance assessment measures/systems in the MFIs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“making Thrive more effective we hope by broadening our understanding of poverty”;</td>
</tr>
<tr>
<td></td>
<td>“contribute to the general knowledge body, reveal areas of future study and improve future programming by Thrive and anyone else with access to the study” (THR1)</td>
</tr>
<tr>
<td></td>
<td>“we did not think to change any system, we have not taken any decision, because the results are so much satisfactory for us right now”; “we are going to a paperless environment (…), what we will try to do is we just try to implement all this set of questions into a paperless environment where we can get some kind of reports through our MIS” (AIM1)</td>
</tr>
<tr>
<td></td>
<td>“we will introduce some questions from questionnaire and we will analyse this data in the future”; “we are going to use our own credit ratings; we are going to work on different aspects to measure the impact of our loans in quite better way” (AIM2)</td>
</tr>
<tr>
<td></td>
<td>“[PPI] allows us to compare and certainly allow our partners to compare perhaps with peer organisations in the context of the countries where they work as well”; “the research has a use in itself, the research has a use in terms of giving us results that we can disseminate to a wider audience, almost to justify our results. But the third thing is that the research begins a process within our partners as well” (LWC1)</td>
</tr>
</tbody>
</table>
|                                    | “as a team we are investing more money into impact, measuring and monitoring impact, scaling up that side of our work”; “now we need to think what we’ll be doing with that learning and the actions that come as result of it”; “to take the opportunity to kind of dig a
<table>
<thead>
<tr>
<th>Rigour</th>
<th>Legitimacy &amp; Credibility</th>
<th>Standards of Rigour</th>
<th>Field Compromises</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Collaboration with University as source of credibility – impartiality and independency</td>
<td>“research is the motivation for us to induce the latest techniques of measuring impact at our level, instead of waiting for universities ... We should also be aware, it may not be as correct as research of an independent researcher, but we must have an idea of where we are moving” (AIM2)</td>
<td>“Most of the first interviews for clients were conducted at the branches” (THR2)</td>
</tr>
<tr>
<td></td>
<td>“a reputable development partner and an academic/research institution, your findings are obviously authoritative and can be accepted as credible” (THR1)</td>
<td>“we have several different methods looking at the same sorts of questions, each giving us an answer which taken in context, we can probably triangulate a lot of the findings and say either we are having a good impact or not. But, certainly of all the methods that we have, we think this is the one that’s most comprehensive, most robust” (LWC1)</td>
<td>“I did the initial contact and then later trained the staff members in inviting clients for the interview” (THR2)</td>
</tr>
<tr>
<td></td>
<td>“we need some evidence, but that evidence should be produced by some external party instead of Akhuwat itself”; “this is a key thing for us, that we should need some information which is prepared by reliable and professional people, so that the world could understand the importance and the existence of Akhuwat” (AIM1)</td>
<td>“the level of scrutiny of our investors is much less than perhaps an academic audience” (LWC1)</td>
<td>“And when I put the data into an Excel sheet” (AIM1)</td>
</tr>
<tr>
<td></td>
<td>“we could potentially be criticised for not being impartial in terms of the evaluation that we did. So what we decided to do was to look for academic collaboration” (LWC1)</td>
<td></td>
<td>“the microfinance institution was inputting the data so obviously a few kind of concerns around that, I guess in terms of the objectivity of entering the data themselves” (LWC2)</td>
</tr>
</tbody>
</table>

Source: Own construction
5.3 Data Analysis: a utilization-focused framework

In case study analysis, “much depends on a researcher’s own style of rigorous empirical thinking, along with the sufficient presentation of evidence and careful consideration of alternative interpretations”. Relying on pre-developed theories, working with the data from the ground up, developing a case(s) description and examining reasonable alternative explanations are general strategies for analysing the available evidence (Yin, 2018, p.165). In the context of the thesis, data analysis is based on use-focused evaluation theories, particularly the utilisation-focused evaluation (UFE) model described in chapter two.

5.3.1 The choice of evaluation model

There were two main factors influencing the selection of this evaluation model. The first refers to the conditions of my participation in the LWC assessment project and its context. Considering that the LWC team was actively involved (ultimately taking the decisions) in the design and implementation of the project, there were several important factors in the definition of the evaluation model that were out of my control and restrained the available options. The second factor, equally important in the decision, was my own practitioner background which led me to consider that a pragmatic approach would be better accepted and assimilated by the MFI managers.

In the appraisal of the different evaluation models presented in chapter two, the comparative analysis developed by Stufflebeam & Coryn (2014) was used as reference, with particular attention paid to the three best classified models – CIPP, UFE and constructivist model.

It is important to refer that this analysis was conducted by me in the context of the research developed in the PhD, independently of the LWC assessment project, and having in mind the analysis of the challenges associated with its implementation. In fact, there was no previous discussion with the evaluation stakeholders, and namely with the LWC team (as project leaders) about the theoretical model to be adopted in the project. There was an intentional decision from LWC to not restrict the design and implementation of the project to a particular theory, and to not reduce the data analysis to a comparison between targeted and real outcomes of the microcredit programmes evaluated.

These decisions do not imply that there was not an implicit programme theory underlying the selection of outcomes to be evaluated in the project. In the words of one of the LWC team members, “it has not been spelled out in the research, in our discussions, but I think intrinsically we have a theory in our own minds of what we expect our assistance is providing” (LWC1, personal communication, June 4, 2018). This approach to the project resulted from
the underlying objective of developing this “initial collaboration” as a first step in a learning process leading to improvements in the assessment and management of social performance in all the LWC partners, 14 at the time of the interview (LWC1, p.c., June 4, 2018).

In addition to this ‘open’ and broad perspective towards the outcomes to be evaluated, by the time I have joined the project (even if the evaluation was still in its early stages), there were already decisions taken regarding the design and methodologies to be employed in the project.

Considering the above, the constructivist model was discarded as main basis for the analysis in the thesis. A constructivist approach is associated with the application of qualitative methodologies and requires a previous agreement between evaluator and the project stakeholders that this is the adequate approach and that all partners will cooperate to accomplish it (Stufflebeam & Coryn, 2014, p. 200).

These conditions were hardly achieved within the context of the LWC project. This type of approach was not considered by LWC when designing the project and, as mentioned above, I have joined the impact team when some of the decisions have been already taken. It was possible to totally review the questionnaire and, thus, the outcomes to be evaluated, and to advise on the sampling methodology and the ethics standards and procedures associated with rigorous academic research (in accordance with the University of Portsmouth ethics policies). However, my position as a PhD candidate made it more difficult to lead the partners (which had very defined objectives linked to obtaining quantitative evidence on the impact of their programmes) to equate other approaches such as the constructivist. In these initial stages, my suggestions towards a mixed methods approach, complementing the quantitative data from the survey with the application of qualitative methodologies did not gather much support from the field partners.

My attention was, therefore, set on the two use-focused models, CIPP and UFE, with the main criteria for decision being related to the level of involvement of the programme stakeholders. It was clear from the beginning of the project that the involvement of all the programme stakeholders, particularly the programme beneficiaries or their representatives, would be difficult to attain, given the limited resources available for the project and the objectives and the approach privileged by the project leader and the field partners.

The microcredit clients of the two MFIs had an important role in the project through their participation in the household survey, but as in the case of the field staff of the MFIs, their participation was passive in the sense that they had no influence in the project decision-making process, and they were not involved in any way in the other stages of the process.
Equally, the focus of the LWC evaluation team was on the assessment of the social outcomes of the microcredit programmes, not aspiring at this stage to a comprehensive evaluation involving context, inputs and process as suggested by the CIPP model. This is not to say that these elements were totally neglected in the evaluation, but they were not the focus of the analysis, being included as part of the case studies description and used to contextualise the findings in terms of the product evaluation.

In this way, Patton’s UFE model and its approach based on the active participation of a selected group of stakeholders (primary intended users), whose interests and intended uses are privileged in the design and implementation of the project, fitted better the LWC impact project in the terms it was developed and this was the evaluation model I have chosen to analyse the implementation process associated with the project.

The selection of the UFE model implied a deeper review of the criticisms found in the literature regarding the model. One of the limitations associated with UFEs is put forward by Patton when identifying the practical implications of the personal factor. The possibility of staff turnover among the evaluation team is considered as sensitive since the model relies strongly on the relationship between the evaluator and the identified primary intended users (Patton, 2008, p.79; Stufflebeam & Coryn, 2014, p.415).

A more significant concern respects the selectivity in the participation of stakeholders by focusing on those with capacity to make and implement decisions and, thus, to use the evaluation. The interests of the intended beneficiaries of the programmes (e.g. microcredit clients), using Greene’s classification, may become secondary in the evaluation. The in-depth involvement of a few purposely selected stakeholders is supported by authors such as Alkin (2004) and Wholey (1987), but opposed by others, namely Fetterman (2003), House (2003) and Stufflebeam (2001) whose preference is for a more inclusive participation in the evaluation.

House (2003, p.54) defended the participation of all stakeholders, giving the involvement of some groups in drawing up the evaluation conclusions as an example of the potential bias in favour of the more powerful stakeholders. Stufflebeam (2001, p.57) also suggested the need to involve all stakeholders, introducing a panel of representatives to be consulted during the study. Both authors agree that minimising stakeholder bias is a responsibility of the evaluator. The possibility of co-option of the evaluator by the user group is a concern for many evaluation theorists who recognise that frequently evaluators, consciously or unconsciously, choose to highlight the outcomes and variables in which the programme is most successful (Stake, 1983a).
In addition to this vulnerability to bias and corruption by the user group, admitted by Patton as a potential “deadly sin of use” (Patton, 2008, p.90), the user group may be unable to guarantee the time, resources and safeguards required for an ethical and credible process of data collection, report and use of results. Under these circumstances, the independency, impartiality and objectivity of the evaluation may be questioned (Stufflebeam & Coryn, 2014).

Pawson and Tilley (1997, p. 13) adopted a different perspective, strongly criticizing Patton for reducing evaluations to a “collection of standard research tasks (...) which depend for their success on the application of sheer craft”. The authors considered that the UFE model, by focusing mainly on the evaluation “ends”, gives little clarification on how to implement an evaluation, especially when compared with the experimental model.

Pawson and Tilley (1997, p. 158-159) defended that proposing a combination of methods to be chosen from a list of possibilities as offered by Patton (2008, p.37), is not sufficient to claim a pluralistic approach in evaluation. Equally, the authors criticised the all-inclusive approaches from evaluators who call for the participation of all programme stakeholders in the evaluation, claiming this could lead to a potential overload of information. In this point, and despite different rationales, they came nearer to Patton’s position.

Pawson and Tilley (1997, pp. 13, 26) suggested that the obscurantism and the active involvement of practitioners associated with the UFE model, and the use branch in general, could bring about co-option by others, namely those paying for the research. In their realist approach, evaluations should be evaluator-led and explicitly theory-led (as opposed to the no theory approach suggested by some of the use branch proponents). Pawson (2014) emphasized the core role of theory in realist evaluations, which aim to seek and refine explanations for programme effectiveness.

Patton vividly replied to the criticisms, particularly condemning the tone employed by the authors of ‘Realistic Evaluation’ (Patton, 1999). In his review of the book, Patton considered Pawson and Tilley’s approach to be elitist in the way they see practitioners as mere sources of information for the evaluator/researcher and they decisively oppose stakeholders’ involvement in the evaluation process. This marks a sharp contrast between the advocates of realist evaluation and any of the use branch theorists.

In the comparative analysis of the two case studies, these criticisms will be revisited within the context of the LWC assessment project.
### 5.3.2 The UFE-based framework of analysis

The challenges and lessons learnt from the LWC assessment project are examined in the thesis with reference to two levels of analysis: the different stages of the project (design, implementation, interpretation and use of findings) and the main features of the UFE model (primary intended users, role of the evaluator and methods employed), as depicted in Table 5.2.

<table>
<thead>
<tr>
<th>Stakeholders / Primary Intended Users</th>
<th>Role of Evaluator</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identification of Primary Intended Users and their motivations Defining Levels of Participation</td>
<td>Role in Definition of Underlying Theory and Outcomes Evaluated Guarantee Engagement of Primary Users</td>
<td>Choice of Methods Preparation of Household Survey (Questionnaire Design, Sampling)</td>
</tr>
<tr>
<td>Building Primary Users Relationship Effective Levels of Participation Lessons Learnt from Process</td>
<td>Guarantee an Ethical, High-Quality Research Guarantee Engagement of Primary Users</td>
<td>Recruitment and Training of Interviewers Quality Control of Survey Computerisation of Data Data Analysis</td>
</tr>
<tr>
<td>Interpretation of Results Perceptions and Use of Results</td>
<td>Guarantee Engagement of Primary Users Report Results Encouragement of the Use of Findings</td>
<td>Report and Dissemination of Findings</td>
</tr>
</tbody>
</table>

Source: Own construction

In the design stage of the impact evaluation project, my main focus was on the identification of the intended primary users of the evaluation, their motivations and expected level of participation in the project. Considering that the “personal factor” is the core element of the UFE model, I have given special attention to the evaluation team members (their personal characteristics, motivations and behaviours), as well as the relationships built during all stages of the project. These factors are critical to the interpretation of the evaluation findings and the use of the project results and lessons learnt (Patton, 2008, p.66).

Equally critical in the UFE model is the role of the evaluator and his/her capacity to adopt a “active-reactive-interactive-adaptive” approach to the evaluation, ensuring that the overall objectives of the primary intended users remain a priority throughout the different stages of the project (Patton, 2008, p.213). Given the partially goal-free approach and the defined evaluation tools, I had a more active role in the stages of preparation of the two waves of the household survey and in the analysis of the results obtained.
From a UFE model perspective, however, the most important features of the role of the evaluator are associated with the pursuit of a continuous engagement with the primary intended users during the project while also guaranteeing ethical and high-quality standards in the evaluation process. In this mission I was conditioned by my consultancy/advisory role and the conditions of the project, including the available resources and the geographical distance between the partners.

The third column in the table refers to the methods employed in the three phases of the project. Although Patton is not prescriptive regarding specific methods in the UFE model, the selection of evaluation tools is essential to the results and their use by the primary intended users (Patton, 2008, p.199).

The LWC team made the initial choice of data collection methods. The decisions regarding their implementation were ultimately taken by LWC, based on my proposals as academic consultant, and following discussion with the partner MFIs. Data analysis was the only stage in which I have led the process, notwithstanding the consultation with other evaluation team members to ensure the reports produced included also the issues of interest to the primary intended users.

In the thesis, I have given particular attention to the field implementation of the chosen methods, with an extensive description of the application of the selected evaluation instruments in the two settings. The identification of the challenges encountered in Pakistan and Zimbabwe, and the description of the distinct responses to these challenges adopted by the partners, are one of the main contributions of the research project. These are essential elements to improve the rigour and credibility of evaluations, which are often neglected in the microfinance literature.

In this UFE-based framework of analysis is important to remind that the last stage of the project – the use of findings – is essential. As described in section 2.3.2, evaluations are conducted to “inform decisions, clarify options and identify potential improvements”, representing in this way a basis for action (Patton, 2008, p.37). It should also be kept in mind that these actions can occur at different levels, including the organisation’s processes (e.g., social performance assessment), the programme’s conditions and mechanisms of delivery and the different stakeholders’ perceptions and attitudes, and that they can be implemented in different moments in time. Therefore, special attention was given both to the findings of the evaluation, which refer to the observed changes in the selected outcomes and potential contribution/impact of the microcredit programme, and to the lessons learnt from the evaluation process.
The matrix presented in Table 5.1 is used as reference in the chapters six and seven when describing the evaluations implemented in Pakistan and Zimbabwe. In this way, for the different stages of the evaluation project, the central issues in a UFE are discussed and form the basis for start answering the PhD first research question.

5.4 The LWC assessment project: an overview

5.4.1 Lendwithcare: the crowdfunding platform

Lendwithcare is a crowdfunding platform created in 2010, by initiative of CARE International UK, a British development NGO.

The model followed by LWC is, in general, similar to the one developed by KIVA, the pioneer and largest microfinance crowdfunding platform created in 2005 in San Francisco (USA). In 13 years, the American platform has reached 3 million borrowers in 81 countries and involved 1.8 million lenders (KIVA, n.d.). A similar model was also replicated by Babyloan, a French platform initiated in 2008 (Babyloan, n.d.).

The three platforms have many aspects in common but also significant differences. One of the main differences refers to their origins. While KIVA and Babyloan were created by a group of individual social entrepreneurs in the USA and France, respectively, LWC is an initiative of a UK-based development NGO, which is part of a large international development network (CARE). CARE has a long history in the provision of microfinance and the development of financial inclusion in developing countries, particularly through their village and savings loans programme (VSLA), which started in 1991. By 2016, the programme has supported the creation of 200,000 groups with approximately 5 million members in 35 countries (CARE International UK, 2018).

LWC started as an alternative and complementary way to work financial inclusion through new technologies, particularly the novelty crowdfunding techniques. The independent department, within CARE International UK, started with a small team of three people, including a microfinance expert with more than 20 years of experience in the sector, playing roles from consultancy to management of microfinance institutions in different countries. His expertise was fundamental in the initial stages of the crowdfunding platform when defining the selection criteria for the local partners and the target countries.

In the initial years of the project, the number of lenders attracted, entrepreneurs supported, and local partners has grown gradually. By June 2018, the platform worked in 11 countries (Table 5.3). It had supported 89,845 low-income entrepreneurs and lent $19.3 million
through the provision of 615,005 microcredit loans. It had 49,267 lenders with active profiles in the platform, mostly based in the UK.20

Table 5.3 – Lendwithcare Portfolio

<table>
<thead>
<tr>
<th>Country</th>
<th>Partner MFI</th>
<th>Established</th>
<th>Legal Status</th>
<th>No. Active Borrowers</th>
<th>Average loan size</th>
<th>PAR 30 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cambodia</td>
<td>BORVOR</td>
<td>2003 (as NGO)</td>
<td>NBFI (since 2017)</td>
<td>3,433</td>
<td>$2,828</td>
<td>2.40%</td>
</tr>
<tr>
<td>Ecuador</td>
<td>FACES</td>
<td>1991</td>
<td>NGO</td>
<td>17,340</td>
<td>$2,564</td>
<td>2.20%</td>
</tr>
<tr>
<td>Ecuador</td>
<td>Cooperativa Santa Anita</td>
<td>2001</td>
<td>Cooperative</td>
<td>3,596</td>
<td>$3,612</td>
<td>1.10%</td>
</tr>
<tr>
<td>Malawi</td>
<td>Microloan Foundation</td>
<td>2002</td>
<td>NGO</td>
<td>26,098</td>
<td>$81</td>
<td>4.60%</td>
</tr>
<tr>
<td>Pakistan</td>
<td>Akhuwat</td>
<td>2001 (as NGO)</td>
<td>NBFI (2017)</td>
<td>820,071</td>
<td>$144</td>
<td>0.26%</td>
</tr>
<tr>
<td>Peru</td>
<td>IDERCV</td>
<td>1990</td>
<td>NGO</td>
<td>2,233</td>
<td>$860</td>
<td>3.92%</td>
</tr>
<tr>
<td>Philippines</td>
<td>LAMAC</td>
<td>1992</td>
<td>Cooperative</td>
<td>42,015</td>
<td>$400</td>
<td>9.20%</td>
</tr>
<tr>
<td>Rwanda</td>
<td>Umutanguha</td>
<td>2003</td>
<td>Cooperative</td>
<td>9,519</td>
<td>$520</td>
<td>5.60%</td>
</tr>
<tr>
<td>West Bank and Gaza</td>
<td>Reef Microfinance</td>
<td>2007</td>
<td>NGO</td>
<td>3,614</td>
<td>$3,512</td>
<td>6.60%</td>
</tr>
<tr>
<td>Vietnam</td>
<td>MACDI</td>
<td>2007</td>
<td>NGO</td>
<td>4,608</td>
<td>$221</td>
<td>1.7%</td>
</tr>
<tr>
<td>Zambia</td>
<td>Microloan Foundation</td>
<td>2008</td>
<td>NGO</td>
<td>8,409</td>
<td>$121</td>
<td>6.7%</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>Thrive Microfinance</td>
<td>2012</td>
<td>NBFI</td>
<td>2,199</td>
<td>$655</td>
<td>9%</td>
</tr>
<tr>
<td>Zimbabwe</td>
<td>Microloan Foundation</td>
<td>2016</td>
<td>NBFI</td>
<td>185</td>
<td>$124</td>
<td>0%</td>
</tr>
</tbody>
</table>

Source: Own construction from Lendwithcare (2018).
Notes: Information updated in June 2018 for all partners, except IDERCV (August 2017) and Akhuwat (case study data, see Table 6.3). Data on Thrive Microfinance differs from data used in the case study which refers to December 2017 (see p. 156).

The LWC model is illustrated below in Figure 5.1. An important feature of the model is the no interest associated with the loans provided by the crowd (lenders) and, likewise, with the loans provided by LWC to the local microfinance institutions. The operational costs of the platform as well as the technical assistance provided to the local partners is covered mainly by donations.

The conditions of the business loans obtained by the borrowers at local level are dependent on the microcredit programmes implemented by each of the partner MFIs. Further information on the programmes and the local institutions is available in the platform website.

20 Information provided by LWC upon request of the researcher.
The partners have in common social objectives, but the programmes are diverse and adapted to the local contexts. Apart AIM in Pakistan, which provides *qard-hasan* loans, all other institutions run interest-based programmes. The differences between the MFIs are not limited to the cost of the loans, with screening criteria, loan sizes and complementary financial and non-financial services varying greatly across the partners (Lendwithcare, 2018).

**Figure 5.1 – Lendwithcare Crowdfunding Model**

Source: Own construction

LWC has remained focused on funding business loans and supporting low-income entrepreneurs in the developing countries. This strategy diverges from the ones defined by Babyloan, which has enlarged its area of action to include developed countries in Europe (Belgium and France), and KIVA which supports loans for all types of uses (including housing, education, consumption, emergencies, etc.) and has been working with local partners in the USA.

One final note to highlight the role of money fungibility in the model. Although the lenders select specific borrowers in the LWC platform, who are monitored in terms of repayment (as the credit risk is supported by the crowd lenders), the funding transfers between LWC and each partner MFI are based on the net balance between funds to receive and to repay regarding the borrowers supported during the period.

### 5.4.2 Lendwithcare Impact Assessment Project

“LWC main purpose is to improve the lives of poor people” (LWC1, p.c., June 4, 2018). After five years facilitating loans to poor entrepreneurs through the crowdfunding platform, the LWC team wanted to understand whether their work was having a positive impact (or not) on the lives of the clients and their households: “as practitioners, we often are more involved in the actual implementation, but we don’t step back enough and see whether the work that we are doing is having positive impact or not” (LWC1, p.c., June 4, 2018).

Initially, the LWC team thought of implementing the evaluation themselves, but feeling that this could generate criticism for impartiality in the process, “what we decided to do was to
look for academic collaboration”. The academic partner could bring expertise, and above all, legitimacy to the evaluation (LWC1, p.c., June 4, 2018).

The evaluation of the outcomes of the microcredit programmes at the client level, although being the more immediate objective for LWC, was not the sole motive for the project. One of the goals of LWC is to help its partner MFIs to “increase the scale and the outreach of the programmes”, by providing not only funding through the crowdfunding platform, but also technical assistance in diverse areas according to the needs of the local institutions: “one of the things that we were always concerned about is what sort of technical expertise can we provide”. The evaluation of the social performance of the MFIs was signalled by the LWC team and by the MFIs themselves as an evident need for the institutions, which in some cases have “really done anything” in this area (LWC1, p.c., June 4, 2018).

LWC1 mentioned in the final interview an additional motivation for the project. The information collected as well as the results of the evaluations implemented are considered as a part of a wider effort in terms of measuring the social performance of their partners and their own - “we have several different methods looking at the same sorts of questions, each giving us an answer which taken in context, we can probably triangulate a lot of the findings and say either we are having a good impact or not.” (LWC1, p.c., June 4, 2018).

In this context, the LWC team has been employing other data collection methods, namely the request of quarterly data on 18 financial and social indicators from each partner, regular evaluation field visits, including interviews with clients and staff of the institutions, and more recently, a pilot experience implementing social ratings in two MFIs through a partnership with a qualified rating agency (LWC1, p.c., June 4, 2018).

The impact assessment project has, thus, been considered by the LWC team as complementary to these different sources, with the triangulation of data allowing for a more complete image of the performance of their partners, which has been suggested by Hulme (2000b) to be an implicit objective for many sponsors of impact assessments in microfinance.

The crowdfunding programme had during the impact project a small team of four people, two of them were directly involved in the implementation of the impact assessment, with the occasional support of the other members of the team. The first contacts with potential partners occurred in late 2014. These contacts included the field partners in different countries and an academic institution in the UK to act as consultant/advisor in the project.

The first MFI selected to participate in the project was Akhuwat Islamic Microfinance (Pakistan), which was joined by THRIVE Microfinance (Zimbabwe) later in 2015. The selection
criteria of the two MFIs, as described in the Introduction (see p. 13), included factors related to the countries (English speaking countries and culturally diverse) and factors on the organisations’ characteristics (interest and motivations to participate, capable local evaluation teams, non-existent or incipient social performance systems implemented, diversity in terms of maturity, size, organisational structure, missions and target populations).

The University of Portsmouth became a formal partner in February 2015. At this stage, some decisions regarding the methods (household survey) and timings of the evaluation had already been made by the LWC team, but the evaluation design was not totally established and issues such as the requirement of a control group for comparison had not been properly considered. The relevance of a strong control group, in the context of a quantitative approach (Barnes & Sebstad, 2000, p.16) was not completely obvious for all the partners involved in the project, and there were (justified) concerns regarding the challenges associated with its field implementation, which are discussed in the following chapters.

In both research settings, business loans delivered through group lending methodologies are the core product of the MFIs, and the focus of the evaluations. The main motivation for the institutions to participate in the project was accountability towards external parties. This was particularly evident in the case of AIM, while for THRIVE, the accountability purposes were complemented with the aspiration to develop the expertise of the institution in the evaluation of their social performance.

In the presentation of the use branch theorists in chapter two, one of the highlighted elements of the models was the focus not only on the outcomes of the programmes evaluated, but also on the lessons learnt through the implementation of the evaluation. This rationale was not foreseen as a priority by all the field partners at the initial stage of the evaluation since their involvement in the project was mainly motivated by accountability-driven purposes.

This initial oversight of the implementation phase is illustrated by examples of indifference demonstrated by the partners towards questions such as the effective duration of the survey interviews, or the reasons for non-participation by the clients and non-clients (when they were invited to participate but were not available for the interviews). This initial behaviour can be explained by the focus on the results of the evaluation, as well as the non-familiarisation with the evaluation process. As the project evolved, so did the attitude of the partners in regard to the use of the evaluation process, which is reflected in LWC1 statement:
“I think our partners are now beginning to take the issue of evaluation research much more seriously as well” (LWC1, p.c., June 4, 2018).

The initial focus on the results of the programmes implied being able to identify the changes in the lives of the clients after successfully applying for and receiving the microcredit loan. Ravallion (2008) observes that a frequently used evaluation tool to assess programme outcomes is a sample survey of the participants. In the context of the LWC assessment project, it was decided to implement a longitudinal survey applied to a sample of clients and non-clients in order to construct a dataset that could be analysed using panel data techniques. This choice of method is similar to the one adopted in previous impact studies in microfinance, including Barnes (2001), Chen and Snodgrass (2001) and Dunn and Arbuckle Jr (2001), which describe evaluations conducted on microfinance programmes supported by USAID.

It should be noted that implementing an RCT to evaluate the microcredit programmes was not considered by LWC at this stage, most likely due to the high cost and complexity associated with this type of evaluation. Furthermore, the nature of the microcredit programmes evaluated, namely their on-going status and established lending methodologies, made it less adequate to develop an experiment (Odell, 2015; Stufflebeam, 2001).

The longitudinal aspect of the survey involved interviewing the same sample of clients and non-clients in successive waves of interviews. Using panel data presents several advantages when compared with cross sectional research, being its application in microfinance evaluations advocated by several authors, including Khandker (1998, p.18) and Garikipati (2017, p.200). Among the strengths of panel data identified by Glewwe and Jacoby (2000, p.276) are an increase in the precision of the estimates obtained relating to changes in the mean values of the variables (smaller standard errors), and the possibility of tracking specific individuals or households. In addition, the repetition of the interviews at different moments in time, increases the accuracy regarding past events when compared to recall data obtained by asking retrospective questions in a one-time survey.

In the implementation of the LWC evaluation project, there was a concern to ensure that the clients who have exited the programme during the period of the project were included in the final dataset. This is important in longitudinal studies since, otherwise, only successful clients will be compared with non-clients, introducing bias in the analysis (Goldberg, 2005, p.14; Sebstad & Chen, 1996).
This was not an easy task as recognised by the evaluation teams of both MFIs, with the interviewers in the two countries facing more difficulties to locate and interview clients who have faced problems repaying their first loans or had their second loan application rejected by the institutions - “issue was only with those clients we have not given a second loan because they had not repaid their first loan quite well” (AIM2, p.c., March 22, 2018); “most of the clients who were in default would not avail themselves for the interviews because some thought they were being called pertaining their non-performing loan” (THR2, p.c., April 23, 2018).

Despite these difficulties, the number of drop-out clients interviewed in both countries was considered satisfactory given the research objectives. In the case of Pakistan, 24% of the clients re-interviewed (106 clients) had finished repaying their first loan and had no intention of applying for a new loan (Afonso, 2018b). In Zimbabwe, 36% of the clients re-interviewed (89 clients) had not applied for a second loan, 35% were finishing or had finished a second loan and 29% had continuously renewed their loans and were repaying a third loan by the time of the interview (Afonso, 2018c).

This PhD thesis focuses on the first two waves of the household survey, although the partners involved have affirmed the intention to continue the project and carry out a third wave of interviews in 2019. The intended continuity of the evaluation is one of the strengths of the project as it will allow to collect and analyse data for an extensive period on the two microcredit programmes, minimising the limitations associated with using short-term data to assess impact on the welfare of the clients, which has a long-term nature (Ravallion & Chen, 2005).

The evaluation tools applied in the LWC impact project were a purposely built questionnaire and a poverty assessment tool administered through a household survey. In both case studies, the first wave of interviews established the study baseline, with the second wave being implemented after the clients have completed at least one loan cycle. In the case of AIM, the borrowers were selected from the four branches which provide the borrowers profiles for the crowdfunding platform. This meant working with four of the around 300 existing branches at that stage. This geographical selection was not relevant at THRIVE, where the two existing branches in 2015 were included in the study.

The first client screening criterion was, thus, related to the geographical location of the clients and derived from the branch selection. The second criterion was being a new client, meaning that this was the first time successfully applying for a loan at the MFI. These new clients had been informed that their loan request was approved and they were about to
receive the money. The third criterion was of non-exclusion, i.e., all new clients in the branches selected were invited to participate in the survey until the goal established in terms of sample size was achieved. As described in the respective chapters (see sections 6.3.2.3 and 7.3.2.3), the recruitment process of the clients was straightforward in its application in both countries.

Not as simple was the recruitment of the non-clients. The objective was to find entrepreneurs with the same characteristics of the clients (type of activities, income and poverty levels, etc.), but who had not applied for a loan at the MFIs. In order to do so, different strategies were adopted in the two countries following the advice of the local evaluation team members: in Pakistan, information collected by the loan officers and clients of the MFI was used to identify potential interviewees, while in Zimbabwe the non-clients were recruited directly at different market places. As in the case of the clients, the process is further detailed in chapters six and seven.

The anticipated difficulties in reaching non-clients and the initial focus of the field partners on client outcomes led to the definition of less ambitious objectives in terms of the number of interviews for the non-clients, and to the inclusion of entrepreneurs located only in the same neighbourhoods of the clients. In this way, Gaile and Foster’s (1996) recommendation to include in the control group eligible entrepreneurs from areas outside the selected locations was not followed, given the additional effort and resources that would have been required to identify, contact and interview these non-clients.

The questionnaires (clients and non-clients) and the poverty assessment forms used in the two waves of the survey can be found in Annexes II and III. The questionnaires were planned to take between 20 and 25 minutes, eventually slightly more in the second survey considering the open questions introduced. The questionnaires were initially prepared for the Pakistan MFI and then adapted to the Zimbabwean institution, taking into consideration the local context and the particular characteristics of the THRIVE programme.

The selection of the outcomes and respective indicators was decided in the first instance by the researcher who, based on her knowledge of the existing theoretical and empirical research on the impact of microcredit programmes (presented in chapter three) and the initial information gathered about LWC and the MFIs, proposed the questions to be included in the client and non-client questionnaires. In both countries, these were discussed with the evaluation teams of LWC and the respective partner MFI. Each partner made comments and suggestions to better relate the questions to the local contexts and their own perceptions of outcomes of interest for each institution.
The process was intuitive, based on the experience of the evaluation team members and the researcher in the sector, and can be considered as a partially goal-free evaluation. Goal-free evaluations are a concept put forward by Scriven in 1972 (Scriven, 1991), which involves gathering data on a wide range of actual effects of the programme and assessing the relevance of these effects to meet the needs of the beneficiaries. In the thesis context, it seems more adequate to consider Marvin Alkin’s interpretation of this strategy as being only partially goal-free, i.e. it is goal-free for the specific targets defined by the managers for the programme outcomes (Stufflebeam & Coryn, 2014).

With regard to the evaluation methods, there was a significant difference between the two countries related to the poverty assessment tool applied. There was an initial decision by the LWC evaluation team to use the Poverty Probability Index (PPI) in the impact assessment project. This poverty scorecard, described in detail in chapter four, was considered to be the “industry’s standard” (LWC1, p.c., June 4, 2018), being recommended by the main international actors in the microfinance sector, including the Consultative Group to Assist the Poor (CGAP) and SPTF. Moreover, the tool was already being implemented by some of the LWC partners, including FACES in Ecuador and the Microloan Foundation in Malawi. The PPI for Pakistan was, thus, integrated in the AIM household survey.

However, the same approach was not an option in the case of the THRIVE survey, since Zimbabwe is one of the countries for which the instrument has not yet been developed (IPA, 2018). Despite no PPI or other externally validated scorecard being available, THRIVE had developed their own poverty assessment tool (THRIVE-PAT), within its broader strategy of assessing social performance. Based on the methodology used by the developers of PPI (and publicly available through the materials published online), the MFI identified a list of eight questions and the respective scores. The data are collected by the loan officers during the loan appraisal process, analysed by the social performance officer and used by managers to monitor outreach targets.

In the survey, THRIVE-PAT was used alongside the purposely built client questionnaire. Additionally, the eight questions were integrated into the non-client questionnaire in order to compute the poverty scores for the control group. In the description of the two evaluations conducted, attention is given to the advantages and challenges associated with the use of PPI or its Zimbabwean alternative.

The questionnaires were administered by independent interviewers recruited from local Universities in Pakistan and Zimbabwe. The teams were different in each wave of the interview, with all the interviewers receiving training from the evaluator/researcher and the
LWC team prior to their field missions. The first week of implementation of the surveys was monitored by the UoP/LWC team, as part of the interviewers’ training and the initial validation of the data collection process. After all interviews were concluded, the data collected was introduced in an excel-based database and made available to the evaluator for analysis.

The baseline survey allowed for the characterisation of the clients, the comparison between clients and non-clients and the identification of differences between segments of clients. The findings were reported to each institution at this stage. After the conclusion of the second wave of the surveys, the data collected in the two surveys were analysed using a double-difference method. Double-difference methods assume that unobserved differences between cases are invariant with time and compare the ‘treated’ and the control groups before and after the programme intervention (Khandker et al., 2010). In addition to the comparison between clients and non-clients, quantile regression was applied to explore heterogeneity across the client sample, which had been identified in the baseline surveys in both countries.

The results of the evaluation were reported through a comprehensive document detailing the analysis and main findings for each microcredit programme, prepared independently by the evaluator. These ‘working’ reports were sent to the partners for comments in order to produce final versions which could be used by LWC and the MFIs in their relationship with external parties, including existing and potential funders. At this stage, two summary reports were also produced, for which the content was adapted to fit the specific needs of the partners.

In the chapters six and seven, a comprehensive description of the different phases of the evaluation process is provided, from the design of the questionnaires to the interpretation, report and use of results. As can be verified in these chapters, the evaluation approach adopted was similar in both settings, although the challenges encountered, and the responses negotiated with the two MFIs were diverse in many aspects. Chapter eight discusses the common project approach to evaluation (second research question), identifying the advantages and limitations associated with the process.

6 Case Study: Akhuwat Islamic Microfinance, Pakistan

The first case study aims to describe and discuss the evaluation of the Akhuwat Islamic Microfinance (AIM) microcredit programme conducted within the LWC assessment project. The project started during the last trimester of 2014 and was implemented by a team including members from LWC, AIM and the University of Portsmouth.
The case study starts setting up the context in which the microcredit programme is developed (section 6.1), followed by a brief presentation of the main features of the institution and the programme being evaluated (6.2). Section 6.3 describes the evaluation implemented using the analysis framework developed in chapter five. For each of the three stages of the project, the participation of the primary intended users and the evaluator, as well as the application of the selected evaluation tools, are presented and discussed. Finally, section 6.4 summarises the main challenges and lessons learnt from the implementation of the project in Pakistan, emphasizing the findings from the case study which answer the first research question.

### 6.1 Macro-Context

Pakistan is a “hard” and misunderstood country. It is “economically backward, corrupt, violent, unjust, often savagely oppressive towards the poor and women (...) ‘yet it moves’ and is in many ways surprisingly tough and resilient as a state and a society.” (Lieven, 2011, p.4). The country is an Islamic Republic created in 1947 to accommodate the Muslim population after the independence of India from the British. Home to more than 209 million people, it is composed by four provinces (Balochistan, Khyber Pakhtunkhwa, Punjab, Sindh) and four federal territories (Federally Administered Tribal Areas, Islamabad Capital Region, Gilgit-Baltistan, Azad Kashmir), which are geographically, economically and socially very diverse.

The country is bordered by India to the east, Afghanistan to the west, Iran to the southwest and China to the far northeast. The troubled relation with its neighbours, particularly India, as well as the Independence War of East Pakistan (now Bangladesh) in 1971 has led to the development of a strong military force (with nuclear power), which balances an unstable political system, based on “patronage and kinship”, and the co-existence of three judicial systems: the law of the state, the law of religion (Shari’ah) and the local community law. This context results in a weak state, restraining the overall economic and social development of the country. (Lieven, 2011, pp.204, 212).

The economy of the country, measured by its Gross Domestic Product (GDP) growth rate, has grown consistently throughout the 2010s, but at a slower pace when compared with the average of the lower middle-income countries (LMIC) and the South Asia countries.\(^{21}\) The evolution seems to be more encouraging regarding the poverty indicators. The most recent available data for the country refers to 2013, when the poverty headcount ratio at the

---

\(^{21}\) GDP growth rate in 2015 of 4.7% for Pakistan, 5.4% for LMIC and 7.5% for South Asia. In 2016, the situation seemed to have slightly improved compared with the LMIC, 5.5% and 5.1% respectively. Retrieved from The World Bank website: [http://databank.worldbank.org/data/home.aspx](http://databank.worldbank.org/data/home.aspx)
national poverty line was 29.5%, which represented a significant improvement from 36.3% in 2011. This positive evolution is also observable comparing the international poverty indicators of Pakistan with those of the LMIC and the South Asia countries. The poverty headcount ratio at the $1.90/day (2011 PPP) poverty line in 2013 was 6.1% for Pakistan, which compared to 15.9% for the LMIC and 15.1% for South Asia.22

In order to understand the context of implementation of the AIM microcredit programme, it is important to consider the development of the financial sector in the country, particularly the figures for financial inclusion. Contrary to the established argument that financial and economic development are positively associated (Beck et al., 2007a), Pakistan has been consistently performing poorly in regard to financial inclusion when compared with countries with similar level of economic development, as illustrated in table 6.1. The table displays just a few of the indicators included in the Global FINDEX survey.

Table 6.1 – Pakistan Financial Inclusion Data

<table>
<thead>
<tr>
<th></th>
<th>2017</th>
<th>2014</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults with account, including mobile accounts, total</td>
<td>21.3%</td>
<td>57.8%</td>
<td>13.0%</td>
</tr>
<tr>
<td>Adults with account, including mobile accounts, female</td>
<td>7.0%</td>
<td>53.0%</td>
<td>4.8%</td>
</tr>
<tr>
<td>Adults with account, including mobile accounts, poorest 40%</td>
<td>14.2%</td>
<td>50.7%</td>
<td>11.2%</td>
</tr>
<tr>
<td>Adults with account at financial institution, total</td>
<td>18.0%</td>
<td>56.1%</td>
<td>8.7%</td>
</tr>
<tr>
<td>Adults with account at financial institution, female</td>
<td>6.3%</td>
<td>51.7%</td>
<td>3.0%</td>
</tr>
<tr>
<td>Adults with account at financial institution, poorest 40%</td>
<td>12.8%</td>
<td>49.3%</td>
<td>6.3%</td>
</tr>
<tr>
<td>Adults borrowing from a financial institution, total</td>
<td>2.3%</td>
<td>8.4%</td>
<td>1.5%</td>
</tr>
<tr>
<td>Adults borrowing from a financial institution, female</td>
<td>1.5%</td>
<td>7.3%</td>
<td>0.3%</td>
</tr>
<tr>
<td>Adults borrowing from a financial institution, poorest 40%</td>
<td>2.6%</td>
<td>7.1%</td>
<td>2.0%</td>
</tr>
</tbody>
</table>


These numbers help explaining the large potential market for microfinance institutions estimated by the Pakistan Microfinance Network (PMN) in 20.5 million clients in September 2017 (Haider, 2017), as well as the classification of the country as underserved market in the 2014 MIMOSA report on market saturation (Javoisy & Rozas, 2015).

Table 6.1 shows that, despite a very positive evolution during the period between the two last surveys, Pakistan still lags far behind the average of the LMIC in terms of access to bank

22 Source: The World Bank
accounts and formal credit. The exclusion from access to formal financial services is especially strong for women and for the poorest segments of the population. It can also be noted in the two right columns of the table that the indicators referring to services provided by formal financial institutions deteriorated between 2011 and 2014, with improvements in access to accounts in the period being mostly driven by the growth of mobile accounts.

The scenario seems to be reverting recently, with significant improvements in all indicators in the third wave of the survey. Interestingly, it can be read from the table that for women and the poor this improvement has been based on financial services provided by formal financial institutions. These are the clients of the microfinance sector, which has been gaining ‘momentum’ since 2015, with a significant growth in breadth of outreach as depicted in table 6.2.

<table>
<thead>
<tr>
<th>Table 6.2 – Pakistan Microfinance Sector 2015 - 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mar. 2015</strong></td>
</tr>
<tr>
<td><strong>No. MFI Branches</strong></td>
</tr>
<tr>
<td>2,587</td>
</tr>
<tr>
<td><strong>Districts Covered</strong></td>
</tr>
<tr>
<td>94</td>
</tr>
<tr>
<td><strong>No. Active Savers</strong></td>
</tr>
<tr>
<td>10,230,527</td>
</tr>
</tbody>
</table>


This growth has been mainly based on the expansion and strengthening of existing institutions, both microfinance banks and microfinance institutions (including AIM), more than on the entry of new competitors in the market. The period has been characterised by more active roles of the Pakistan Microfinance Network and the Government in the promotion of financial inclusion. Ahmed (2016) suggested that the changes introduced in the regulation of the sector and the launch in May 2015 of the National Financial Inclusion Strategy by the State Bank of Pakistan have provided a more suitable environment and incentives for the development of the microfinance sector.

In an effort to professionalise the sector and avoid delinquency crisis as the one occurred in 2009, a sector credit bureau was implemented, and new regulation introduced for the microfinance institutions. In 2017, the financial operations of these MFIs, frequently created

---

23 Unlike the delinquency crises occurred in other microfinance markets (e.g. Andra Pradesh, Morocco and Nicaragua), the situation in Pakistan was constrained geographically to Punjab and affected particularly one institution. It worked, however, as a strong alert to the dangers associated with uncontrolled growth and client over-indebtedness, and motivated institutions, the national network and the government to introduce changes in the legislation and self-regulation mechanisms to prevent future crises (Burki, 2009; Chen et al., 2010)
as NGOs, were required to transform into non-financial banking institutions under the supervision of the Securities and Exchange Commission of Pakistan.

6.2 Micro-context

6.2.1 AIM and Islamic Microfinance

AIM is an Islamic MFI based in Lahore, capital of Punjab. It started operations in 2001 with the mission to “alleviate poverty by empowering socially and economically marginalised segments of society through interest-free microfinance and in the process harnessing their entrepreneurial potential and enhancing their capacity through economic and social guidance”. The institution does not have political or religious ties to other organisations, but it is a faith-based institution where compliance with Islamic principles guides the strategy and procedures of the MFI (Khan et al., 2017, p.18).

As in many other Islamic MFIs, the institution has slowly grown during the first decade, gradually expanding its services to other provinces and diversifying its product mix. Other credit products (housing, education, health, marriage, emergencies and debt repayment) were introduced, as well as non-financial services, including education, health and a clothes recycling programme.24

The institution started working as an NGO, funded mostly by donations, but in July 2017, as a consequence of the above referred changes in the microfinance sector regulation, its financial and non-financial operations were formally separated. Microcredit became the core activity of AIM Islamic Microfinance, a non-bank financial institution, while the non-financial activities remained in the sphere of Akhuwat NGO.

Table 6.3 presents the main activity indicators for the financial institution. In the second semester of 2017, AIM became the largest MFI in Pakistan in terms of number of active borrowers (market share of 15.8%), and geographical coverage. Its market share was, interestingly, lower in terms of gross loan portfolio (7.7%), only fifth in the national ranking (Haider, 2017). This derived from a smaller average loan size ($144), a strong sign that the institution works with vulnerable segments of the population (Bos & Millone, 2015; Morduch, 1999).

24 Marriage loans aim to help widows supporting the costs associated with their daughters’ wedding and dowry. Liberation loans are destined to repay previous debts with private moneylenders (at higher interest rates). The clothes programme employs exclusively members from the transgender community, one of the most marginalised groups in Pakistan (Khan et al., 2017).
Table 6.3 – AIM Main Indicators

<table>
<thead>
<tr>
<th></th>
<th>June 2017</th>
<th>June 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. Branches</td>
<td>686</td>
<td>820,071</td>
</tr>
<tr>
<td>No. Districts</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Gross Loan Portfolio</td>
<td></td>
<td>$118,534,055</td>
</tr>
<tr>
<td>PAR30days</td>
<td>0.26%</td>
<td>42% [1]</td>
</tr>
</tbody>
</table>


By comparison, in March 2015, AIM was the third largest MFI in client outreach, with 355,880 active borrowers served by 300 branches across 59 districts (PMN, 2015). In a period of little over two years, the number of active borrowers and branches increased approximately 130%.

The extraordinary growth of the MFI stems, partially, from increasing partnerships with regional governments, particularly the Punjab regional government, which have been developing their own interest-free microcredit programmes. These are designed to promote local entrepreneurship among the poor segments of the population, traditionally with lower levels of literacy and limited access to permanent paid jobs. The programmes provide funds for the loans, and an additional amount for AIM to cover its operational costs.

Other fund sources of the institution include the application fees charged to their borrowers, individual and collective donors (including their own borrowers) and, marginally, LWC which funded less than 2% of the gross loan portfolio in 2016. The distribution of AIM’s income and expenses from June 2002 to June 2016 is presented by Khan et al. (2017, pp. 25-27), who refer to a cost coverage ratio of 131% by June 2016. It is important to consider, however, that taking into account the particularities associated with the qard-hasan microcredit model, this ratio includes all donations applied into the MFI’s operations, and, in this sense, it differs from the conventional OSS (operational self-sustainability) ratio.

6.2.2 Microcredit programme: Family Enterprise Loan

Despite the wider offer of credit products, the portfolio of the institution is dominated by qard-hasan loans to low income entrepreneurs through a product called ‘Family Enterprise Loan’. The offer of interest-free benevolent loans is one of the distinctive features of AIM in the context of the Pakistani microfinance market where, surprisingly, Islamic microfinance has not made much impact so far.25 It is also a unique case among Islamic Microfinance, in

---

25 Pakistan Microfinance Network classified only 16% of the Active Borrowers as Islamic microfinance clients in September 2017, being the majority AIM clients (Haider, 2017).
general, as most institutions choose to offer murabaha (trade loans) and, in a lower number of cases, musharaka and mudaraba (profit and loss-sharing instruments) (Harper & Khan, 2017).

The AIM borrowers pay no interest or penalties in case of default, with the only monetary cost associated with the loan being a fixed application fee of PRK 200, paid up-front. The fee amount is independent of the loan amount requested. The MFI follows a group lending methodology in several features similar to many conventional MFIs. The candidates are requested to form a group of three to six members who will be jointly liable for the loans of each member of the group. The loan evaluation process follows many of the practices advocated by the main international actors such as the CGAP. Among them are the analysis of income and expenditure both at the business and household levels, and the consultation of the sector credit bureau to verify the indebtedness level of the loan applicant. Figure 6.1 summarises the lending process.

Figure 6.1 – AIM Lending Process (2015-2018)

As depicted in the figure above, the process starts and finishes at the local branch but includes visits of the loan officer to the home and/or business of the client, group meetings and disbursement of loans at the local mosque or other religious site (in some areas, churches or Hindu temples), and the approval of the loan at the Area Office.

One important characteristic of the programme is the physical proximity to the clients. The branches have a short radius of operations (1.5 miles) which helps branch managers and loan officers to become familiar with the territory where they work. The local branches are small and decorated with basic furniture, and they are many times located in spaces provided by donors, which contributes to maintain low operational costs.

At the beginning of the evaluation project, each area office coordinated six local branches, and this was the level of the organisation in which the loan data was introduced in the information system (local branches worked based on paper). Above the area office, there were regional offices, operating under the direct supervision of the headquarters in Lahore, and each responsible for six area offices. This information was collected during the first field visit to the institution in March/April 2015, when the first wave of the household survey was
implemented. The organisational structure did not suffer significant changes during the evaluation project, although the number of branches strongly increased along with the geographical coverage of the programme as described in the previous section of the thesis.

If there are similarities of AIM microcredit programme with others from traditional MFIs, there are also significant differences, being one of them the family approach. Independently of who is managing the business, other member of the family co-signs the loan contract (normally the spouse). This approach is expected to encourage the family to be involved, or at least, supportive of the business, and to facilitate reaching female clients.

The signature of contracts and the loan disbursements are public acts carried out in religious sites, mostly mosques. In this way, the MFI respects Islamic principles asserting that contracts must be written, and their signature must have witnesses, reinforcing the moral obligation of borrowers to comply with the loan repayment. The disbursement ceremonies are also used to provide ‘social guidance’ on topics such as the importance of education or community cohesion. Finally, the MFI encourages their clients to become themselves donors by making voluntary donations to the institution (Khan et al., 2017).

Pakistan is a country where “the majority of the population is deeply conservative and steeped in different Muslim traditions”, tending to favour and trust institutions with a religious character (Lieven, 2011, p.126), particularly when compared with foreign secular institutions. The compliance with the Islam principles and the links of AIM to the local mosques can, therefore, be seen as a strength, potentially increasing the donations to the institution and the demand for loans, and contributing to excellent loan repayment rates.

6.3 LWC Evaluation Project

6.3.1 Introduction

The AIM evaluation project aimed to assess the microcredit programme developed by the institution at the client level. The impact assessment focused on the four branches in the Punjab province which provide the borrowers profiles uploaded to the crowdfunding platform. Three of these branches were located in the city of Lahore (Badami Bagh, Kot Khawaja Saeed and Khana Nau), and one in the smaller city of Kasur (30 miles from Lahore).

The first contacts between LWC and AIM occurred in the end of 2014, with a pilot questionnaire prepared by the two institutions being tested in January 2015 with a sample of 30 clients. The University of Portsmouth joined the project in February as academic consultant. The core evaluation team was then complete and included two elements from the LWC team which are responsible for the selection and monitoring of the crowdfunding
platform field partners; two elements from AIM, being one a representative from the management team and an operational member; and the University of Portsmouth evaluator/researcher. In different moments of the evaluation project, the LWC coordinator, members from AIM Board, and the two PhD supervisors were also involved.

Figure 6.2 shows the timeline of the project, identifying the project major milestones, including the formal meetings of the partners and the main steps on the implementation of the household survey (which are described in the remaining of the chapter).

![Figure 6.2 – LWC/AIM Project Milestones](image)

Source: Own construction from field notes

It is important to refer that in between meetings, there was a flow of email messages addressing different issues, from the preparation of the field visits to the discussion of the survey results, as well as the parallel activities developed during the project. There were also several bilateral meetings LWC/UoP to discuss the evaluation project taking place either in London or Portsmouth.

The description of the project presented below is based on the analysis framework developed in chapter five, focusing on the central ideas of the UFE model in each of the three main stages of the project.

### 6.3.2 Design of the Evaluation

In the UFE model, the design stage is essential to set up the use of the evaluation lessons and findings by involving the primary intended users (Patton, 2008, p.51). In the Pakistani case study, this stage of the process was affected by time constraints and the conditions of implementation of the project, which translated into a short period of joint reflection of the three partners regarding the design of the project. The experience and knowledge of the partners by the LWC evaluation team as well as the flexibility of the academic consultant to
adapt to the decisions previously taken were important to ensure the best possible research design.

From the beginning of the project, the LWC evaluation team expressed their commitment to use the evaluation findings: “we wouldn’t have got involved with the research if there wasn’t a practical use for it for us or for our partners” (LWC1, p.c., June 4, 2018). Although the initial focus in the communications between the partners was on identification of the outcomes of the programme, the relevance of the lessons learnt through the implementation process became gradually more important, and this was emphasized by the LWC and the AIM evaluation teams in the final interviews.

6.3.2.1 Primary Intended Users: motivations

A utilisation-focused evaluation as described by Patton (2008) starts with the identification of the evaluation project stakeholders, and among them, those who are going to be involved in the implementation of the evaluation and will be able to make use of its findings – the primary intended users. Figure 6.3 shows the stakeholders map for the LWC impact project, with stakeholders being clustered in three groups linked to each of the project partners, and considering a residual group including other external stakeholders.

Figure 6.3 – LWC/AIM Evaluation Stakeholders Map

Source: Own construction
LWC was the commissioner and leader of the evaluation project. LWC aimed to assess the impact of the microcredit programmes funded by the crowdfunding platform and, in this way, gather evidence of the merit of the initiative to present primarily to its supporters, including its ‘mother’ organisation, CARE International UK, and the individual and collective crowd-funders, but also “giving us results we can disseminate to a wider audience, almost to justify our intervention” (LWC1, p.c., June 4, 2018). The evidence was also expected to attract new lenders, and the experience gained during the project to be of use to other LWC partner MFIs, having in mind the extension of the evaluation project.

The first drive for the evaluation was, thus, LWC own accountability, both internally and externally. There was, however, a second motivation associated with the provision of technical assistance to the field partners, which is part of LWC mission, encouraging them to “begin the path for themselves of starting to evaluate their work” (LWC1, p.c., June 4, 2018). Ultimately, the feasibility of the project was dependent on the involvement of the partner MFIs. They would have to be willing and capable to have an active participation in the implementation of the project, since the recruitment of external local consultants was not considered in this first stages of the project.

As previously mentioned, AIM in Pakistan was the first MFI involved. The commitment of the institution to the impact assessment project was expressed by AIM Executive Director in several occasions, including partner meetings and public events, ensuring the fundamental buy-in from management. The institution is rapidly growing in number of clients and geographical coverage, which translates into greater public scrutiny and an increasing need for additional funding. This process has been hampered by some distrust from external actors, particularly the research community whose recognition is seen by the AIM managers as important to legitimate the success of the programme. The manager of AIM (AIM1) expressed this distrust:

Some people from different schools of thought, like academia, researchers, PhD scholars, microfinance professionals all over the world, they come to us and they always ask the same question: either it is a financially sustainable model? So, we said yes, this is a sustainable model, but they don’t believe. They say that OK, but this model is only suitable for the Islamic countries and you cannot apply it in other parts of the world. So we said OK, we need some evidence, but that evidence should be produced by some external party instead of Akhuwat itself. (AIM1, p.c., March 22, 2018)
Therefore, having access to externally validated evidence of the impact of their microcredit programme became a priority for the MFI managers, who were very motivated to participate in the evaluation.

Three other stakeholders connected to AIM play a decisive role in the implementation of the project, despite not having decision making power. These are the internal evaluation team (operational level) and the clients and non-clients participating in the household survey. Non-clients are a particularly sensitive group as their main characteristic is not being linked to the MFI and, in this way, having little or no incentive to be part of the project (Hulme, 2000b). Although the compensation of the non-clients for their participation in the surveys was a hypothesis discussed during the first field trip, the LWC and AIM evaluation teams decided not to give any monetary compensation or gift, assuming this would be seen as inappropriate by the interviewees. As discussed further in this chapter, this opinion partially changed during the implementation of the surveys as the difficulties to locate and interview the non-clients materialised.

Theoretically, other groups interested in the evaluation results are all AIM existing and potential clients, considering a scenario in which the evaluation results are used to improve the microcredit programme, increasing the benefits for the borrowers. Likewise, other AIM funders will have interest in gaining access to valuable information on the institution to fundament their investment choices.

The third partner in the evaluation project is the University of Portsmouth, acting as academic consultant. Adding to the inherent interest on the research topic and the development of partnerships with microfinance practitioners, the active participation in the evaluation project was also crucial for me as PhD candidate. Finally, other actors with indirect interest in the evaluation results include competitors of AIM and other actors in the Pakistani microfinance sector, including regulators, Islamic MFIs in other countries, microfinance consultants and the academic community.

Considering that the primary intended users (PIU) of the evaluation should be those who will most likely use the evaluation results, having the decision-making power to implement changes in the programme, the LWC team and AIM management team were identified as PIU, and I, as evaluator, gave special attention to building the relationships with these two partners.

6.3.2.2 Role of Evaluator

At the early stages of the project, my attention as evaluator focused on two main objectives: encouraging the engagement of the PIU in the necessary tasks related to the design of the
evaluation and guaranteeing that the defined research design would comply with high standards in terms of rigour and ethics.

There was a pre-disposition of the two field partners to be actively involved in the project, but there were different expectations regarding the requirements associated with the implementation of the project and the results that could be realistic obtained given the context and conditions of implementation. Therefore, it was important to stimulate the participation of the partners but, simultaneously, inform them of the limitations associated with the design and methods used in the project and manage their expectations regarding the external credibility of the results.

This implied gradually building a relationship of trust among the members of the evaluation teams. As introduced in section 6.3.1 the team included me as evaluator/academic consultant, two members of the LWC team based in London, and two AIM representatives located in Lahore (one member of the management team and a project manager, responsible for the operations of AIM with the crowdfunding platform). The geographical distance as well as the cultural and professional background differences between the evaluation members made this a demanding task which reflected, for example, in the quote by LWC1: “sometimes the MFIs found it difficult to communicate with researchers from outside” (LWC1, p.c., June 4, 2018).

In this context, it was very helpful that one of the members of the LWC team had British/Pakistani nationality, having a strong understanding of the context of the country and the specific microfinance programme, and being able to speak the local language. By being perceived by the MFI as someone who could understand both perspectives (their own and the evaluator/researcher), he became pivotal in all communications and a crucial element to guarantee the buy-in from AIM management.

In building the relationship between the partners, it was also helpful the development of parallel activities to the evaluation project. In the initial stages of the project, these initiatives included the preparation by the University of Portsmouth of microfinance market briefs on countries of interest to LWC and the co-authorship of a chapter focusing on the AIM experience for a book on Islamic Microfinance (Harper & Khan, 2017).

In the management of the expectations of the partners, I have given particular attention to the language used in all the contacts, clarifying the concepts used and working to develop a common language among the team regarding the evaluation. Equally important at this stage of the process was emphasising rigour in the application of the selected methodologies and
advising on the ethic procedures which should be adopted in order to ensure objectivity, independency, and subsequently, credibility of the project and its results.

6.3.2.3 Household Survey: preparation

The household survey included two evaluation tools, the Poverty Probability Index (PPI) and a purposely built questionnaire. The PPI was included with the objective of assessing the outreach of the programmes and the targeting strategies of the MFIs by measuring the poverty levels of the clients at baseline. The re-application of the tool in the second wave of the survey allowed for identifying changes in the poverty levels over the period, complementing the questionnaire, which was designed to assess the changes in a selection of relevant client outcomes, both at the business and household levels.

The preparation of the household survey involved a series of tasks. One of these was the selection of the outcomes of the programme to be evaluated, followed by the definition of the specific questions to be included in the questionnaire. Given the partially goal-free approach adopted, this selection was firstly conducted by me, based on the review of the microfinance impact literature (presented in chapter three) and the information gathered about the institution and its microcredit programme. The proposed questions were then discussed with the partners in order to adapt them to the local context and the outcomes of interest for the institutions. In this process, a first important negotiation between the partners related to the planned duration of the questionnaires, taking into consideration the negative relation between the length of the interviews and the comprehensiveness of the information (Barnes & Sebstad, 2000, p.42) and the expected limited availability of the respondents. As a result, some of the questions I have initially proposed were dropped. This negotiation process allowed to identify some differences between LWC and AIM interests, particularly in regard to gender related questions, which were considered important by the LWC team, but eventually were the questions taken from the baseline questionnaire.

The final client questionnaire ended up with 16 questions, while the non-client questionnaire included 13 questions. In both cases, the application of the questionnaire was complemented by the 10 questions of the PPI form (see Annex II). The questionnaire included demographic data on the interviewee and his/her household, and both closed and open questions. In the baseline survey, two open questions were introduced to better understand the expectations of clients and non-clients regarding the future evolution of their businesses and households. In the second questionnaire, the open questions aimed to capture the perception of the respondents on the changes occurred in their lives during the period between interviews.
Before application, the questionnaires were translated to Urdu, the official language used by the institution using the back translation technique. The translation to Urdu was done by a staff member of AIM, then the questionnaire was translated back to English by the LWC bilingual team member and the two versions were compared. This inexpensive technique allows for the identification of potential errors (Gaile & Foster, 1996; Barnes & Sebstad, 2000).

The sampling procedure was defined jointly by the LWC team and the evaluator. At baseline, a sample of new clients who had their loans approved and were about to receive the money was selected. This strategy implied that every new client in the four participating AIM branches, since the beginning of the survey, was invited to participate in the study. No other criteria were set to make the sampling process as random as possible. The goal for the number of interviews (500) was set with reference to the number of active clients in the branches, aiming to achieve statistically significant results. Having in consideration the firm expectation of AIM evaluation team members of encountering difficulties locating and interviewing non-clients, the goal for the control group number of interviews was set at 100, a much more modest objective which resulted in an unbalanced dataset. The interviews were to be conducted until the established numbers were achieved.

Besides the considerations regarding the control group, I have made at this stage a number of recommendations regarding the ethical procedures of the interviews. Since the new clients were all clients supported by LWC through the crowdfunding platform, they had already been questioned and had given their consent to participating in research. There was, however, a need to communicate and confirm acceptance regarding the specific objectives and conditions of the project, which was done at the beginning of the interviews, along with other ethical procedures regarding the confidentiality and storage of the information and the right of the participant to not answer any of the questions or withdraw from the interview. In the interviews, the interviewers presented themselves stressing their independency from the MFI and explaining the objectives of the research, as it was simulated during the training sessions before each round of the survey. Finally, a question regarding the consent of the respondents to be contacted in the future and invited to participate in the second wave of the household was included in the questionnaire.

In the preparation of the household survey, I have planned the training sessions for the interviewers who were to be recruited by LWC to implement the survey. The training of the interviewers and the monitoring of the first applications of the questionnaires were the main purpose for the first field trip to Pakistan between 30th March and 5th April 2015.
The recruitment of the local interviewers was led by LWC which supported the related costs, but there was an active involvement of AIM in the process by contacting local Universities to gather candidates and choose the final team of five interviewers. These were finalists or recent graduates from Master Courses in the area of Finance from the Lahore Institute of Management and Technology, who had no formal link with the MFI.

### 6.3.3 Implementation Process

The implementation process began with the field trip of the LWC team and the researcher to Pakistan in March/April 2015. Beyond the training and kick-start of the baseline survey, the visit was also used by the researcher to experience *in loco* the culture of the country and to learn more about the institution. It included visits to the branches involved in the study, where it was possible to meet field staff and a few clients of the institution.\(^{26}\) Although this was a short visit, the experience was important to better understand the context of implementation both of the microcredit programme and the evaluation project. Among other questions, the field trip was important to clarify the characteristics and procedures associated with the microcredit programme. In the same sense, the co-writing of a book chapter on the institution with the evaluation team members from LWC and AIM helped consolidating the information previously collected.

#### 6.3.3.1 Primary Intended Users: behaviours

The LWC evaluation team was throughout the implementation of the project involved in all the stages of the process, with its members alternating their attention between the two MFIs participating in the project. In the case of Pakistan, the member of the team with Pakistani nationality was considerably more involved in the AIM evaluation, leading the two field visits to the country and, as mentioned above, being the pivot of all communications between the evaluation team members. His previous relationship with the institution managers as well as the knowledge of the local language granted him the respect and trust of the AIM evaluation team members, which facilitated the implementation of the impact project. In terms of prompting action and use, it became evident that the intervention of this LWC member was essential, confirming Patton’s premise regarding the relevance of the personal factor.

The AIM evaluation team was very committed regarding all aspects related to the implementation of the household survey. In certain periods, the evaluation demanded a strong time commitment from the project manager involved (AIM2). During the two periods

---

\(^{26}\) The visits included the branches in the city of Lahore (Badami Bagh, Kot Khawaja Saeed and Kahna Nau). The visit to Kasur, closer to the Indian border, was not authorised for security reasons.
of interviews, he estimates that half of his working time was dedicated to supporting and monitoring the recruited interviewers during the application of the questionnaires. This was even more intense, although for a shorter period, during the computerisation of data, which was conducted by him in both waves of the survey (AIM2, p.c., March 22, 2018). AIM participation was also very active in the parallel activities involving the three partners.

There was, however, a more passive attitude whenever feedback or comments on the produced reports were requested. This was particularly noticeable when the preliminary report based on the baseline survey data was shared within the team and both LWC and AIM were invited to comment on this internal report. The LWC evaluation team shared their comments, but the few remarks obtained from AIM resulted from direct questions raised during the formal partner meetings in September 2015 and November 2016. In both cases, the answers gave valuable indications to the evaluator, and they were taken in consideration when preparing the second questionnaire.

6.3.3.2 Role of the Evaluator

During the implementation of the project, I had two main roles. One was related to the implementation of the surveys, preparing the questionnaires for the second wave of the survey, leading the training of the interviewers and analysing the collected data.

The analysis of the baseline survey results was shared with the partners through the above mentioned internal preliminary report. This report was mainly descriptive, characterising the clients and non-clients in regard to the studied variables, but included also an analysis of the differences and similarities between these two groups and among different segments of clients. The focus in the analysis on the heterogeneity of the client sample was deliberate, since the recognised diversity of the sector is often not explored at an empirical level and there was an expectation that a more detailed knowledge of the different client segments could open opportunities to improve the services provided by the institution, and consequently, the potential impact of its intervention.

The other relevant role I had was the continuous effort to guarantee the involvement of the evaluation teams and the deepening of the relationships between the members using multiple ways of communication and developing parallel initiatives. The internal report was one of the instruments used to strengthen the relationships between the partners and cross-validate the interests and perceptions of the PIU. It was presented to the AIM project manager in the partners meeting in September 2015 (London). This event was further explored with the invitation of the AIM representative to visit Portsmouth where he met
some of the University of Portsmouth (UoP) researchers and presented AIM to students of the university.

During the implementation of the impact project, two other parallel activities should be highlighted. The first was the visit of AIM Executive Director to Portsmouth, in November 2016, to participate in a project meeting and a workshop organised by the Development Studies Group of the University on the experience of AIM as an Islamic microfinance provider. The second regards the introduction by the UoP, as organisers of the 5th European Research Conference on Microfinance, of a plenary session on Islamic Microfinance with the participation of LWC as session moderators and AIM as an example from the field. In the two initiatives, the involvement and active participation of a representative of AIM Board, strongly signalled the commitment of the institution towards the partnership and the evaluation project.

6.3.3.3 Household Survey: implementation and analysis

The household survey was implemented in two waves: April-May 2015 and February–May 2017. The baseline survey was applied to a sample of 500 AIM new clients and 100 non-clients. The survey was repeated in 2017, after all the clients have finished repaying the first loan; 63% of them had a second loan approved or about to be approved and an additional 13% were thinking of applying to a second loan. The second questionnaire included some questions identical to the first questionnaire, new queries about the changes in the selected outcomes, and two open questions aiming to capture the perception of the interviewees regarding any observed changes. The PPI form was also completed in both waves of the survey.

Table 6.4 shows the geographical distribution of the respondents, giving a first impression on the attrition figures, which as expected in Pakistan affected more the sample of non-clients.

<table>
<thead>
<tr>
<th>Branch</th>
<th>Clients 2015</th>
<th>Clients 2017</th>
<th>Non-clients 2015</th>
<th>Non-clients 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badami Bagh</td>
<td>79</td>
<td>74</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td>Kot Khawaja Saeed</td>
<td>105</td>
<td>91</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>Kahna Nau</td>
<td>132</td>
<td>107</td>
<td>32</td>
<td>10</td>
</tr>
<tr>
<td>Kasur</td>
<td>184</td>
<td>175</td>
<td>19</td>
<td>12</td>
</tr>
</tbody>
</table>

27 The 5th European Research Conference on Microfinance took place in Portsmouth, between 12 and 14 June 2017. It was co-organised by the University of Portsmouth, the European Microfinance Platform (e-MFP) and CERMI (Centre for European Research on Microfinance).

28 The second wave had a complementary phase of interviews with non-clients in November 2017.
The first wave of the survey was completed after seven weeks of interviews. By 18 May 2015, AIM confirmed the end of the interviews and the beginning of data introduction. Although the questionnaires were translated to Urdu, with the interviews carried out in Urdu and the local language Punjabi (when necessary), the interviewers filled in the questionnaires directly in English.

Between May and July, there was a frequent exchange of messages between the partners to clarify questions regarding the data treatment. By 8 August, the final databases (clients and non-clients) were shared by AIM. This allowed me to analyse the data and prepare the preliminary survey report. The report set the baseline for the longitudinal study.²⁹

Among the results, there were two especially relevant in the context of the evaluation process. The first is related to the use of PPI. The computed poverty scores were used to identify the poverty levels among the clients and compare them with the non-clients. Since AIM had not been using the tool in recent years, the integration of PPI in the baseline survey corresponded to a first application, not allowing at this stage to track changes in the poverty levels.

The PPI scores were computed for all respondents, and the correspondence tables provided by the Grameen Foundation were used to identify the likelihood of each group being poor considering the national poverty line, as well as the international poverty lines of $1.25/day and $2.50/day 2005 PPP (Purchasing Power Parity).

No statistically significant differences were found between clients and non-clients for the vast majority of the indicators included in the study, which was encouraging in terms of the strength of the comparison between the two groups, despite the difference in the sample size.

For the clients, the PPI scores concentrated between 40 and 90, with an average of 63.8.³⁰ This figure represented a probability of the average LWC/AIM client being poor of 5.1% and 1.3% when using the national and the $1.25/day poverty lines, respectively. The poverty likelihood increased to 54.8% using the $2.50/day poverty line.

---

²⁹ Preliminary report available upon request. This is a project internal document.
³⁰ PPI scores range between 1 and 100. As higher the score, lower is the probability of the household to be considered poor in reference to the selected poverty line.
This result initially surprised AIM and LWC, but the discussion that followed enabled this result to be contextualised and better understood. A first important factor to consider is that PPIs are tools developed at country level, with comparisons using national averages as reference. Those clients located in the more economically developed areas, such as the Punjab in Pakistan, by the nature of the tool, score the maximum in the location question as they are expected to have a lower probability of being poor than those living in poorer regions. This sensitivity to regional disparities is similar to the situation described by Boucher (2014, p.15) in Peru where the existence of differences in government policies, infrastructure development and patterns of consumption across the regions is expected to influence the relationships between poverty and the 10 indicators included in the PPI for the country.

In addition to the regional disparities, the PPI for Pakistan is outdated. It was developed in 2009, based on national statistical data from 2005/2006, and unlike countries such as Ecuador, there was no update of the tool during the period of implementation of the evaluation project. As its own developer recognises, this factor reduces the accuracy of the instrument in the identification of poverty levels (Boucher, 2014; Schreiner, 2010).

The fact that the scorecard was outdated meant that also the supporting materials were outdated, including the look-up correspondence tables which still used the international poverty lines of $1.25/day and $2.50/day (2005 PPP), while the new reference values are $1.90/day and $3.20/day (2011 PPP). This situation led the evaluator to analyse the results and report them in the final report using the $2.50/day poverty line, closer to the new values.

These limitations require caution in the way results are presented and contextualised and help explain the reluctance of many managers from MFIs regarding PPI (Boucher, 2014), situation that is exacerbated if there is only a one-time application of the tool. This critical perspective on the instrument echoed in the opinion of the AIM manager, who considered the non-statistical methodologies applied by the institution during the loan application process more trustworthy in the identification of poverty levels than “something that has been obsolete” such as the PPI for Pakistan developed by Mark Schreiner and his team (AIM1, p.c., March 22, 2018).

The appraisal of the poverty level of the AIM loan applicants was “not formally designed, we did not do like a PPI questionnaire that gives us some scorecard” (AIM1, p.c., March 22, 2018). The process relied mainly on the information collected by the loan officer during the screening process when questioning the applicant and other persons of interest and observing the living conditions of the household.
It is important to highlight, however, that the limitations identified are not sufficient to explain why the LWC/AIM clients were not among the poorest segments of the population according to the average PPI scores. There are two other elements that need to be acknowledged. One respects to the sample of the LWC/AIM clients. The comparison of the average loan size calculated for the study sample and for the total loan portfolio of the institution at baseline, indicates that there might be a slight difference in the risk levels associated with the borrowers selected to the crowdfunding platform (Afonso, 2016).31

The other aspect relates to the target of the institution as presented by a management representative during the first field trip. AIM targets low income entrepreneurs with its microcredit programme, the economically active poor as described by Robinson (2001), who are expected to integrate the upper segment of the 40% poorest in the country. For the poorest segments of the population, microcredit is not considered to be an effective tool, which is in line with Khandker’s (1998) analysis of the Bangladeshi experience. AIM works with these poorer populations through Zaqat donations and the provision of non-financial services, including the clothes recycling programme.

The second result to be highlighted in the baseline survey refers to the differences observed between segments of clients, particularly when adopting a gender perspective. Statistically significant differences between male and female clients were found for many of the studied variables. Female clients were, on average, older but less experienced in the business. They had, on average, lower personal incomes and were more likely to be poor than male clients. There were, however, no statistically significant differences when analysing income and expenditure at the household level (Afonso, 2016). These were not surprising results considering the macro-context described above and the findings of several microfinance studies focused on gender inequalities in Pakistan (Asim, 2009; Safavian & Haq, 2013; Zulfiqar, 2017).

On his first look to these results, AIM2 suggested that the identified income gap might have resulted, at least in part, from the lesser time dedicated to the business by women clients as they were more often part-time entrepreneurs. To check this hypothesis, a question on the time dedicated to the business was included in the second questionnaire. Interestingly, although the average time dedicated to the business was lower for the female clients (60h/week compared with 67h/week for the male clients), the difference was not statistically significant and the share of female entrepreneurs reporting less than 40h/week was small

31 Average loan size reported by the institution to Mix Market for 2014 was $116, while the sample average loan size was $197 (1USD=101.80PKR - 04.08.2015).
(9.8%), indicating that this is probably not the main factor explaining the income gender gap (Afonso, 2016).

Significant gender differences were also found with respect to savings. The survey data suggests that women valued savings more than men and saved more frequently. This conclusion is similar to previous studies, including Ashraf, Karlan and Yin (2006), who conducted a randomised control trial in the Philippines, finding that women were more likely to take up savings products than men, especially if there was a commitment component in the product.

In October 2016, the preparation of the second wave of the survey started, with me preparing an initial draft of the second questionnaire, which was discussed and agreed with the partners. The second questionnaire was slightly longer than the baseline survey, which was accepted by the field partners based on the experience during the first questionnaire.

The following month, there was a partner meeting organised at UoP. In the meeting, following discussions between me and the LWC team, the LWC representative in the meeting proposed to include in the second field trip, planned to February 2017, the implementation of focus groups with clients and branch managers. The objective was to complement the data collected from the survey with qualitative data. Given that the proposal was well received by the AIM board member present in Portsmouth, I have started preparing the focus groups considering the implementation of three groups with female clients, male clients and branch managers, respectively. Along with the definition of the objectives and initial list of topics to be addressed, I have also prepared recommendations regarding the practical conditions (place, translation, etc.) and the ethical procedures to be considered by the partners in the implementation of the focus groups. However, unforeseen events affecting the second field trip resulted in the cancelation of these plans.

These unforeseen events included a delay on the visa process which shortened the visit of the evaluator from two weeks to five days, with priority being given to the training of the interviewers and the start of the household survey; the occurrence of a suicide bomb attack on the first day in Lahore, limiting the team movements in the following day; and time constraints from the AIM evaluation team as they were involved in the preparation of a large disbursement event involving approximately 10,000 borrowers, which was scheduled to happen in the week before the visit.

Despite these limitations, it was possible to organise a meeting with six branch managers, including three from the impact project branches and three from other AIM branches located in Lahore. The meeting was used in the evaluation as a complementary source of information.
in regard to the lending process and the characteristics of the clients in the different branches, allowing for cross-validating the data collected in the visits to the branches and the questionnaires.

The second wave of interviews initiated on the third week of February 2017 and unfolded for 15 weeks, double the time dedicated to this stage at baseline. Attrition problems were expected, and this simple time statistic illustrates part of the challenge. As table 6.4 illustrates, the attrition rate was especially strong for the non-clients group (48%), being very reasonable in the case of the clients (10.6%).

The interviewers, a new batch of graduate students recruited by LWC with the logistic support of AIM, confirmed that they had difficulties locating and securing the participation of the non-clients. In the final interview, AIM2 refers also to some problems with a small group of clients who have delayed their repayments during the period or had a second loan not approved by the institution (AIM2, p.c., March 22, 2018). These difficulties were aggravated by internal problems in the branch of Kahna Nau (reported to the researcher by AIM in April), which resulted in the dismissal of its branch manager and some of the loan officers and affected the interviewing process. Consequently, when the ‘final’ dataset was sent by AIM in the beginning of June, there were no interviews with non-clients from Kahna Nau, and the number of valid non-clients interviews was of just 39.

The potential bias associated with the available data, predominantly from Badami Bagh and Kot Khawaja Saeed, was confirmed when applying non-parametric tests to this sub-group of non-clients. They differed from the baseline control group, namely in the poverty levels, with their average PPI scores being significantly different (higher) from the client group in both waves of the survey. The impact of this bias on the credibility of the results were the main argument used to convince the AIM evaluation team to try to locate and facilitate the interviews with some of the non-clients from Kahna Nau. These additional interviews took place in November and allowed for the final number of interviews with non-clients to raise to 52.

6.3.4 Interpretation and Use of Findings

6.3.4.1 Primary Intended Users: perceptions and intentions

The main results of the evaluation were included in the final report, which was prepared following discussion with the partners (Afonso, 2018b; see Annex IV). As it happened with the preliminary report, comments were mostly provided by the LWC team and were accommodated in the final version.
This report was prepared having in mind a broader audience but included information on the methodologies employed in the evaluation, and as so, some of the language used was necessarily technical. In this way, the report served a part of the public that LWC was addressing, particularly CARE International UK and other LWC corporate supporters such as Hogan Lovells and Bluefrog, but it was too long and technical to present to their individual crowd-funders and to use as a marketing tool. Considering this as one of the needs of the PIU, the evaluator prepared a shorter version of the report focusing on the main results (Afonso, 2018a). This brief overview of the findings was used by LWC in their social media communications and was presented at the LWC funders Christmas workshop, which took place in December 2017. The full report can be consulted in Annex IV.

The extended version of the report has been shared and discussed by the LWC team with other teams and country offices of CARE, which have been showing interest in the type of partnership established with the University of Portsmouth. According to LWC1, CARE have been implementing evaluations of their programmes in different areas of development (education, health, water and sanitation, etc.) using primarily the services of consultancy firms. The partnership established within the LWC evaluation project has been perceived as offering advantages in terms of the impartiality of the analysis and the cost associated with the implementation of the project (LWC1, p.c. – part 2, June 4, 2018). Equally, the managers of the MFIs considered the project to be more credible due to the participation of an academic partner as illustrated by the following statements: “they (MFIs) can now present these findings to other investors and say look we had our work analysed by an impartial academic researcher and these are the results” (LWC1, p.c., June 4, 2018); “information which is prepared by reliable and professional people” (AIM1, p.c., March 22, 2018).

Looking back to the overall motivations of the LWC team identified in chapter five (5.4), the complementary objective of collecting additional information on the MFI, in line with Hulme (2000b), is of particular relevance in AIM due to the institutional culture and contextual factors which make the MFI less comfortable sharing information with external parties.

The results of the evaluation project in the two selected countries were also evaluated by the LWC team in regard to the continuation of the present project, with the confirmed implementation of a third wave of the survey in Pakistan and Zimbabwe, and the extension of the project to other field partners. These issues are dealt with more detail in chapter eight when discussing the common project approach to evaluation subjacent to the LWC impact project.
Looking into the AIM evaluation team perceptions and intentions regarding the evaluation, in the final interviews with the two AIM team members they were both very positive regarding the usefulness of the evaluation and confirmed the intention of the institution to continue developing the project, supporting the implementation of the third wave of the survey in 2019 (AIM1 and AIM2, p.c., March 22, 2018). The representative of management stressed the overall positive results of the evaluation at this stage and their expectation that these results will be even more positive when the third wave of the survey is implemented: “the results are quite satisfactory right now, but we want more improvement than that” (AIM1, p.c., March 22, 2018). He reiterated the intention of the institution to share the final report and results with their partners in Pakistan, especially the regional governments which are the main funders of the microcredit programme, and with all others interested in the performance of the institution, including donors and academics.

Both members of the team also referred to the internal benefits for the MFI of participating in the project. They shared the plans of the institution to, along with the implementation of a new information system, create their own internal evaluation system, which will allow them to generate evidence of the results of their work. Although recognising the different levels of credibility of the data produced internally and externally, AIM1 emphasized the importance of not being dependent of projects led by external institutions – “research is the motivation for us to induce the latest techniques of measuring impact at our level, instead of waiting for universities”, which have their own objectives and timings. In this decision, the participation in the evaluation project was essential as it gave them an understanding of what the process involves and how it can be practically implemented. AIM2 mentioned the intention to use part of the questions included in the household survey in their future internal evaluation system: “we will try to use some of the questions of the questionnaire in our MIS to tell the impact” (AIM1 and AIM2, p.c., March 22, 2018).

The decision to internalise the evaluation process is not surprising given the culture of the institution and its recent growth, which grants AIM the access to resources that most probably were not available prior to the evaluation project.

6.3.4.2 Role of Evaluator

The most active and time-consuming tasks for me as evaluator concentrated in these final stages of the project in the interpretation of the results of the household surveys and the preparation of the reports. These two tasks implied, however, two different attitudes. In the data analysis and interpretation of the results, I have led the process and independently interpreted the findings, looking for depicting an impartial portrait of the changes in the lives
of the clients and non-clients during the evaluation period. In this perspective, I have included in the reports also the limitations associated with the research, giving particular attention to the unbalanced dataset between clients and non-clients, and to the weaknesses resulting from an analysis based on average indicators.

The attitude was more reactive and adaptive, as prescribed by the UFE model, in the decisions related to the results of interest to the PIU to be included in the reports, as they were understood by the evaluator. The results of interest became clearer as the relationship between the partners evolved and there was a better knowledge of the evaluation team members and the culture of the institutions. Indeed, this growing knowledge about the institution and its intentions has led me to adjust the process regarding the final report of the longitudinal study, with no internal report being prepared at this instance. Instead, the partners were invited to comment on a working version of the final report.

6.3.4.3 Household Survey: results and reports

The results of the survey were overall positive for the institution with the average business profits, income and poverty indicators, among others included in the study, improving during the period for the new AIM clients, as Table 6.5 exemplifies.

Table 6.5 – Client Average Monthly Income and Expenses (Survey)\(^{32}\)

<table>
<thead>
<tr>
<th></th>
<th>Personal Monthly Income</th>
<th>Household Monthly Income</th>
<th>Household Monthly Expenses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clients</td>
<td>Clients</td>
<td>Clients</td>
</tr>
<tr>
<td>2015</td>
<td>13,380PKR</td>
<td>24,479PKR</td>
<td>17,007PKR</td>
</tr>
<tr>
<td>2017 (2015 values)</td>
<td>16,041PKR</td>
<td>29,377PKR</td>
<td>17,434PKR</td>
</tr>
<tr>
<td>Real Variation (%)</td>
<td>+19.9%</td>
<td>+20%</td>
<td>+2.5%</td>
</tr>
</tbody>
</table>

Source: Own construction from LWC impact project data
Notes: Data includes 270 observations (those with available data for the three indicators)

Although the limitations associated with the quantitative survey do not allow for definitive conclusions on the formal impact of the microcredit loan, the data analysis based on the application of quantile regression (Table 6.6 below) showed that for some segments of the clients, the loan was one of the factors contributing to the changes observed.

\(^{32}\) The number of non-clients with data available for the three indicators was very small, thus, it was not included here. For the 46 non-clients interviewed in the second wave of the survey with data available for the personal monthly income, the real variation of the indicator was of +16%.
Table 6.6 – Quantile Regression Results - AIM

Dependent variable: PPI Variation  
Sample: 499 – Included observations: 383

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Type</th>
<th>Coefficient (Quartile)</th>
<th>Coefficient (Quartile)</th>
<th>Coefficient (Quartile)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (2017)</td>
<td>Q</td>
<td>0.006200 ***</td>
<td>0.006694 ***</td>
<td>0.009906 ***</td>
</tr>
<tr>
<td>Female</td>
<td>D</td>
<td>0.053218</td>
<td>0.006750</td>
<td>0.031259</td>
</tr>
<tr>
<td>Married</td>
<td>D</td>
<td>-0.110948 **</td>
<td>-0.041616</td>
<td>-0.093870 *</td>
</tr>
<tr>
<td>Badami Bagh Branch</td>
<td>D</td>
<td>0.089220 ***</td>
<td>0.046650</td>
<td>0.100315 **</td>
</tr>
<tr>
<td>Kahna Nau Branch</td>
<td>D</td>
<td>0.012617</td>
<td>-0.006522</td>
<td>0.017976</td>
</tr>
<tr>
<td>Kot Khawaja Saeed</td>
<td>D</td>
<td>0.157151 ***</td>
<td>0.080715 *</td>
<td>0.074695</td>
</tr>
<tr>
<td>Trade</td>
<td>D</td>
<td>0.028370</td>
<td>0.037280</td>
<td>0.029608</td>
</tr>
<tr>
<td>Services and Education</td>
<td>D</td>
<td>-0.016557</td>
<td>-0.039921</td>
<td>0.011931</td>
</tr>
<tr>
<td>Production and Construction</td>
<td>D</td>
<td>0.034542</td>
<td>0.040377</td>
<td>0.002536</td>
</tr>
<tr>
<td>Primary School</td>
<td>D</td>
<td>0.067060 *</td>
<td>0.025416</td>
<td>0.032394</td>
</tr>
<tr>
<td>Secondary School</td>
<td>D</td>
<td>0.076846 **</td>
<td>0.042123</td>
<td>0.000686</td>
</tr>
<tr>
<td>Higher Education</td>
<td>D</td>
<td>0.046482</td>
<td>-0.000256</td>
<td>-0.004743</td>
</tr>
<tr>
<td>Household Size</td>
<td>Q</td>
<td>-0.016481 ***</td>
<td>-0.005000</td>
<td>-0.001460</td>
</tr>
<tr>
<td>Business Time (2017)</td>
<td>Q</td>
<td>0.002395</td>
<td>-0.000742</td>
<td>-0.001830</td>
</tr>
<tr>
<td>Working Hours/Week (2017)</td>
<td>Q</td>
<td>0.001761 **</td>
<td>0.002213 ***</td>
<td>0.002754 **</td>
</tr>
<tr>
<td>PPI_2015</td>
<td>Q</td>
<td>-0.006205 ***</td>
<td>-0.005910 ***</td>
<td>-0.007338 ***</td>
</tr>
<tr>
<td>Loan</td>
<td>D</td>
<td>0.046636</td>
<td>0.054611</td>
<td>0.198767 ***</td>
</tr>
</tbody>
</table>

Source: Afonso (2018b)
Notes: Q – quantitative variable; D – dummy variable (1, 0). *** - significance level of 1%, ** - significance level of 5%, * - significance level of 10%

As indicated in Table 6.6, in the quantile regression, the dependent variable used in the analysis was the variation of the PPI scores during the period. Notwithstanding the limitations of the tool identified previously, which advise against the determination of client poverty levels per se, the analysis of PPI variations for different segments of clients and non-clients over time generated useful data. In particular, the relatively stronger variation of the average PPI scores for the more marginalised groups, particularly women and illiterate clients, potentially contributing to reduce the poverty gap for these groups, can be seen as an encouraging result in terms of the social mission of the institution (Afonso, 2018b).

Another relevant finding, which was consistent in both waves of the survey, was the identification of heterogeneity among the sample clients. This heterogeneity should be acknowledged in the evaluation of the microcredit programmes as it would be expected to have consequences on the potential impact of the microcredit loans. The analysis was conducted by applying non-parametric tests (Mann-Whitney) to sub-groups of clients based on characteristics such as gender, location, educational level, type of activity, house ownership and programme participation. The statistically significant differences found were
included in the final report, which can be consulted in Annex IV. In the report, the methodology applied is described, including the list of the independent variables analysed. Table 6.7 shows the differences found to be statistically significant in terms of gender (female) and educational level (illiterate).

Table 6.7 – Mann-Whitney Tests for Gender and Educational Level

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Female</th>
<th>Average Male</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPI Variation</td>
<td>0.12</td>
<td>0.08</td>
<td>20,336 **</td>
</tr>
<tr>
<td>Real Personal Income Variation</td>
<td>1.01</td>
<td>0.36</td>
<td>13,415 ***</td>
</tr>
<tr>
<td>Loan Amount</td>
<td>19,750 PKR</td>
<td>20,120 PKR</td>
<td>20,016 ***</td>
</tr>
<tr>
<td>Age</td>
<td>41.5 years</td>
<td>37.8 years</td>
<td>17,571 ***</td>
</tr>
<tr>
<td>Working Hours/Week</td>
<td>60 hours</td>
<td>67 hours</td>
<td>11,118 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>14,971 PKR</td>
<td>22,158 PKR</td>
<td>10,061 ***</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>60.1</td>
<td>65.7</td>
<td>18,539 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Illiterate</th>
<th>Average Other</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPI Variation</td>
<td>0.13</td>
<td>0.06</td>
<td>21,641 **</td>
</tr>
<tr>
<td>Age</td>
<td>41.2 years</td>
<td>37.2 years</td>
<td>18,323 ***</td>
</tr>
<tr>
<td>Business Time</td>
<td>10.4 years</td>
<td>9.5 years</td>
<td>21,839 **</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>57.9</td>
<td>68.8</td>
<td>15,587 ***</td>
</tr>
<tr>
<td>PPI 2017</td>
<td>62.5</td>
<td>70.7</td>
<td>17,658 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>17,435 PKR</td>
<td>21,469 PKR</td>
<td>16,886 ***</td>
</tr>
<tr>
<td>Household Income 2017</td>
<td>28,535 PKR</td>
<td>34,815 PKR</td>
<td>10,505 ***</td>
</tr>
<tr>
<td>Household Expenses 2017</td>
<td>16,989 PKR</td>
<td>20,578 PKR</td>
<td>19,460 ***</td>
</tr>
</tbody>
</table>

Source: Afonso (2018b)
Notes: *** - significance level of 1%; ** - significance level of 5%

These results corroborate the findings from the quantile regression presented previously in regard to the variation of the PPI scores, i.e., there was a relatively higher increase of the scores for the average female and illiterate clients, in comparison with the male clients and the clients with formal education, respectively.

Client heterogeneity is, however, not an easy topic for MFIs in general, especially when the microcredit model in place is based on very low operational costs. Any customisation of the programmes to accommodate the differences between client segments will likely imply changes and increasing costs. This expectation, along with the accountability-driven motivation of the MFI, helps explaining the focus of its representatives on the average analysis and global results. In their communications with the partners, AIM managers did not give much importance to the identified heterogeneity or to the possibilities of exploring more in-depth the findings of the survey using qualitative methodologies.

6.4 Discussion

The previous sections of the chapter describe the evaluation process implemented in Pakistan, based on the UFE model framework. This description provides the necessary elements to start answering the first research question and sub-questions introduced in
chapter one, identifying the challenges and lessons learnt from the implementation of the impact project in this specific setting.

Research Question 1.1: What are the motivations of different stakeholders regarding the design and implementation of an evaluation process?

The first issue to be addressed in this regard is the motivations of the stakeholders to participate in the evaluation. This is a topic rarely covered in the description of evaluations, as exemplified in chapter three with the papers of Navajas et al. (2000) and Pitt and Khandker (1998). This omission is likely to be associated with the assumption that the participant institutions were willing to do so.

In the case of the LWC evaluation project, not surprisingly, the decision to implement the assessment project fits within the mainstream accountability-drive observed in the microfinance sector. The desire to provide evidence of the outcomes of the work developed was not, however, the only motivation for the crowdfunding platform. A complementary goal for LWC was to help the partner MFIs to increase their scale and outreach, by providing also technical assistance. The area identified by LWC and the partner MFIs as requiring more support was the measurement and management of social performance.

This impact evaluation project is part of a broader effort of LWC in this matter in which the triangulation of different sources of information is central. The LWC team requires the field partners to provide quarterly data on a group of 18 financial and social indicators; visits regularly the MFIs, where among other diligences, the LWC representatives visit and interview a sample of borrowers and staff; and more recently, has been facilitating the access of the MFIs to social ratings provided by a qualified ratings agency (LWC1 and , p.c., June 4, 2018), but with this evaluation project they were seeking a “much more rigorous way to look at the impact of LWC” (LWC2, p.c., June 4, 2018).

In the final interview, LWC1 confirmed the intention of the institution to eventually “extend this research to all our (14) partners” (LWC1, p.c., June 4, 2018). This perspective was presented as one of the reasons to adopt what was identified in chapter five as a ‘partially goal-free approach’, not discussing previously with the partners the underlying theory of change associated with each microcredit programme or the specific outcomes and evaluation criteria to be used in the project. The LWC team left to me as academic consultant the initiative to suggest the latter. The discussion between the partners in the preparation of the first wave of the survey in Pakistan as well as the desk research I have made on the institutions allowed to understand the main implicit objectives of LWC and likewise, of the partner MFI.
Looking at the same question from the perspective of the MFIs, in many cases, the institutions participate in the evaluations uniquely to comply with the requirements of external donors or investors (Woller et al., 1999, p.72), and in several projects such as the one described by Podems (2007), without a genuine interest in the process. These reservations towards evaluation may arise from unfamiliarity with the evaluation process, resulting in scepticism about the worth of the evaluation or fear of being judged. They can also derive from suspicious regarding the intentions of the external evaluators or their capability to understand the specific context of the institution. In these cases, “evaluation anxiety and fear” can result in reluctance to conduct evaluations (Patton, 2008, p.44).

In what concerns the Pakistani MFI, this was not the case, with a strong buy-in from the top management, which was accompanied by a strong commitment of the evaluation team members, particularly at operational level. In the case of AIM, there were a number of conditions that could have hindered the participation of the MFI. The institution had never conducted an evaluation of its microcredit programme and was “doing nothing” in terms of assessing social performance, according to LWC1 (LWC1, p.c., June 4, 2018). Adding to this, the previous experiences with academics seemed to have resulted in mutual mistrust: “sometimes the MFIs find it difficult to communicate with researchers from outside” (LWC1, p.c., June 4, 2018); “they don’t believe” (AIM1, p.c., March 22, 2018), which could have been reinforced by the geographical and cultural differences between the evaluation team members.

It did not happen this way, which cannot be dissociated from the increasing need felt by the MFI managers to provide evidence of the social results of its microcredit programme and to professionalise operations, in a context of intensive growth and transformation of the NGO into a regulated non-financial banking institution. The accountability perspective, external and internal, remained the main drive for the AIM management throughout the project, which translated into the attention given to the final report – “we want this report. It will be very useful for us”, and into the availability to gather and apply the necessary resources for the implementation of the evaluation (AIM1, p.c., March 22, 2018).

Equally decisive for the MFI commitment was the role of the LWC team member with dual nationality, who was clearly trusted by the AIM managers. Without this factor, the general mistrust or lack of understanding regarding outside researchers would have been more difficult to overcome. This analysis is in line with the importance attributed in the UFE model to the personal factor, and also explains my continuous effort in this matter, notably by encouraging and participating in parallel activities to the project involving the AIM board and the AIM and LWC evaluation teams.
The final note regarding the motivations of the stakeholders is to acknowledge that AIM clients and non-clients, although critical for the successful implementation and the results of the evaluation, were not consulted in regard to the project design and execution and were also not included in the initial feedback process, after the application of the second wave of the survey. This feature of the project is mirrored in the UFE approach adopted in the PhD research and illustrates what is considered one of the limitations of the use branch evaluation models, the fact that information is largely provided to the most powerful stakeholders (Wilson and Mertens, 2012) creating opportunities for co-option by the user groups (House, 2003; Pawson & Tilley, 1997; Stake, 1983).

Given the conditions of implementation of the project, it was difficult to consider a wider participation of AIM clients. The continuity of the project should open opportunities to bring this question into the discussion of the next stages of the evaluation, however, considering the characteristics and declared objectives of the Pakistani MFI, significant changes to the process at this level should not be expected.

*Research Question 1.2: What are the main challenges in the field implementation of an evaluation in microfinance?*

The field implementation of the evaluation presented several challenges. It is, therefore, interesting to realise that all evaluation team members from LWC and AIM, when asked about the main challenges in the implementation of the project, identified one specific challenge, the control group.

Setting up a comparison group with similar characteristics to the clients was pushed initially by me taking into account the methodologies to be applied in the project defined by LWC. It appeared at first sight an achievable task, considering that the areas of the branches included in the research are highly populated, with strong entrepreneurship focus and low levels of financial inclusion. The AIM evaluation team was, however, less optimistic about the participation of the non-clients, and there was at the beginning some disregard of the importance of the comparison group in the evaluation.

In the initial conversations among the team in the first field trip (April 2015), it was suggested by a member of the AIM team to give a small gift to incentivise the participation of the non-clients, without this having the connotation of a payment. Hulme (2000b) considered that non-monetary gifts in this context can be seen as ethical, in the sense of compensating for the time and other costs the interviewees may support to participate in the interviews.

The suggestion, however, has not been followed up in the implementation of the surveys in Pakistan. This possibility was brought up again after the application of the second survey,
when the attrition problems were detected, but it was late to significantly change the situation at this stage. There was, as mentioned above, an effort to minimise the problem with a second phase of interviews with non-clients in Kahna Nau, aiming to increase the numbers and balance the composition of the group.

Apart from the consensus regarding the control group, other challenges were identified (and valued) by the evaluation team members in the final interviews. LWC1 considered time management (“finding the time”) as the major challenge for the LWC evaluation team, and to a lesser extent for the field partners (LWC1, p.c., June 4, 2018). LWC2 underlined the need to conciliate different interests and perspectives on the evaluation, particularly in the initial stages - “try to put together those initial questionnaires that also met the requirements of different people”, and having in mind the different locations and cultural backgrounds of the partners (LWC2, p.c., June 4, 2018).

AIM1 described the implementation of the evaluation as a series of challenges - “these were all challenges that we faced” (AIM1, p.c., June 4, 2018), given the unfamiliarity of the institution with the process, which were successfully overcome by the partners. The operational member of the team was asked to detail some of these challenges, identifying questions related to the quality of the questionnaires (especially during the baseline survey), and the difficulties to guarantee the participation of the respondents, particularly the non-clients, in both waves of the survey. Cultural characteristics resulting in reluctance to share personal information were pointed out as a major challenge for the interviewers in the field – “because Pakistani culture, people don’t share their information to any person”, which had consequences both in terms of attrition and the quality and completeness of the questionnaires (AIM2, p.c., March 22, 2018). These questions were relevant in the implementation of the surveys, despite the attention given to them during the training of the interviewers.

Not mentioned by the PIU in the final interviews, but of importance in the process, was the challenge of language. Although English is spoken by many people in Pakistan, being this one of the reasons pointed by the LWC team to choose AIM to participate in the project (LWC1, p.c., June 4, 2018), AIM clients had low levels of literacy, with 47% of the client sample not having formal education (Afonso, 2016), and the visits to the branches and the group meeting with branch managers showed that the field staff in general did not speak (fluently) English. The questionnaires were translated into the local language, but the interviewers immediately translated the information to English which introduced issues related to the interpretation of the answers, especially in the open questions. This issue was also anticipated and addressed during the training but still constituted a challenge.
In the case of AIM, a different challenge arose from the focus on the LWC supported branches, introducing a placement bias, which is relevant when thinking of generalising the results to other branches of the institution. It was expected that the selected branches were in general similar to most of the existing AIM branches at the time of the first survey since they were mainly located in urban areas of the province of Punjab. However, the geographical expansion of AIM to different provinces in the following years, including rural and remote areas with average poverty levels higher than in Punjab, alters this analysis. Moreover, even within the branches included in the study, heterogeneity was present in some of the indicators. This situation implies a cautious approach to the transferability of the results.

Finally, it is important to refer the challenges associated with the nature of the LWC evaluation project, and particularly the need to manage the trade-off between a pure accountability approach which is associated with independency and rigour and the focus on improvements of the programme and the institution, which implies a stronger involvement of the partners aiming to increase the usefulness of the project results. The case of AIM illustrates well the complexity and the tensions introduced by this dual objective of the impact project. The priority given by both PIU to external accountability goals, as described above, puts a higher weight into guaranteeing the independency of the evaluation, but the local context and the limitations of resources led to a direct involvement of the MFI in some of the implementation stages, e.g. data entry, which should ideally be externalised.

The example given shows how the field conditions interfere and shape the balance between the two goals, with the learning/improvement perspective prevailing in this instance. This is a similar conclusion to Greeley (2005, p.49) in his analysis of the evaluations implemented within the Imp-Act project.

Different objectives and levels of rigour can be associated with an impact assessment (White, 2009b), which calls for increased transparency in the communication of the project process and results, being crucial in this effort the development of a common language regarding the evaluation.

*Research Question 1.3: Can poverty scorecards such as the Poverty Probability Index (PPI) be a useful evaluation tool for MFIs?*

In the case of Pakistan, most of the main challenges identified in the literature regarding PPI (presented in chapter four) seem to converge and contribute to the negative perception of AIM managers towards the instrument.
These main challenges include theoretical and empirical issues and relate to accuracy over time, sensitivity to within-group inequalities, cost of implementation and property of the tool (Bhat & Yamini, 2012; Boucher, 2014). From these, the cost of implementation, although important, seemed to be secondary to AIM managers compared to the other factors.

Before discussing these challenges in the context of the LWC impact project, it is important to contextualise the application of the tool in the institution referring to another crucial factor for the results obtained, which is highlighted by Boucher (2014), the motivations for using the PPI and, subsequently, the model of implementation.

For AIM, the decision to use the PPI was external, being the application of the tool proposed by LWC as part of the evaluation project. The institution was not using PPI at the time of the project, so this was looked upon as a first application. In the previous sections of the chapter, the model of implementation was presented: the tool was applied to a sample of new clients from specific branches of the institution, with data collection and data analysis being conducted by the project independent interviewers and the evaluator, respectively. In addition, data was also collected for the sample of non-clients.

In the project, this choice of implementation model derived naturally from the decisions made by LWC within the evaluation project, but this is not always as straightforward. Boucher (2014) in his study of the implementation of the tool in five MFIs from Ecuador and Peru found five different models in regard to issues such as sampling (for example, Fodemi applied the tool to all clients while Espoir and Prisma used samples) and data collection, with four of the institutions collecting data through internal resources (normally, the loan officers) and only Prisma using a third party for this effect.

The model applied in AIM allowed to solve the two concerns associated with the external motivation raised by Boucher. One of these concerns refers to the incentives for the institutions to over-state the poverty levels of their clients in order to demonstrate they are targeting poor populations and achieving their social mission. The other is related to the rigour in the implementation, being expected that it will be lower if the application of the tool is imposed from the outside (Boucher, 2014, p.19). By being applied by the independent interviewers recruited for the household survey and being the collected data analysed by me as the evaluator, both questions were solved or strongly mitigated.

The results of the application of the tool at the baseline survey, included in the preliminary report, illustrated the questions associated with accuracy over time and within-group sensitivity (Afonso, 2016). As previously mentioned, the PPI scorecard for Pakistan was outdated. It was based on data from the Pakistan Social and Living Standards Measurement.
Survey (PSLM) 2005/6 (Schreiner, 2010) and was not updated since. In contrast, the scorecard for Ecuador was updated in 2015 (national survey data from 2013/14) and the new version of the scorecard for Peru was published in 2018, using data from 2016 and the improved methodology adopted by IPA (IPA, 2018). Time considerations raise a parallel question associated with the international poverty lines used as reference in the analysis. Older versions of the scorecards employ also the outdated poverty lines 2005 PPP. To minimise the situation, I have used the $2.50 poverty line in the analysis of the AIM data, closer to the currently used poverty lines (2011 PPP).

The accuracy over time is particularly relevant considering that this is recognised by Mark Schreiner as one of the major limitations of the tool (Schreiner, 2014a). Considering that the credibility of the estimates for the poverty levels of the households is questionable, the use of PPI in the evaluation focused on the changes of the scores during the period of the project, and not on the estimation of absolute poverty levels.

The application of PPI in AIM suffered as well from the sensitivity of the tool to strong variations within group. A common situation at this level regards the existence of regional disparities (Boucher, 2014). The fact that PPI is based on national averages potentially results in underestimation of the poverty levels of those respondents living in more economically developed areas, such as the Punjab in Pakistan.

These two main issues are reinforced by questions related to the ownership of the tool as stressed by Bhat and Yamini (2012). The update of the questionnaires and its adaptation to the local contexts is not ‘in the hands’ of the MFIs and it is not an automatic process. It is dependent on the availability of new national survey data and implies costs associated with the re-design and test of the questionnaire. The process is, thus, inherently political as gathering funds and support requires the involvement of different actors in the country’s microfinance sector. In the case of Pakistan, there is updated national data from the PSLM 2013/14, so the question seems to be mainly political and helps contextualising the decision taken by AIM regarding the application of the tool.

AIM will not further employ PPI as a management tool preferring to develop its own poverty measurement tool. In this process, the loan application procedures in place and the lessons learnt during the evaluation project are fundamental references. There are two main arguments put forward by AIM1 for this decision. The first, in line with the accuracy over time limitation, is that “PPI form is something which has been obsolete”, with some of the questions not making sense in the present context, especially if the tool is to be used in targeting decisions. The second is related with the discussion around the property of the tool.
and its implications, with the AIM manager questioning who “prepared the PPI”, “how they get the data” and pointing out that “nobody bothered to edit it” (AIM1, p.c., March 22, 2018). These declarations seem somehow strong given the transparency of the PPI promoters and the related data and materials available online (Boucher, 2014, p.45), but they should be considered taking into account the apparent mutual suspicious regarding the academic community, expressed in the previously referred “they don’t believe” (in the sustainability of the AIM microcredit model) and the observation of LWC1 regarding the difficulties of MFIs communicating with researchers from different cultural backgrounds (LWC1, p.c., June 4, 2018).

From the exposed above, it may be concluded that from the point of view of the AIM managers, PPI is not a useful evaluation tool. Despite this, it should be noted that the variation of the poverty scores was a variable of interest in the analysis and the project reports.

Research Question 1.4: How do different stakeholders perceive and use the evaluation results?

Focusing on the two main PIU, the LWC team and AIM managers, the perceptions and (intentions of) use by these two stakeholders were in line with the expectations of the researcher given the strong motivation and commitment towards the project. In the case of AIM, the overall positive results contributed to a sense of satisfaction, confirmation of merit and legitimisation of the work, which naturally LWC and AIM were looking forward to communicating to their funders and supporters. This was reflected in the different initiatives described in the chapter, from the LWC funders workshop to AIM sharing the final report with the Pakistani regional governments funding the institution.

The positive results had also as consequence the decision by the AIM managers to not introduce in the short-term changes in the microcredit programme – “we have not taken any decision, because the results are so much satisfactory for us right now” (AIM1, p.c., March 22, 2018). Elements of both AIM and LWC teams assigned especial relevance to the third wave of the survey to consolidate the findings and support future decisions regarding the programme and the funding (AIM1, p.c., March 22, 2018; LWC2, p.c., June 4, 2018).

The use given to the findings of the AIM evaluation was, in this way, in line with the primary accountability-driven purpose of the PIU. The absence of immediate concrete actions regarding the programme cannot be considered as surprising, being a common outcome of many programme evaluations. Also recurrent is the recognition of the participation in the process as valuable for those involved and their institutions (Patton, 2008, p. 101).
This recognition has been also evident within the LWC assessment project and has led to the decision by the LWC team of extending the project to other field partners. The experience gained implementing the evaluations in Pakistan and Zimbabwe will be used to replicate the project in other countries, introducing the necessary adjustments to the local contexts. In the interviews, the LWC evaluation team referred explicitly to make improvements related to the control group - “that is probably something we need to address when we extend this to the next countries” (LWC1, p.c., June 4, 2018), with LWC2 verbalising also the wish to further explore the topic of women’s economic empowerment (LWC2, p.c., June 4, 2018).

The LWC evaluation project contributed to a mindset change in the institutions involved, particularly in the Pakistani partner, with LWC1 considering that “our partners are now beginning to take the issue of evaluation research much more seriously” (LWC1, p.c., June 4, 2018). The participation in the project represented for AIM an opportunity to learn, to better understand the processes involved in the evaluation and the advantages and challenges associated with the different stages of the project, as well as to study the possibilities of integrating the evaluation process in their own operations.

Evaluation seem to have become a management priority in AIM, with the lessons learnt during the period of the project being used in the development of the new management and information system (MIS) of the institution. AIM2 gave a concrete example of this use when talking about the potential integration of some of the questions/indicators used in the survey in the new MIS (AIM2, p.c., March 22, 2018).

It should be noted that the answer to this research is partial and incomplete due to the continuation of the project, with an on-going analysis by the PIU of the evaluation results.
7 Case Study: THRIVE Microfinance, Zimbabwe

Chapter seven describes and discusses the evaluation project developed in Zimbabwe with the objective of assessing the changes in the lives of the microcredit clients of THRIVE Microfinance supported by Lendwithcare (LWC). The process replicated the one previously implemented in Pakistan. Therefore, there are similarities between the two settings in terms of the approach to the evaluation and the methodologies applied, which is reflected in a description of the case in a slightly less detailed description of those aspects already included in chapter six. As much as possible, these similarities as well as the main differences are flagged throughout the chapter, paving the way to the cross-case synthesis presented in chapter eight.

The chapter is organised as follows. Section 7.1 gives a brief overview of the external environment in which THRIVE operates, followed in section 7.2 by the presentation of the institution and its microcredit programme. In common with the previous chapter, section 7.3 describes and discusses the evaluation project, using the UFE-based framework of analysis. Finally, section 7.4 summarises the main challenges and lessons learnt from the implementation of the evaluation in Zimbabwe, addressing the first research question of the thesis.

7.1 Macro-context

The Republic of Zimbabwe is a landlocked country located in Southern Africa, with borders with South Africa, Mozambique, Botswana and Zambia. The Inter-Censal Demographic Survey 2017 estimates the population resident in the country at around 13.6 million people, with 68% living in rural areas and approximately 10% residing in the province of Harare (one of 10 provinces in the country). The population is mainly Christian (84%) and ethnically homogeneous, with an insignificant presence (less than 1%) of immigrants (ZIMSTAT, 2018).

The country is a recognised sovereign nation since 1980, subsequently having been ruled by the political party ZANU-PF. From 1987 until his resignation in November 2017, Robert Mugabe was the President of Zimbabwe. The recent history of the country has been dominated by economic instability, aggravated by the 2008-2009 hyperinflation crisis and episodes of extreme climate conditions, such as the drought in 2015. These events have been gradually eroding the wealth of the country, affecting all economic sectors and leading to the development of a significant informal sector composed of poor entrepreneurs managing “survival” businesses (Murisa & Chikweche, 2013, p.7).
The 2010s have been particularly difficult from an economic perspective, with the emergence of the cash crisis. Between 2014 and 2016, the GDP growth rates in Zimbabwe were, respectively, 2.1%, 1.7% and 0.6%; consistently below the equivalent averages for the low-income countries (6.1%, 4.7% and 4.2%). The situation has not improved during the implementation period of the evaluation project (2016 – 2017). The announcement of the introduction of bond notes by the Central Bank of Zimbabwe in May 2016 to deal with the cash crisis resulted in a double standard for the currencies, with bond dollars rapidly depreciating to the US dollar. The situation instigated fears of a new hyperinflationary period (not confirmed so far) and increased the levels of uncertainty towards the future, contributing also to the political developments that resulted in the resignation of Robert Mugabe (Burke & Graham-Harrison, 2017).

Despite the gloomy economic picture associated with the last decade, there are also positive indicators that, beyond the political results of the elections in the summer 2018, bring hope of a better future for the country. One of the most positive factors is the high literacy levels of the population, well above 90% in the whole country and particularly strong in the urban areas (99.3% for Harare in 2014).

Another important factor in this context has been the changes regarding financial inclusion. Looking at the most recent figures included in the Global FINDEX survey (table 7.1), the situation in Zimbabwe seems to have significantly improved with regard to access to formal finance.

Table 7.1 – Zimbabwe Financial Inclusion Data

<table>
<thead>
<tr>
<th></th>
<th>2017</th>
<th>2014</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults with account, including mobile accounts, total</td>
<td>ZIM (1)</td>
<td>LIC (1)</td>
<td>ZIM</td>
</tr>
<tr>
<td>Adults with account, including mobile accounts, female</td>
<td>51.7%</td>
<td>29.9%</td>
<td>29.0%</td>
</tr>
<tr>
<td>Adults with account, including mobile accounts, poorest 40%</td>
<td>43.6%</td>
<td>25.5%</td>
<td>16.3%</td>
</tr>
<tr>
<td>Adults with account at financial institution, total</td>
<td>28.2%</td>
<td>24.5%</td>
<td>17.2%</td>
</tr>
<tr>
<td>Adults with account at financial institution, female</td>
<td>23.3%</td>
<td>20.5%</td>
<td>15.3%</td>
</tr>
<tr>
<td>Adults with account at financial institution, poorest 40%</td>
<td>13.2%</td>
<td>16.4%</td>
<td>7.4%</td>
</tr>
<tr>
<td>Adults borrowing from a financial institution, total</td>
<td>4.0%</td>
<td>7.1%</td>
<td>4.0%</td>
</tr>
<tr>
<td>Adults borrowing from a financial institution, female</td>
<td>3.0%</td>
<td>6.4%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Adults borrowing from a financial institution, poorest 40%</td>
<td>0.9%</td>
<td>5.6%</td>
<td>2.2%</td>
</tr>
</tbody>
</table>


33 Source: World Bank
34 Source: Zimbabwe National Statistics Agency.
After a negative period between 2011 and 2014, with the deterioration of all indicators included in table 7.1 (especially for the poorest segments of the population), the percentage of adults owning a bank account has changed dramatically in more recent years. This evolution has been mainly driven by the development of mobile banking, following the trend observed in other African countries. In Zimbabwe the growth of mobile banking was also boosted by the cash crisis in 2016, which eventually led to financial inclusion statistics significantly better than those verified for the average of the low-income countries.

Interestingly, the analysis is different in regard to formal credit, with the percentage of adults belonging to the poorest 40% of the population who have successfully applied for a formal loan decreasing substantially, to values far from the average for the LIC (as depicted in table 7.1). These figures illustrate the unwillingness of the formal finance institutions, including the MFIs, to serve the informal sector. In 2017, many institutions privileged salary-based lending, with the percentage of productive loans provided by MFIs decreasing from 73% by the end of 2016 to 54% one year later (Reserve Bank Zimbabwe, 2018).

The situation is also reflected in the ‘weaknesses’ of the microfinance sector in terms of outreach and financial sustainability (Murisa & Chikweche, 2013). By the end of 2017, the sector served 323,286 active borrowers, with 38% of these being women (Reserve Bank Zimbabwe, 2018). The market remains underserved, especially in rural areas. This underdevelopment of the sector cannot be dissociated from the economic shocks experienced in the end of the 2000s, when the number of active MFIs decreased from 180 in 2007 to 75 in 2009 (Murisa & Chikweche, 2013, pp.9-10), or the cash crisis in 2016.

“The cash crisis meant the average person on the street was highly illiquid and had minimal disposable income. Most of our borrowers trade on a cash basis and can only thrive if a ready medium of exchange is in the possession of potential buyers. This was far from the case, purchases mainly being done via point of sale machines (which our borrowers could never dream of having) or mobile money which for a while remained unpopular.” (THR1, p.c., May 18, 2018).

The cash crisis brought along uncertainty and an increase of the risks associated with the development of small businesses which translated into a higher risk of loan delinquency. As a consequence, the microcredit market ‘slowed down’, with MFIs avoiding approving riskier loans and entrepreneurs more averse to risk refraining from applying for the loans. At the same time, the requirement of working with mobile money contributed also to an increase of the operational costs of the institution.
Despite this more cautious approach by socially focused MFIs, the period between the two waves of the survey was characterised by an increase of the competition in the sector, with the number of registered MFIs reaching 183 by December 2017, a similar size to the period before the hyperinflation crisis (Reserve Bank of Zimbabwe, 2018). Older actors in the development finance sector, such as MicroKing (now Microcred), became active once again and a few new actors, including the Microloan Foundation, entered the market. Nevertheless, the lack of external investment in the sector has contributed to increasing interest rates, attracting financial providers with a pure commercial approach (Murisa & Chikweche, 2013). This idea was corroborated during the informal conversations maintained during the field visits to Zimbabwe, with the perception of managers and staff of the institution being that the majority of the financial providers in the sector are not socially-driven and charge higher interest rates than THRIVE.

7.2 Micro-context

7.2.1 THRIVE Microfinance

THRIVE Microfinance was created by a British social investor in January 2012 with the mission of “providing training and credit to women excluded from the mainstream financial sector in a manner that is both socially responsible and financially sustainable”. It is a for-profit private company, however, all surpluses are retained in the business “for the benefit of its borrowers”. The MFI is registered as a non-banking financial institution under the supervision of the Reserve Bank of Zimbabwe.

The MFI started working in Willowvale (Harare), offering business loans to female entrepreneurs. Initially the lending methodology was based solely on group lending, but it was soon enlarged to also allow for individual loans. The second branch opened in Chitungwiza (25km south of Harare) in July 2015 and approximately two years later, the third branch was created in Whitecliffe (Harare). In 2017, the MFI diversified its portfolio with the introduction of micro-housing and school fees loans and expanded its geographical coverage to rural areas by opening sub-offices in Marondera (70 km from Harare), Bindura (88km from Harare) and Juru (50km from Harare).

By 31st December 2017, the MFI employed 63 persons and it had provided training and credit for more than 10,000 women. The number of active borrowers amounted to 5,503, with a gross loan portfolio of $1,344,610, of which 90% corresponded to solidarity group loans. The

35 From these MFIs, only five are licensed as deposit-taking institutions.
36 Retrieved from the website of the institution.
average loan size was $1,167 (for the borrowers applying within a group). Individual loans tended to be of higher amounts and were usually provided to existing clients who have been part of a group. 37

During the growth of the institution, the microfinance crowdfunding platforms KIVA and Lendwithcare have been important sources of funding. THRIVE is LWC field partner since August 2015. By 31st October 2017, 553 group loans reaching 1,820 clients had been funded, corresponding to an aggregate lending amount of $850,000. 38

One of the distinctive features of THRIVE, when compared with commercial MFIs, is its human resources policy, which does not provide for payment of performance-based monetary incentives. The loan officers, called group relation officers (GRO) are grouped in teams, which are responsible for a given number of loan applicants. In the training stage of the process, different GROs from the same team can provide the different sessions, with the group being allocated to the GRO who has conducted the field visit to the home/business. From the informal conversations with staff, training seems to be an essential component of their job satisfaction and is perceived as crucial in building a trust relationship with the clients.

Another important element of the human resources policy of the institution is recruitment. The field staff is mainly recruited from a pool of newly graduated students from local Universities who have the opportunity to have a work placement at THRIVE as part of their undergraduate course. These “attachees” (as they are named in THRIVE) work for a period of one year in one of the branches of the institution, receiving a monetary compensation. After this period, and depending on the needs of the institution, they may be invited to integrate the permanent staff of THRIVE.

7.2.2 Microcredit Programme

THRIVE microcredit programme has been inspired by the experience of the Grameen Bank in Bangladesh and similar MFIs. Like the experience described by Murisa and Chikweche, THRIVE programme can be considered an example of “imitative entrepreneurship”, representing an “incremental redesign” of a product which has been successfully implemented in other markets (Murisa and Chikweche, 2013, p. 14).

---

37 Data provided by THRIVE upon request.
38 Data provided by LWC upon request. Beyond the crowdfunding platforms, which provided around 30% of the MFI’s funding (2016), the Zimbabwe Microfinance Wholesale Facility and the MFI shareholders were the main funders.
THRIVE lending model is based on group lending with joint liability of the members and includes compulsory savings and training. Although the institution offers also individual loans, the core of the programme is based on group lending, with individual loans representing a solution mainly to existing clients looking for larger amounts of credit. All clients are women.

The following description of the microcredit programme refers to the conditions at the time of the first field trip to Zimbabwe when the baseline survey was implemented (2016), corresponding to the conditions applied to the sample clients in the evaluation. The applicants for the group loans were required to form a group of 3 to 10 members. Group members attended an initial information meeting and received a subsequent visit from a GRO. If the process was set to advance, the group members were invited to the training course. During the five weeks of the course each member was required to save $10 in a specific bank account at a commercial bank.

The training consisted of six compulsory two-hours sessions focused on group formation and financial training. The MFI charged $1 for each session, which represented the only direct cost associated with the application process. In the final stage of the process, there were two additional free sessions, with the first consisting of a conversation with existing borrowers and the second, attended only by the group leaders, focusing on leadership skills.

The loan appraisal was based on the information collected by the GRO in the field visit and by the different members of the credit team during the training sessions. This implied that the final decision on the loan approval integrated opinions from different persons, not just the GRO conducting the field visits.

Figure 7.1 illustrates the MFI lending process. Once the loan was approved and the contract signed, the group members were expected to meet monthly, before the repayment date. The first meeting was scheduled in advance and took place at THRIVE branches. In the monitoring process, GROS were supported by community representatives, who were former and active borrowers of the MFI invited to collaborate in the dissemination of the programme and the monitoring process, for what they received a small monetary compensation. These were normally experienced borrowers with a good repayment record and a good reputation in the institution.
As competition in the sector increased and repayment rates were negatively impacted by the difficult economic situation, THRIVE managers introduced changes to the training programme in 2017 aiming to curtail the time associated with the loan appraisal process and reduce costs. The training sessions were cut to three, one on group formation and two related to financial training, to be completed in the period of one week. This change allowed for the duration of the loan application to decrease from one month to two weeks (with the objective of further reduce it to a maximum of 10 days). 39

This microfinance plus approach is particularly demanding in terms of the operational costs of the institution, and consequently, its funding. This situation translated into the institution charging relatively high interest rates. By the time of the baseline survey, the nominal interest rate for the majority of the loans was 3%/month (flat rate) for the first loan and 5%/month on declining balance in the following loan cycles. In response to higher default rates and operational costs, the nominal interest rate has increased to 5%/month (flat rate) in 2017, which represents an EIR (Effective/Real Interest Rate), considering the training fee and the compulsory savings, that exceeds 100%/year.

The need to charge higher interest rates has been addressed in informal conversations and meetings with the MFI managers, who considered this situation as worrying and contrary to the interests of the clients. The only mitigation factors in the analysis are the small scale of the loans, their short duration (average of 6 months), and to some extent the fact that THRIVE

39 Information collected during the second field trip in June 2017.
is still cheaper than most of the MFIs and financial services providers operating in the same geographical areas.

Regarding the microcredit programme, it is important to highlight that by the time of the two surveys, THRIVE clients were all women and a significant part of them developed trade-related businesses. One of the most common activities was cross-border trading, buying and selling different products (often food or clothing) in the neighbour countries, including South Africa, Mozambique, Namibia and Zambia. The concentration of TRIVE portfolio on trade clients seems to be reducing with the expansion in 2017 to rural areas outside Harare, which represented by the end of that year 12% of the active clients.

7.3 LWC Evaluation Project

7.3.1 Introduction

The THRIVE evaluation project aims to assess the microcredit programme developed by the institution in the branches funded by LWC, which at the time the project started were Willowvale and Chitungwiza. THRIVE is one of the most recent LWC field partners (2015), but its managers immediately showed interest in participating in the evaluation project. The first formal meeting between the three partners took place in London, in September 2015.

The approach to be followed in the evaluation was similar to the one adopted in Pakistan, with the implementation of a purposely built questionnaire and a poverty scorecard, which were adapted to the local context and the resources available. This adaptation to the local conditions, as described below, introduced some differences in the process and implied distinct strategies to overcome the challenges encountered, even when these seemed identical at first sight (e.g. forming the control group).

The evaluation team was formed by two elements from the LWC team which are responsible for the selection and monitoring of the crowdfunding platform field partners; three elements from THRIVE, two representatives from the management team and an operational member; and the University of Portsmouth researcher. In different phases of the evaluation project, the LWC coordinator and the two PhD supervisors were also involved.

One of the managers and founder of THRIVE lives in London and divides his time between the two countries, while the other manager involved in the project is a Zimbabwean national permanently based in Harare. The degree of involvement of each of the managers varied along the project. Other particularity of the project in Zimbabwe was related to the LWC team and the absence of one of their evaluation team members for a period of 10 months (maternity leave). This member was temporarily replaced, fact that did not affect significantly
the development of the project since the new team member was internally recruited in CARE International UK and had previously worked in LWC, being familiar with its objectives and operations.

Figure 7.2 shows the timeline of the project, identifying the major milestones, including the main steps on the implementation of the household survey.

Source: Own construction

It is important to refer that, like in Pakistan, between the formal meetings of the partners included in the timeline, there was an extensive exchange of emails between the evaluation team members. In the case of THRIVE, there were also Skype calls with the THRIVE team located in Harare and bilateral meetings LWC/UoP in the UK where the evaluation was discussed.

Similar to the Pakistani case study, the following sections were structured using the UFE-based framework of analysis presented in chapter five.

### 7.3.2 Design of the Evaluation

Considering that the evaluation project implemented in Zimbabwe was a replication of the one initiated in AIM one year before, much of the design of the evaluation was pre-defined according to the decisions taken for Pakistan. The project design had necessarily to be adapted to the characteristics of THRIVE, a much smaller and resources restrained institution compared with AIM, and the context which was particular challenging in Zimbabwe during the period between the two waves of the survey, as described above.

#### 7.3.2.1 Primary Intended Users: motivations

The analysis regarding the identification of the primary intended users is very similar to the one presented for Pakistan. A similar project stakeholders map was produced (Figure 7.3),
with the stakeholders identically clustered in three groups linked to each of the project partners and a residual group, which includes other external stakeholders.

Figure 7.3 – LWC/THRIVE Evaluation Stakeholders Map

![Stakeholders Map]

Source: Own construction

As described in the previous chapter, the implementation of the project was driven by LWC’s accountability purposes towards CARE International UK and the existing crowd-funders supporting the platform, with a secondary objective of building capacity in social performance assessment within the partner MFIs. In this sense, the lessons learnt from the evaluation process were expected to be of interest also to the other LWC field partners.

THRIVE managers were enthusiastic from the start regarding the evaluation and the potential participation of the institution in the project. In the final interview, one of the MFI managers refers to two broad expectations from the MFI perspective, which were the main motivations to join the project – gaining a better understanding of the poverty characteristics in the geographical areas where they work, and “reveal the extent to which training and micro-credit were impacting livelihoods as an intervention” (THR1, p.c., May 18, 2018).

Throughout the project, it was noticeable an explicit commitment to the integration of social performance in THRIVE management priorities. The institution had started implementing a number of related measures, including the creation of an exclusively dedicated staff position (social performance officer), the introduction of client satisfaction surveys and the development of an internal poverty scorecard. Recognising the usefulness but also the
shortcomings of this poverty assessment tool, THR1 expressed in his interview how they are looking forward to the development of a “standard PPI that fits the global reporting template” (THR1, p.c., May 18, 2018). To this purpose, THRIVE has been working with other actors in the Zimbabwean microfinance sector to persuade the managers of PPI to develop the scorecard for the country.

The latter citation reflects also the perception of the MFI managers regarding the relevance of gathering standardised and credible evidence of the outcomes of their work. This issue is particularly important for the institution, which is still a young MFI in need to attract new funders in order to increase the amount of available funds and diversify the risks associated with funding.

Similar to the project in Pakistan, three other stakeholders connected to the institution were crucial in the evaluation – the operational member of the evaluation team, the clients and the non-clients participating in the household survey. Although they have influenced the results of the evaluation, their participation was primarily reactive (to the invitation received), since they had no decision-making power over the evaluation project or the microcredit programme. The definition and implementation of strategies to encourage the participation of non-clients in the household survey (further discussed in section 7.4.3) revealed to be very challenging, with the results of the project being affected by these difficulties encountered to set up the control group.

The third partner (major stakeholder) in the project was the University of Portsmouth as detailed previously in the AIM case study. The other groups theoretically interested in the evaluation results are comparable to the ones identified there, including the existing and potential clients (benefiting from eventual improvements in the programme), THRIVE funders, other microfinance institutions operating in the Zimbabwean market, regulators, microfinance consultants and the academic community.

The identified primary intended users (PIU) in the evaluation project were the LWC team and the THRIVE management team.

7.3.2.2 Role of the evaluator

Having as reference the UFE model, I have identified the PIU, encouraged their active involvement in the design of the evaluation and the compliance with recommended ethical procedures. Given the strong motivation of the THRIVE evaluation team regarding the project, my efforts focused on the establishment of working relationships with the Zimbabwean team members and the management of expectations from the partners regarding the achievable results of the project.
In building the working relationship with the two evaluation team members located in Harare, I have used the internal poverty scorecard developed by the institution (THRIVE-PAT) as the subject in the initial contacts, gaining a better understanding about the tool and offering recommendations regarding its application.

In parallel, the relationship between the UK-based THRIVE manager and the LWC team developed along with the process of consolidation of THRIVE as LWC partner and led to the participation of the MFI in other LWC projects. Among these projects was the pilot implementation of social ratings through a partnership of LWC with a qualified rating agency based in Barcelona - Inclusion [Social Ratings]. This project run along the evaluation project, with the social performance report prepared by the agency being publicly shared in November 2017 and presented to all LWC partners in a workshop in March 2018, in London.

As in the case of Pakistan, the management of the expectations regarding the evaluation results implied giving particular attention to the language used in the communications between the partners, emphasizing the importance of being rigorous and following ethical procedures in the application of the methodologies selected.

7.3.2.3 Household Survey: preparation

The household survey included the application of the purposely built questionnaire and the THRIVE-PAT. Considering that the new clients selected to participate in the study had just completed the loan application process and the THRIVE-PAT questions were part of that process, it was decided to not repeat them in the client questionnaire (or separately), using the data previously collected by THRIVE to compute the respective poverty scores.

This decision made possible to include further questions, which had not been considered in Pakistan, without increasing the planned time for the interviews of 20 to 25 minutes. In the case of the non-clients, though, the questions necessary to compute the poverty scores were introduced in the questionnaire (not being asked separately as it happened in Pakistan). The final versions of the client and the control group questionnaires had 23 and 22 questions, respectively (see Annex II).

The selection of the programme outcomes to be evaluated and the redaction of the specific questions to be used in the questionnaires started from the analysis of the recent experience in Pakistan. In the case of THRIVE, there was a closer alignment of the interests of the PIU regarding the outcomes of interest for the institutions, so the main adjustments made to the

---

40 The preparation of the questionnaires for the baseline survey in Zimbabwe concentrated in the initial months of 2016, after the first application of the survey in Pakistan had been completed and the data analysed and shared with the respective evaluation team (September 2015).
questionnaires resulted from the adaptation to the local context (e.g. the Pakistani ‘committee’ was replaced by ‘ROSCA’ in the answer options to the question on savings mechanisms, and in the potential justifications to applying for the loan at the institution, the religious-based motivation was excluded in Zimbabwe).

As in Pakistan, the questionnaires included demographic data on the interviewee and her household, and both closed and open questions; the latter aiming to capture the perception of clients and non-clients in relation to the expected (baseline) and the observed (second questionnaire) changes in their lives during the period of the study.

In Zimbabwe, there was no need to translate the questionnaires to the local language since English is one of the official languages of the country and it is generally understood by THRIVE clients. During the application of the questionnaires, the interviewers used both Shona (most commonly spoken local language in the province of Harare) and English to guarantee that the clients fully understood the questions, but all answers were written down in English. This is a procedure similar to the one followed by the institution in their training sessions, in which the loan officers speak mainly in Shona but integrate English into their discourse, with all the supporting materials being produced in English.

The goals for the number of interviews, both with clients and non-clients, were defined taking into account the number of active clients at the time of the baseline survey (2,520 by 31 December 2015) and the expected difficulties to locate, and persuade to participate, a group of non-clients with similar characteristics to the new THRIVE clients. These goals were set at 350 interviews for the client sample and 150 for the control group. As in the case of AIM, all new clients who had their loans approved and were about to receive the money were invited to participate in the study, and the interviews were conducted until achieved the established objectives. No other criteria were considered.

The recruitment process was straightforward in the case of the clients, who in most cases were first contacted in the last session of training, when their loan appraisal process was close to be concluded. At this time, the ‘quasi-new clients’ were given generic information about the research project and the survey and they were invited to participate (after the signature of the loan contract).

The process was more challenging in the case of the non-clients. The objective was to recruit women entrepreneurs with similar characteristics to the THRIVE clients, on average, but who had not applied for a business loan at the institution. The complexity associated with the task led the partners to discuss and decide the strategies to adopt during the first field visit to Harare.
The trip took place between 14 and 23 March 2016 and coincided with the first evaluation of the field partner by the LWC team. The programme of the visit included the training of the interviewers, a pilot test of the client questionnaire, visits to the Chitungwiza branch and informal contacts with clients and staff of the institution. Unlike Pakistan, in Zimbabwe the first field trip was preparatory of the baseline survey, with the application and validation of a small number of questionnaires (five). This situation resulted mainly from the timings of the visit, which were set by LWC considering the partner evaluation process.

The recruitment of local interviewers was supported by LWC following the proposal of THRIVE to invite four of the ‘attachees’ that had recently finished their work placement in the institution and were about to conclude their undergraduate course. These students were independent from the institution at the time of the survey, but they were familiar with microfinance and THRIVE microcredit programme, which could contribute to a higher quality of the interviews. Considering that the clients to be interviewed were all new clients, there was a minimum probability of them previously knowing the interviewers, who were instructed to present themselves as University students recruited by LWC for the project, independent from the MFI.

By the time of the first visit to Zimbabwe, the selected interviewers were still in their last period of exams at the respective universities, with two of them studying at the University of Bulawayo, around 435km from Harare. This implied that the training of the interviewers had to be scheduled to the weekend (Sunday) and that only the two interviewers located in Harare managed to conduct five pilot interviews in the following days. These were validated by me and led to small adjustments in the wording of the questions. Complementing the training session, I have prepared a questionnaire guide for the interviewers, addressing the issues thought to be relevant for each question and the interview process in general. The effective start of the baseline survey happened two weeks after (April 2016).

The field trip was also the moment to coach the interviewers and the local evaluation team on the recommended ethical procedures related to the interviews. Similar to Pakistan, the clients had already given their consent to participate in research projects coordinated by LWC. The concern was, therefore, to ensure that the specific objectives and conditions of the project were communicated and accepted by the interviewees and that they were made aware of other ethical procedures. These issues were generally explained to the potential respondents in the presentations of the project at the training sessions, and they were reinforced in the beginning of the interviews. Finally, within the questionnaire, a question was included regarding their consent to be contacted in the future regarding the survey.
7.3.3 Implementation Process

7.3.3.1 Primary Intended Users: behaviours

The LWC team was actively involved in all stages of the process. However, unlike the AIM project in which the LWC member with Pakistani nationality was present in all meetings and all events related to the evaluation, the participation of the two members of the evaluation team (and the temporary member) was more distributed in the case of Zimbabwe. There was also a different involvement of the LWC team in one of the stages of implementation – the computerisation of the data – which was conducted in London under direct supervision of the LWC team.

THRIVE managers, similar to what have happened in AIM, were keen to show their permanent commitment towards the project. There was a stronger involvement of the institution, and particularly of the operational member of the evaluation team, during the field implementation of the two waves of the survey. During the LWC/UoP field trips to train the interviewers and validate the questionnaire first applications, THR2 estimates to have dedicated 75% of her working time to the project. During the survey implementation, the time requirement, which was mainly associated with the quality check of the questionnaires, was of 60% in the first week of interviews, decreasing to 30-40% in the following weeks as the interviewers became more familiarised with the questions and the process. The participation of the THRIVE operational member in each wave of the survey ended after sending the questionnaires to the UK (THR2, p.c., April 23, 2018).

Also similar to the experience in Pakistan was the attitude from THRIVE managers regarding the provision of feedback on the reports produced, with the comments obtained resulting mainly from direct interpellations at the partner meetings. This situation seems to indicate higher effectiveness of verbal communication, which is one of the main lessons learnt from the implementation process and is further discussed in chapter eight in the context of the continuation of the LWC evaluation process and its extension to other partners.

7.3.3.2 Role of the evaluator

My role as evaluator in the implementation stage was especially active in two stages: the training of the interviewers and validation of the first interviews during the field trips, and in the analysis of the data collected in the two surveys. During the implementation of the surveys, I have received information from the field mainly from the communication with THR2. Likewise, my participation in the computerisation of data was indirect, providing guidelines on the process and recommending the fields to be included in the database.
In the first survey, data entry was outsourced by LWC to a data agency. This decision accelerated the process but ended up not being the best option for the project. The quality checks conducted in the field by THR2 focused on the completeness of the questionnaires and the identification of errors, but not as much on the standardisation of the answers. As an example, in the question related to the average monthly expense on education, some of the interviewers introduced the school fees for the term, and explicitly mentioned this reference period. In this case, previously to the introduction in the database, the amount needed to be transformed into the monthly equivalent, which was not taken into account by the data company. Given that different persons introduced the data and they were not familiar with microfinance or the objectives of the project, the indications given by the LWC team were insufficient to avoid inconsistencies.

The lack of experience of all parties involved in this stage helps explain the situation, which was corrected in the second wave of the survey with a different solution used for the introduction of data (presented below). In the first survey, I have validated and corrected the information in the database, checking all the questionnaires for the issues identified as problematic. This procedure implied extra time dedicated to the process and it preceded the analysis of the data and the production of the first report based on the baseline results.

In the case of Zimbabwe, there were fewer opportunities to develop parallel activities involving the three partners as the ones described for AIM in the previous chapter. I have worked to maintain the communication between the partners throughout the project, especially in the periods between surveys. Besides promoting the involvement of the PIU, this strategy allowed me to gain insights regarding the economic and political situation in the country and its effects on the institution and the microcredit clients.

7.3.3.3 Household Survey: implementation and analysis

The two waves of the survey were implemented in the periods April-June 2016 and June-August 2017. The period between surveys was approximately one year; less than the equivalent period in Pakistan due to the different loan cycles of the two microcredit programmes. THRIVE microcredit loans have an average duration of six months, which implied that a part of the clients (29%) were already starting a third loan cycle by the time of the second interview.41

It should be noted that the interviews with the clients were scheduled to happen at THRIVE offices since it was deemed unfeasible to conduct them (as desired) at the houses/businesses

---

41 35% of the clients had finished repaying the second loan and 36% completed the first loan cycle and had not applied for further loans by the time of the interviews (Afonso, 2017).
of the clients. This important logistical constraint, with potential implications in terms of response bias, resulted of the combination of two factors: the geographical dispersion of the residences, a part of them situated in new and non-planned areas in the periphery of Harare, and the time and financial constraints associated with the project in Zimbabwe. To compensate for the time and the cost of transport, the clients participating in the interviews were offered a small monetary compensation.

The baseline questionnaire was applied to a sample of 341 clients and 157 non-clients. The results were compiled into a preliminary report similar to the one produced in Pakistan, focusing on the characterisation of the clients and non-clients and the identification of significant differences between sub-groups of clients. The report was prepared as a working document circulated between the partners and included recommendations regarding the implementation process (Afonso, 2017).

The initial analysis of different groups of clients focused on the potential differences between the clients of the two branches and between cross-border traders and those with other businesses. Cross-border traders represented 23% of the initial client sample and it was expected that given that the nature of their work was different compared with other borrowers, there would be differences between the personal characteristics of the two groups. This was not confirmed by the statistical tests implemented, with the few statistically significant differences found reflecting the inherently different working conditions (namely, number of working hours and working alone/with employees).

There were, however, statistically significant differences in many of the studied variables when comparing clients and non-clients, which illustrates the difficulties felt in the constitution of the control group. In the first field trip, the members of the evaluation team agreed on three possible strategies to locate the potential respondents: invite the participants at THRIVE information meetings who did not proceed with the loan application; ask the community representatives to signal non-clients in their neighbourhoods with similar characteristics to the clients, who could be then contacted by the interviewers; and direct contact by the interviewers in the same areas.

These strategies were, however, not effective when implemented according to the THRIVE team who proposed, as alternative, to focus the attention of the interviewers in market/shopping areas where it would be easier to identify and contact potential

---

42 The client sample is slightly inferior to the goal set (350) as nine of the received client questionnaires from Zimbabwe were excluded by the evaluator. Eight of these cases corresponded to data duplication, with the same clients being interviewed by two interviewers in different moments of time.
respondents. Similar to the clients, the non-clients participating in the interviews received a small monetary payment to compensate them for the opportunity cost associated with the time spent in the interview.

The clear differences between the two groups were not totally anticipated and they were only identified by me when analysing the dataset received from LWC. Clients were, on average, significantly older than non-clients (40 compared to 34 years old) and lived in larger households (4.7 members vs 4.2). There were significant differences regarding house ownership, with 29% of the clients living in their own house and 31% renting, while the majority of the non-clients lived in rented houses (72%), and only 7% owned a house. Clients were also significantly more likely to have made home improvements (31% vs 10%) and to have bought fixed assets (56% vs 45%) during the previous 12 months (Afonso, 2017).

With regard to economic activity, clients were, on average, more experienced in their businesses (7 vs 5.5 years) and more likely to employ others (48% vs 24%). They were also significantly more positive regarding the performance of their businesses, with 42% declaring their business sales to be growing during the previous year, compared with only 26% reporting decreasing sales. Conversely, 45% of the non-clients declared their businesses to be decreasing and, inversely, only 26% affirmed that their businesses were growing. Moreover, the comparison of the averages for the two groups THRIVE-PAT scores and income indicators indicated that the non-clients were more likely to be poor than the clients (Afonso, 2017).

The subsequent analysis of these results led to the conclusion that these differences resulted partially from the concentration of the effort to interview non-clients in shopping areas, and thus, on trade entrepreneurs.\footnote{Looking at the sub-samples of clients and non-clients interviewed in the two waves of the survey, 86% of the non-client declared trade as their main activity, while this figure was only 62% for the clients (including cross-border traders in both cases) (Afonso, 2017).} Another factor that may have contributed to these results was the small token given to the respondents, which was, most likely, relatively more significant to the poorer entrepreneurs located in these markets.

In March 2017, in a meeting between the partners in London, THRIVE manager reinforced their commitment towards the evaluation, despite the negative consequences for the institution and its clients of the deterioration of the economic situation in Zimbabwe. The second wave of the questionnaire in the summer 2017 was implemented with the recruitment of a new batch of four interviewers using the same strategy of the baseline survey. As they were ‘attachees’ of the MFI when the sample clients were already involved
with the institution, there was a special attention in the distribution of the interviews to avoid that the clients associated the interviewers with THRIVE.

In this second wave of the survey, the field visit took place after the interviewers finished their final exams period which allowed for a longer training period and the validation of a higher number of interviews by the evaluator and the LWC team member.

The second questionnaire included some questions that were identical to the first questionnaire, new queries about the changes in the selected outcomes, and two open questions aiming to capture the perception of the interviewees regarding any observed changes. In order to reduce response bias, the interviewers introduced these questions without hinting at any particular area or cause of change, elaborating on the question only if the interviewee was not answering. As mentioned before, also included in the client and non-client questionnaires were the questions required to compute THRIVE-PAT scores. The questionnaire forms are available in Annexes II and III.

Table 7.2 shows the sample distribution in the two surveys.

<table>
<thead>
<tr>
<th>Branch</th>
<th>Clients 2015</th>
<th>Clients 2017</th>
<th>Non-clients 2015</th>
<th>Non-clients 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Willowvale</td>
<td>180</td>
<td>133</td>
<td>117</td>
<td>89</td>
</tr>
<tr>
<td>Chitungwiza</td>
<td>161</td>
<td>112</td>
<td>40</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>341</td>
<td>245</td>
<td>157</td>
<td>110</td>
</tr>
</tbody>
</table>

Source: Own construction

In the second questionnaire, 245 clients and 110 non-clients participated in the interviews. Considering the application of two waves of the survey, attrition was predicted, and expected to be particularly significant in Zimbabwe given the economic conditions during the period. The numbers in the table above corroborate this expectation, with attrition reducing the sample size by 29% compared with its initial size. Interestingly, the problem similarly affected both clients and non-clients, contrarily to what have been experienced in Pakistan where the problems of attrition affected mainly the control group. In the case of the clients, attrition affected mainly those clients who were in default and “would not avail themselves for the interviews” (THR2, p.c., April 23, 2018).

The presence of all the interviewers and evaluation teams at the start of the second survey, along with the experience gained during the application of the first questionnaire, contributed to a smoother process in the implementation of the second questionnaire, which correspondingly improved the quality of the data. There was a lower number of missing
values and measurement errors in the questionnaires (compared with the first survey), and the open questions provided valuable descriptions of the respondents’ perceptions of the changes in their lives.

In the second questionnaire, as mentioned above, LWC assumed the computerisation of the data through the work of a volunteer of the institution and the direct supervision of one of the evaluation team members. The process of data introduction was slower, dependent on the availability of the volunteer, but it was more consistent and trustworthy. The resulting dataset was analysed using a difference-in-difference method and quantile regression (to explore the heterogeneity within the client sample). I have subsequently prepared a comprehensive report of the findings, bearing in mind the objectives of the PIU, especially LWC. As previously referred, although the report was to be shared and discussed with the partners, it was prepared as a public report, i.e. potentially shareable with third parties.

7.3.4 Interpretation and Use of Findings

7.3.4.1 Primary Intended Users: perceptions and intentions

The mixed results obtained after the second wave of the survey in Zimbabwe were regarded by the LWC evaluation team with some caution. The findings (see Tables 7.3 and 7.4 below) were considered as “not as positive” or “despite everything, the results in Zimbabwe are not negative” (LWC1, p.c., June 4, 2018). LWC2 attributed an increased importance to the next round of the survey for the team to be able to confirm (or not) the initial results and be in better conditions to take decisions regarding the participation of LWC as investor in the microfinance institution (LWC2, p.c., June 4, 2018).

In the case of THRIVE, the delay in the implementation of the second wave of the survey as well as the relatively longer time taken to computerise the data collected resulted in the discussion of the results and production of reports to be closer to the deadline for submission of the PhD thesis. The discussion of the use of the evaluation findings was, therefore, based mainly on THRIVE managers planned actions and intentions, both in the short and long-term, as they were expressed in the final partners meeting (February 2018) and the written interview (May 2018).

The reaction to the results of the evaluation by THRIVE managers was of some disappointment – “we are disappointed with the overall result, but we are not surprised” considering the economic context and the increased difficulties observed in the contact with the MFI clients (THR1, p.c., May 18, 2018). Notwithstanding, the results were not what the THRIVE managers had expected when they joined the project. In informal conversations with
the two managers during the first field trip, they have expressed the expectation of increased business income (even if in small scale) and increased savings for the clients after one-year of participation in the programme. They were not expecting significant changes in poverty levels and other social outcomes given the short period between the surveys.

The evaluation findings show the accomplishment of the second goal, with the proportion of clients declaring saving regularly increasing from 53% at baseline to 80% in the second round of the questionnaire, while the average proportion of non-clients saving regularly did not change substantially in the period (53% at baseline, 59% in the second questionnaire. Differently, the changes in the income and expenses indicators were negative (see Table 7.2 below), with the average personal income from businesses decreasing 16% in the period for the sample clients. Non-clients have experienced a larger decrease (28%) on their personal business income (Afonso, 2018c; Afonso, 2018d). The expectation of positive changes in the clients’ business income was, thus, frustrated, which contributes to the strong motivation of the THRIVE managers to continue the project and implement a third round of the household survey, now after a more stable period in the country.

Considering the differences identified in the baseline survey between the client and non-client groups, there was not much emphasis on the comparison of the two groups results in the first version of the final report shared with the partners. However, during the partners meeting in Portsmouth, THRIVE manager expressed an interest in knowing more about the non-clients group. The findings at baseline showed that the non-clients were, on average, poorer than the clients, representing a segment of population that THRIVE would like to serve under its social mission, but apparently was not reaching.

As a final note on the use of the evaluation findings, it should be noted that the reports produced within the LWC evaluation project, as well as the report on the social performance of the institution prepared within the social ratings project mentioned above, were perceived in THRIVE as different components of the social strategy of the institution, complementing themselves and allowing the managers to have a deeper understanding of the work and results of the institution. In this way, these pieces of information were helpful both to improve the microcredit programme and to gather evidence that can be provided to present and future funders.

7.3.4.2 Role of the evaluator

Similar to Pakistan, the interpretation of the findings of the two waves of the survey and the preparation of the final report were the most time-consuming tasks for me during the project. The data analysis and interpretation of the results were carried out independently
of the opinions and comments of the other partners, with a working version of the final report being shared with LWC and THRIVE in December 2017. The presence of one of THRIVE managers in the UK made it easier to hold one last meeting in February 2018 to discuss the results of the evaluation and the continuity of the project. As in other occasions during the project, the opportunity to meet and talk through the questions facilitated the clarification of doubts and the understanding of the specific interests of the PIU at this stage of the project. This was particularly important with THRIVE given the mixed results obtained in the evaluation.

Following this meeting (as in the Pakistan case study), I have responded to the feedback provided by the MFI manager preparing a summary report which included the main results for both clients and non-clients (Afonso, 2018d). In the meeting, and in this short report, I have given special attention to the pre-project differences between the groups and their consequences in the analysis and results. The report was in this way an opportunity to highlight the implications of the methodologies applied and the conditions of implementation of the project.

7.3.4.3 Household Survey: results and reports

The findings of the second survey in Zimbabwe were considered mixed, in the sense that they reflected both positive and negative changes in the lives of the clients of the microcredit programme during the period studied. Tables 7.3 summarises the variation on the main income and expenses indicators for clients and non-clients.

<table>
<thead>
<tr>
<th>Table 7.3 – Average Monthly Income and Expenses (Survey)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Monthly Income (Business)</td>
</tr>
<tr>
<td>Clients</td>
</tr>
<tr>
<td>2016</td>
</tr>
<tr>
<td>2017</td>
</tr>
<tr>
<td>Variation (%)</td>
</tr>
</tbody>
</table>

Source: Own construction from LWC impact project data
Note: Current values were used as inflation was not significant in the period of the survey.

The results were more positive in regard to the average changes of PAT with improvements of the average scores both for clients (from 23.5 increased to 24.2) and non-clients (from 14.1 to 20.9). These average results, however, conceal a diversity of situations. In the case of the clients, for 50% of the sample the scores did increase, indicating a lower probability of their households to be considered poor, but there were also 42% who experienced a deterioration of the scores (Afonso, 2018c).
This diversity of situations was explored through the application of quantile regression, dividing the total sample in four equal quartiles with the variation of the personal income as dependent variable. The methodology was then replicated using the variation of the PAT score. Table 7.4 summarises the results for both applications.

Table 7.4 – Quantile Regression Results - THRIVE

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>Type</th>
<th>Coefficient</th>
<th>Coefficient</th>
<th>Coefficient</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Quartile 1 (0.25)</td>
<td>Quartile 2 (0.50)</td>
<td>Quartile 3 (0.75)</td>
<td>Quartile 1 (0.25)</td>
</tr>
<tr>
<td>Age_LN</td>
<td>Q</td>
<td>0.113</td>
<td>0.386 **</td>
<td>0.722 ***</td>
<td>0.261 ***</td>
</tr>
<tr>
<td>Loan (Client)</td>
<td>D</td>
<td>0.569 ***</td>
<td>0.328 *</td>
<td>0.112</td>
<td>-0.001</td>
</tr>
<tr>
<td>Willowvale</td>
<td>D</td>
<td>0.018</td>
<td>0.014</td>
<td>0.049</td>
<td>0.061</td>
</tr>
<tr>
<td>Married</td>
<td>D</td>
<td>-0.071</td>
<td>0.011</td>
<td>0.006</td>
<td>0.081</td>
</tr>
<tr>
<td>Widow</td>
<td>D</td>
<td>-0.189</td>
<td>-0.191</td>
<td>-0.186</td>
<td>-0.004</td>
</tr>
<tr>
<td>Primary School</td>
<td>D</td>
<td>0.036</td>
<td>-0.024</td>
<td>-0.444</td>
<td>0.047</td>
</tr>
<tr>
<td>Secondary School</td>
<td>D</td>
<td>0.061</td>
<td>0.090</td>
<td>-0.246</td>
<td>0.153 **</td>
</tr>
<tr>
<td>Rented House</td>
<td>D</td>
<td>0.059</td>
<td>0.026</td>
<td>0.119</td>
<td>0.141 **</td>
</tr>
<tr>
<td>Own House</td>
<td>D</td>
<td>0.039</td>
<td>0.065</td>
<td>0.208</td>
<td>0.164 ***</td>
</tr>
<tr>
<td>Business Time_LN</td>
<td>Q</td>
<td>0.113 ***</td>
<td>0.155 ***</td>
<td>0.051</td>
<td>-0.019</td>
</tr>
<tr>
<td>Working Hours/Week_LN</td>
<td>Q</td>
<td>0.195</td>
<td>0.207 *</td>
<td>0.199 *</td>
<td>0.146 ***</td>
</tr>
<tr>
<td>Work Alone</td>
<td>D</td>
<td>-0.046</td>
<td>-0.187</td>
<td>-0.170</td>
<td>0.054</td>
</tr>
<tr>
<td>Shocks</td>
<td>D</td>
<td>-0.048</td>
<td>-0.087</td>
<td>-0.242 **</td>
<td>0.017</td>
</tr>
<tr>
<td>Saving Regularly</td>
<td>D</td>
<td>0.210</td>
<td>0.200</td>
<td>0.218</td>
<td>0.121 **</td>
</tr>
<tr>
<td>PAT_2016_LN</td>
<td>Q</td>
<td>0.438 ***</td>
<td>0.336 **</td>
<td>0.319 *</td>
<td>-0.658 **</td>
</tr>
<tr>
<td>PI Business 2016_LN</td>
<td>Q</td>
<td>-0.692 ***</td>
<td>-0.730 ***</td>
<td>-0.756 ***</td>
<td>0.037</td>
</tr>
</tbody>
</table>

Source: Adapted from Afonso (2018c)

Notes: Sample – 355; Observations included: 273; Q – quantitative variable, D – dummy variable (1,0) *** significance level of 1%, ** significance level of 5%, * significance level of 10%

The description of the independent variables is included in the evaluation report included in Annex IV.
Not being the objective of the thesis to explore in detail these results, it is important to highlight that the quantile regression results indicated that the microcredit loan was not statistically significant for the variation of the PAT scores for all the client segments during the period. The analysis pointed, however, to a positive contribution of the loan in the variation of the personal income from the business for those clients who experienced a larger decrease of the business income during the period (quartile 1), meaning that the loan might have functioned as a cushion to these clients compared with non-clients.

The overall negative results in terms of income, associated with other results included in the report such as the decrease of the number of clients buying assets or doing home improvements compared with the previous period, led THR1 to conclude that “the macro-economic difficulties during the test period set a very low upper limit of what any intervention could do to improve livelihoods” (THR1, p.c., May 18, 2018).

But not everything was negative. In the project reports, it can be verified that two of the distinctive features of THRIVE, the training and the savings, seemed to be appreciated by the clients. Even considering that there was some response bias (resulting from the location of the interviews and the clients’ perspectives of applying for future loans), there were different indicators pointing out to this conclusion. Training was considered useful or very useful by 99% of the sample clients and it was mentioned in different ways by clients in the open questions. Equally, albeit the difficulties of the period, the savings frequency seems to have changed substantially, with an increase from 53% to 80% of the clients declaring to save regularly (Afonso, 2018d; see Annex IV).

Another interesting finding was related to the activity mobility of the clients, with 28.5% changing their main business during the period between surveys. These corresponded to situations of change to a different activity or, most frequently, the initial activity being complemented with other that became more profitable. This mobility between businesses was similarly verified in the non-clients group and was identified by one of THRIVE managers as a common strategy to deal with the uncertainty of the period. If anything, the evaluation results seem to demonstrate the resilience of THRIVE clients and their positive attitude, despite all, with 68% of the clients considering their life to be better compared with the time of the first round of the survey. Again, not neglecting the existence of response bias, the figure was quite different (29%) for the non-clients (Afonso, 2018c; Afonso, 2018d; see Annex IV).

Although the results in Zimbabwe were strongly affected by contextual factors, there were still a number of findings that raise important questions about the programme and call for
further analysis. For example, the analysis of the impact and the cost-benefit relation of training was not possible to develop within the quantitative approach adopted in the LWC impact project. Therefore, the use of qualitative methodologies such as clients focus groups, was being studied in the scope of the continuation of the project.

7.4 Discussion

As in the previous chapter, the description of the evaluation process implemented in Zimbabwe within the framework of the UFE model provided the basis to address the first research question (and sub-questions). In each question, and whenever justifiable, a parallel with the discussion developed for the Pakistani case was established.

Research Question 1.1: What are the motivations of different stakeholders regarding the design and implementation of an evaluation process?

The motivations for the LWC team to initiate and develop the evaluation project were discussed in the first case study (see section 6.5). The focus will therefore be on the perspective of the MFI. In the presentation of the LWC evaluation project in chapter five, it was asserted that there was not a previous discussion between the partners explicitly defining the specific objectives of the evaluation. Underpinning all discussions was, however, a common interest in assessing the changes experienced by the clients of the microcredit programmes after the first loan cycle, and to do so using credible methodologies.

The initial expectations declared by the MFI manager in the written interview show that, in the case of THRIVE, there was a dual approach to the evaluation in the sense that there was a specific interest in documenting the changes associated with the participation in the programme (confirmatory approach) – “reveal the extent to which training and micro-credit were impacting livelihoods as an intervention”, but also an exploratory approach expressed by the wish to gain insights regarding the characterisation of poverty in the geographical areas covered by the institution – “understand what characterises poverty in the context of its operating catchment area and broadly Zimbabwe” (THR1, p.c., May 18, 2018).

This latter motivation represented one of the differences relatively to the Pakistani case; the learning perspective, which was present from the beginning of the project. Throughout the participation in the evaluation, THRIVE managers were expecting to build internal capacity in what refers to the evaluation methodologies and processes, and they were aiming also to improve the knowledge about the context in which the institution works, especially in regard to the characteristics and needs of the poor populations.
The difference in motivations can, at least partially, be explained by the context and cultural differences between the institutions, but most likely it also resulted from the different levels of maturity. AIM had been implementing its microcredit programme in the Lahore area for more than 15 years, whereas THRIVE had started operations only 3 years before. Furthermore, the AIM board and management team were composed of experienced and influential locals, while THRIVE was founded by foreign social investors who were supported by a young local team.

The learning perspective regarding the poverty local context led to an increasing interest of THRIVE managers in the control group, which resulted in special attention being given to this group during the implementation of the second wave of the survey to ensure their participation and in the interpretation of the evaluation results.

The focus on non-clients had an additional reason linked to the broader strategy of the institution concerning the measurement and management of social performance. Both THRIVE managers involved in the evaluation reiterated several times, in informal conversations and partner meetings, that THRIVE has a social mission and, thus, the social performance of the institution matters. It has to be measured and its management should lead to improvements benefiting the clients of the institution. This improvement approach fits with one of the purposes of evaluation put forward by Patton (2008, p. 139), the formative improvement and learning, and it could also be associated with a CIPP evaluation in its intent of identifying improvement areas and recommended actions (Stufflebeam & Coryn, 2014, p. 336).

The emphasis given in this discussion to the formative motivations of THRIVE managers to participate in the evaluation should not be seen as an understatement of the also present accountability purpose. The latter was undeniably important in the participation decision, although being less vocalised than in the case of AIM, which can be understood in the context of the gloomier perspectives in Zimbabwe regarding the evaluation results. The cautious perspective of the MFI managers was strengthened by the preliminary report, based on the baseline survey, in which I have identified methodological limitations associated with the location of the client interviews and the composition of the control group. Nonetheless, the accountability purpose was still important and contributed to the strong desire to continue the project and implement the third wave of the household survey, after what THRIVE managers expected to be a more stable economical period in the country (as it was communicated to the other partners in different occasions).
Research Question 1.2: What are the main challenges in the field implementation of an evaluation in microfinance?

The implementation of the evaluation in Zimbabwe presented several challenges, some of them similar to those encountered in Pakistan, others of a completely different nature. There were questions common to both evaluations associated with LWC and its leader role, namely the time management of the project.

In the case of Zimbabwe, the adjustment of the project timings to the time restrictions of the LWC team had practical implications in the implementation of the baseline survey. The training of the interviewers was shortened, and the evaluator could not validate the initial application of the questionnaires by the interviewers as intended (apart from a small pilot with five clients). As a result, the quality control of the questionnaires fell mainly on the shoulders of THRIVE operational team member (THR2), with the long-distance support of the evaluator.

This situation was particularly demanding and reflected in the quality control of the survey. The quality check of the questionnaires focused on the completeness of the answers and identification of errors, not much attention being given to the standardisation of criteria in the answers, such as in the example of the expense in education described above. This happened in part because by the time of the field trip and largely during the interviews period, the decision about who would be responsible for the data entry had not been made by the LWC team. Unlike Pakistan, where AIM assumed the task, THRIVE did not have the resources to do so, and LWC ultimately decided to use an external data agency.

This factor was initially not taken into account and revealed to be a problem only in the data entry phase. The use of a data agency had the advantage of externalising the process, increasing the credibility of the evaluation findings, but the unfamiliarity of the different staff of the agency with the project and the specific survey led to inconsistencies in the dataset, which had to be addressed during the data cleaning process.

The data cleaning process was, thus, more complex and longer than initially expected, especially comparing with the AIM evaluation. Data cleaning complications are not a unique problem, although it is rarely mentioned in the literature. One notable exception is Bolnick and Nelson (1990) who referred to delays in the project and reduction of the sample size resulting from the data cleaning process.

It should be noted that, as described in the previous sections of the chapter, the quality control, data entry and data cleaning processes were reviewed and substantially improved
in the second round of the survey, contributing to a significant improvement of the quality of data.

Another important challenge in Zimbabwe, which did not occur in Pakistan, was the location of the interviews with the clients. My recommendation to conduct the interviews at the house or business of the client, where the clients should feel more comfortable and the validation of part of the information could be done immediately by observation, was not followed for the reasons presented in section 7.3.3. An alternative solution would have been to use community or religious spaces in the different areas where the clients lived, but in the end the decision taken by LWC and THRIVE was to conduct the interviews with borrowers in the MFI offices in Willowvale and Chitungwiza. This option naturally increased the possibilities of response bias.

Comparable to the AiM evaluation, the constitution of the control group and non-client attrition were also an issue in Zimbabwe, although presenting different challenges. In the THRIVE evaluation, the attrition problem was relatively less important than in Pakistan in regard to the non-clients, but in return, there were more difficulties to re-interview those clients in default, which given the circumstances corresponded to a significant number. In many cases, these clients did not answer the phone when the interviewers tried to schedule the second interview or were not willing to go to Thrive offices. In a small number of these cases, the interviews were conducted over the phone.

The constitution of the control group, as previously described, was a complicated task in which the timings and distance between the evaluation team members, as well as the lack of experience in similar situations, did not help. As a result, the data collected from the group of non-clients, still providing useful information for the MFI as discussed in the chapter, introduced limitations in the comparative analysis of the two groups. This was a different problem from Pakistan, where the observable characteristics of the two groups were similar at baseline.

The control group question was discussed on several occasions between the partners with the identification of different strategies to deal with this situation. However, cost and time considerations, as well as the expectations in terms of expected level of rigour of the evaluations led to a decision by the LWC team to interview in the second round the same group of non-clients as planned initially. The observed differences between the two groups were to be accommodated in the analysis and interpretation of the findings. This decision from the project leader as well as the lack of interest of THRIVE managers in regard to the development of more sophisticated analysis techniques of the survey data (expressed
verbally in the project meeting in February 2018) led me to not develop further the analysis within the PhD timeline, focusing on the preparation of the two reports.

In the final interviews, THR2 suggested that the utilisation of the same group of interviewers in the two rounds would have helped in the location of the homes/businesses of the non-clients in the second questionnaire and the interviewers would have benefited from the accumulated experience with the respondents (THR2, p.c., April 23, 2018). This option was, however, not considered as the first batch of interviewers became part of the permanent staff of the institution, maintaining a relation as GRO with some of the interviewees. If adopted, it would have implied an opportunity cost associated with the time dedicated to the survey and a ‘total’ internalisation of the evaluation project, which was not in the interest of the partners considering the objectives in this first phase of the evaluation project.

Research Question 1.3: Can poverty scorecards such as the Poverty Probability Index (PPI) be a useful evaluation tool for MFIs?

The response to this question was clearly positive if the opinion of THRIVE managers in this regard is considered, and contrary to the position assumed by AIM managers presented in the previous chapter. There are a number of reasons justifying the openness of THRIVE to the international microfinance community and the desire of its managers to comply with “global reporting standards” (THR1, p.c., May 18, 2018), using instruments such as the PPI, one of the most widely promoted tools in the sector to assess the depth of outreach of the microfinance institutions.

As already referred in this chapter, THRIVE is still a young institution, thus, requires additional funding to support the expansion of its operations; both in what refers to the funding of the loan portfolio and the structure of the institution (physical branches, equipment, human resources, etc.). This pressing need for funds has been reinforced in the context of the aggravated economic and cash crisis in Zimbabwe, and the associated (increased) repayment difficulties of the clients. The economic and political situation of the country contributes also to the restriction of available funds within the country, pushing THRIVE managers to seeking external funders, as in the case of LWC and Kiva.

An additional reason for this open attitude relates to the ownership and management structure of the institution, with the presence of foreign members both on the board and within the management team. Beyond an expected difference in attitude towards external investment, these foreign elements give the institution a distinct understanding of the advantages and limitations associated with this type of funding, and potentially facilitate the contacts between the parties. In this way, they can help overcome any anxiety or distrust.
towards outsiders that may exist in the institution. These circumstances are substantially
different from those experienced in the Pakistani MFI, which has a strong local basis,
fundamental to gain the trust of its own clients and attract donations from local supporters.

In this context, the development of the PPI for Zimbabwe was considered an important step
in the integration of the Zimbabwean microfinance market in the international community,
facilitating the access of the Zimbabwean MFIs to resources not available at the local level.

The application of PPI in THRIVE would have conciliated both the internal and external
motivations referred by Boucher (2014). However, neither PPI nor other externally validated
poverty assessment tools were available in Zimbabwe by the time of the project, which led
THRIVE managers to promote the development of an internally constructed poverty
scorecard, following the methodology applied by Mark Schreiner. The first note in this regard
should be to praise the initiative, which demonstrates the commitment of the institution’s
managers towards measuring social performance. The second note is to be alert to the
limitations of the exercise, which was implemented by staff with limited statistical skills and
not subject to validity tests as happens in the case of PPI. The external credibility of the
instrument is, thus, low and its utilisation in the determination of the absolute poverty levels
of the clients or in the comparison of the outreach between different MFIs is not advisable.

Despite its limitations, the analysis of the variation of the scores over time was useful to
complement the results of other indicators on the observed changes on the household
poverty levels, assuming that the indicators included in the index, although might not be the
best estimators of poverty, were related to it, and improvements in the scores could be seen
as an indication of improvements in the likelihood of the household being poor.

It should be noted that in the construction of THRIVE-PAT, province level data from the
Poverty, Income, Consumption and Expenditure Survey (PICES) 2011/12, conducted by the
Zimbabwe National Statistics Agency (ZIMSTAT), were used as statistical reference. This local
nature of the tool is expected to minimise the sub-group criticism pointed to the PPI and
other scorecards based on national survey data (Boucher, 2014), as demonstrated by Dinh
and Zeller (2010) for a locally developed poverty measurement tool in Northern Vietnam (F2
tool).

The efforts made by THRIVE managers to the development of a PPI for Zimbabwe, as well as
the process of construction and implementation of THRIVE-PAT, showed that the model of
implementation of the poverty scorecard will most likely be different from the one used in
Pakistan during the evaluation project. It should be expected that it will be closer to the
experiences in Ecuador of Fodemi (Boucher, 2014) and FACES, one of LWC field partners. Both institutions have integrated the PPI questions in their loan application process, with data being collected by the loan officers for all clients of the institution.

*Research Question 1.4: How do different stakeholders perceive and use the evaluation results?*

Similar to the broad approach to social performance advocated by LWC (LWC1, p.c., June 4, 2018), THRIVE managers have considered the evaluation project as one of the key elements in their decision-making process, which should be considered jointly with other sources of information. Among these additional sources of information were the social rating report and a number of instruments used by THRIVE to understand the perceptions of clients, including satisfaction surveys, exit surveys and the training sessions. Therefore, in Zimbabwe there was already a pre-disposition to listen to the clients and their needs, although this ‘bottom-top’ communication strategy was affected by the resources limitations of the institution, both in terms of time available to conduct these tasks and staff skills.

The differences between THRIVE and AIM in regard to the ‘bottom-top’ communication, similar to the learning motivation, can be partially explained by institutional characteristics, namely the size and the hierarchical structure of the organisations. The larger size and higher degree of formality in AIM, along with other cultural issues, made it more challenging to give voice to the clients when compared with a ‘lighter’ and more horizontal organisation like THRIVE.

In both settings, the collection of data from non-clients of the institution in parallel with the client survey, was important despite, as exposed in the previous sections of the chapter, the two groups of clients and non-clients in Zimbabwe turned out to be different with respect to many of the studied indicators, including those related to income, consumption and poverty. This was not a positive outcome from a formal impact assessment perspective as the identification of distinctive features between the two groups at baseline weakened the possibilities of attributing the observed changes in the lives of the clients to the participation in the programme. But, in line with Greeley’s (2005) suggestion that the participating MFIs in the Imp-Act project privileged “the organisational learning agenda to improve social performance” over a “proving agenda” (p.49), so did THRIVE managers which was most likely linked to the methodological compromises and limitations referred to in the discussion of the research question 1.1 (see p.186).

---

44 Information from FACES collected by the researcher during a field visit to the institution in February 2016.
The collection of data on the control group in the two waves of the survey had, however, a positive outcome as it allowed the MFI to gain insights on a segment of poor female entrepreneurs that apparently were not being served or, most likely, were underserved by the institution. In a more exploratory perspective of extending the outreach of the MFI to more vulnerable populations and reinforcing its social mission, this information could be critical to understand if the existing programme or other to be developed can answer the needs of these potential clients.

It should be noted that by the time of the baseline survey, THRIVE was experimenting a new microcredit programme (ISAL – Internal Savings and Loans Groups), with distinctive characteristics, aiming to attend poorer entrepreneurs developing more basic businesses. The programme was based on the widely spread informal savings and credit group model but included training, similar to THRIVE main microcredit programme. The first results were, however, not positive and the project was discarded. The main issue identified was a low take-up of the loans after the training period, which one of the MFI managers believed had to do with the combination of the maximum amount of the loans ($300) and the time for approval, considered too low and too long respectively, when compared with what these clients could access from informal lenders (aspects often more valued than the cost associated with the loans). The insights obtained from the evaluation project (which can be complemented with the foreseen application of the third wave of the survey), have given the opportunity to re-assess the ISAL experience and make evidence-based decisions regarding the more suitable strategies to potentially serve the more vulnerable segments.

Taking into account the final interview with the member of the management team, it should be expected that the participation in the impact project will be further explored - “for dynamic environments such as those prevalent in Zimbabwe and broadly the poorer regions of the world, evaluation of outcomes should be a continual process”, supporting development practitioners in their management decisions, including if “they should put more effort/resources or re-think interventions” (THR1, p.c., May 18, 2018).
8 Cross-Case Analysis: Evaluation Challenges and a Common Project Approach

In this chapter, the implementation of a set of common elements in the evaluations developed within the Lendwithcare (LWC) project is discussed, starting by comparatively analysing the findings in both countries. Given the diversity of contexts, institutions and programs, it should be once more emphasized that there was no intention of comparing the evaluation results of the two MFIs participating in the LWC project as they are not directly comparable.

The focus of the cross-case analysis was on the evaluation process itself, including the methodologies applied, and on the partnership established between LWC, the University of Portsmouth and the MFIs to deliver the project (see section 8.2.1).

It is equally important to clarify what is meant by a ‘common project approach’ to evaluation. Answering to a question about the common assessment approach adopted in the project, one of the LWC team members started by affirming that evaluations should always, as much as possible, “have a tailored approach depending on the context of the country” (LWC1, p.c., June 4, 2018) and the specific programme being evaluated. This seems to be a consensual notion in the microfinance sector and evaluation in general.

The common project approach in the LWC evaluation project translated into the application in the two settings of a longitudinal survey to a sample of new clients of the MFIs and non-clients living in the same geographical areas of the clients. The data analysis, although with adaptations to the specificities of each dataset, also had common aspects, including the exploration of heterogeneity among the MFIs’ clients. There were, however, a number of decisions regarding the design and implementation of the project that were adapted to the particularities of the local context taking into consideration the resources available. These decisions were highlighted in the description of the two programme evaluations. Examples included the decisions regarding the sample size, the language of the questionnaires, the recruitment process of the interviewers, the local of the interviews, and the strategies to locate and contact the non-clients.

Any attempt to harmonise strategies, methodologies and procedures in evaluation will most likely be subject to dissenting voices. This was referred to in chapter three in regard to the harmonisation attempt by the authors of the RCTs included in Banerjee et al. (2015), which as described by Odell (2015) was challenged by practitioners considering that this effort may have led to neglect the differences between the programmes and contexts, treating the microcredit programmes as the ‘same’. It is, thus, fundamental that a common project
approach to evaluation is carefully thought and is clearly communicated to the different audiences interested in the process and its results. Transparency in evaluation is essential to make it worthwhile exploring the potential gains associated with the implementation of a common project in the implementation of the evaluations.

8.1 Cross-case synthesis: similarities and differences

At first sight, the two evaluation settings could hardly be more different. The two MFIs operate in two countries located in two different continents - Asia and Africa, which translate into significant cultural, ethnic and religious differences. The religious factor is especially relevant because, as described in chapter six, Islam influences the development of the AIM microcredit programme in different aspects (demand, credibility of the institution, repayment rates, etc.), whereas in Zimbabwe the majority of the population is Christian, but the religious factor seems not to be so significant for THRIVE operations.

Another relevant difference between the two sites, which had practical implications on the implementation of the evaluation, refers to their geography. The branches included in the study are located in urban areas but with different characteristics. Lahore is a megacity, with more than 11 million inhabitants and a very high population density, which facilitates a relatively straightforward implementation of the AIM lending model (based on branches with small radius of operations) and of the household survey in terms of locating the homes/businesses of the respondents.

In turn, Harare is a much smaller city (1.6 million people), and above all, dispersed geographically, with comparatively a lower population density. As referred in the previous chapter, lack of resources, including time, to overcome this logistical constraint led to the decision by the LWC team to conduct the interviews with clients at THRIVE offices, offering a small monetary compensation for their time and the cost of transport. In this context, the initial contact to present the research project and the introduction made by the interviewers in the beginning of the questionnaires gained more importance to, at least, minimise the disadvantages of the interview setting. These were not limited to the potential response bias, including also the loss of possibilities to validate by observation some of the answers of the respondents, namely those related to the household assets and housing characteristics.

The description of the two programme evaluations in the previous chapters allows for the identification of two other relevant contextual differences. One is related to the economic development levels of the countries, which translate into Pakistan being classified as a lower-middle income country while Zimbabwe is among the low-income countries. The distinction between the two countries goes beyond the GDP figures, particularly with respect to the
higher level of uncertainty associated with the economic situation in Zimbabwe during the period of the research, which had a direct impact on the evaluation results for THRIVE.

The second factor is related to the levels of literacy in both countries. Zimbabwe clearly outperforms Pakistan in this respect, which reflects in the enormous difference in the client education statistics of the two MFIs. While 48% of the AIM clients had no formal education (67% in the case of women), with a further 16% having received only primary education, 91% of the THRIVE clients in the survey sample had achieved secondary education or above (see Annex IV). This is probably one of the reasons explaining the more developed and detailed answers given by the THRIVE clients to the open questions included in the questionnaires.

The differences between the two case studies are also noticeable at the level of the institutional characteristics and the microcredit programmes implemented. Analysing the evaluations conducted in the two institutions, it is important to take into consideration that, on the one hand, AIM is a large and mature MFI, with a strong local focus, experienced staff and a highly formal hierarchical structure. On the other hand, THRIVE is a small and young institution, founded by external investors, employing a capable but less experienced team, and with a comparatively more horizontal organisational structure.

These differences have several consequences. Among them, they influence the relationship of the MFIs and their evaluation teams with partners from other locations and backgrounds, and they affect the views of the MFIs’ managers regarding evaluation and its goals. Ultimately, the institutional culture contributed (along with other factors) to the decision of AIM managers to develop their own internal evaluation system, not having to rely on external parties, and at the same time, maintaining control over the process. Equally, they help explain the interest of THRIVE managers in the continuation of the evaluation project led and financially supported by LWC, and their enthusiasm towards externally validated evaluation tools, such as the PPI.

Finally, the two institutions implement microcredit programmes with different characteristics in terms of the credit product:

- The AIM programme is an example of Islamic microfinance based on interest-free loans, whereas THRIVE illustrates the conventional interest-based microcredit model;
- The loan cycle was significantly longer, on average, in AIM (15-20 months) compared with THRIVE (6 months);
AIM adopted an only-credit approach, while in Zimbabwe the THRIVE programme was an example of “microfinance plus” (Maes and Foose, 2006), including compulsory savings and training prior to the loan approval;

THRIIE was gender focused (all the microcredit applicants and clients were women), with AIM adopting a family approach by involving other family members in the loan process (the microcredit applicants and clients might be male or female). In June 2016, 42% of the AIM clients in Pakistan were women, which represented an increase of the female clients’ percentage relatively to the precedent years (Khan et al., 2017).

It should be noted that these characteristics of the programmes were in this initial phase of the project more significant *per se* in the analysis of the evaluation results for each MFI. In the design and implementation of the evaluation project, given the focus on main outcome indicators, these differences implied some adjustments, but they were not a deterrent from the perspective of adopting a common project approach in the two settings. The situation would have been more complex if the objective was to evaluate the processes and mechanisms of change associated with each programme.

The contextual, institutional and programme differences identified above illustrate the features of the two settings that made them interesting cases in the context of the research. The complexity associated with the contextual and institutional factors in the two countries was expected to introduce challenges in the implementation of the project which would allow for testing the application of the common project approach in the emergence of unexpected events.

Unexpected events did occur. Looking back at the description of the cases, it is easy to recognise that the ‘troubled’ second field visit to Pakistan had several implications for the preparation of the second wave of the survey and limited the possibilities of developing further qualitative work with AIM clients and field staff. Equally, the internal problems in the AIM branch of Kahna Nau led to no interviews being conducted with non-clients from that area in the second round of the survey, which was minimised after my insistence, supported by the LWC team, with a second phase of interviews in November 2017, six months later than planned. In the case of Zimbabwe, the deteriorating economic situation and the consequences of the cash crisis (including the measures adopted by the Zimbabwean government to minimise the problem), generated a period of economic uncertainty which affected strongly the small businesses developed by the majority of THRIVE clients and, consequently, the results obtained in the evaluation.
Although the referred episodes had strong implications both to the evaluation process and the evaluation results in the two cases, they did not discourage the partners, particularly LWC, in regard to the application of a similar methodological approach. This positive perception of the partners seems to indicate that the factors in which the two cases are similar had a greater weight in the process.

Among the similarities between the two cases there are also contextual factors, especially related to the status of financial inclusion and the microfinance sector in the countries. The microfinance markets in Pakistan and Zimbabwe, although in very different stages of development, are both growing, with increasing competition but still large segments of the population underserved, especially in the rural areas of the two countries.

According to the results of the latest Global FINDEX Survey, the overall access to formal finance in Pakistan and Zimbabwe has substantially improved in the period 2014-2017, however, this progress was accomplished through the growth of bank accounts (mobile accounts in the case of Zimbabwe), not reflecting in the same way in the credit data. In Pakistan, there was a slight improvement in the access to formal credit by the poorest segments of the population, which seems to reflect the increasing dynamism of the microcredit sector. However, access to finance remains at low levels and well below the average numbers for the countries with similar economic development levels, as pointed out in chapter six (World Bank, 2018).

Notwithstanding the different scales and idiosyncrasies specific to each country, the national microfinance networks have increasingly become more active and, in parallel, the government authorities in the two countries have channelled monetary funds to financial inclusion strategies and created or altered regulations specific to the microfinance sector. This increasing dynamic of different local actors in the microfinance sector and their interactions is at its early stages in Zimbabwe, but it has been fundamental in Pakistan as described in the respective case study (see section 6.1). In the case of Pakistan, it is important to stress the increasingly active role of the regional governments directly funding microcredit programmes, which are normally administered by MFIs, including AIM as mentioned previously.

In regard to the external factors, the two countries share also a negative feature related to political instability, which opens doors and aggravates the effects of other external events such as the cash crisis in Zimbabwe, the terrorist attacks in Pakistan or the extreme climate episodes which have affected both countries in recent years. The year of 2017 was politically significant in both countries with the resignation of the Pakistani Prime-Minister Nawaz
Sharif in July, after his condemnation in a corruption case linked to the Panama Papers, and the Zimbabwean President Robert Mugabe in November after a military operation launched by the Zimbabwe Defence Forces (Burke & Graham-Harrison, 2017; Rasmussen, 2017). Interestingly, Pakistan and Zimbabwe are preparing to hold elections in 25 July and 30 July, 2018, respectively.

The similarities between the two settings can also be found in factors linked to the institutions. Despite the differences mentioned above, both microcredit programmes are based on group lending methodologies and aspire to have a social impact, improving the lives of their clients by giving them access to business loans to develop their entrepreneurial activities. This shared social focus translates into an interest in assessing the performance of the institutions regarding a range of common outcomes, which facilitated the task of developing a similar evaluation methodology. Many of the questions included in the surveys were the same or were slightly modified in the hypotheses of answer (see the example given in chapter seven regarding the savings mechanisms), with the differences between the two cases being established in the interpretation of the results.

Also shared by the MFIs was the incipient stage in terms of performance measurement in general, even if THRIVE managers were taking some steps to measure social performance. At the beginning of the project, both institutions had management information systems not much sophisticated, with AIM branches working on a paper basis and THRIVE database not providing individual loan data for the members of the solidarity groups. Adding to this, there was also some unfamiliarity with the evaluation process and what it may imply.

The similarities mentioned so far were significant, but probably the most important factor present in the two settings, which contributed to the successful implementation of the evaluations, was the firm commitment of the MFI managers throughout the project. This strong motivation, based on the need to produce credible evidence of the outcomes of the programmes implemented, was one of the reasons which led the LWC team to select these MFIs for the project and it relates to Patton’s (2008, p.66) “personal factor”. Not only there was the buy-in by the management of the two institutions, but also the evaluation members of all the partners (PIU) were involved and committed to the process, with the leadership by LWC as well as the stability of the teams during the period being fundamental for the results obtained.

Underlying the project was a ‘subliminal’ intention of use of the evaluation, common to all those involved in the evaluation with decision-making power over the programmes. This intention of use was voiced by LWC1 in the final interview when declaring that the LWC team
would have not implemented the project if it was not clear “from the beginning that this is of practical use to all parties involved” (LWC1, p.c., June 4, 2018).

8.2 Insights from a UFE-based comparative analysis

8.2.1 Relevant elements in the evaluation design, implementation and use of findings

Much of the argument relating to the use of the UFE model has already been presented in chapters six and seven when describing and discussing the two individual case studies and answering the first research question. The objective of this section is to emphasize the elements in the two cases for the different stages of the evaluation which are identified as crucial in the LWC project and may be of use in the implementation of evaluations in other contexts.

8.2.1.1 Design

The design of an evaluation is determinant for the outcomes attained in the other stages of the evaluation, and the credibility and use of its results. Duvendack et al. (2011) relate some of the field implementation problems, namely attrition and response bias, to poor designs. It is, therefore, a crucial stage in which the first negotiations between the evaluation partners occur, and those involved in the project meet and begin a working relationship based on the evaluation project.

In the context of the thesis, the evaluations conducted within the LWC impact project were considered as inter-organisational partnerships, in the definition presented by Bendell (2011, p.14) - “an arrangement between one or more separate organisations to pursue a common activity or interest where risks and benefits are shared”. This broad definition encompasses diverse individual objectives as well as different levels of power and participation between the partners.

In what refers to the design stage, the preparation of the LWC evaluation project was led largely by LWC, with the decisions taken being influenced by two main factors. One of these factors was linked to the legitimacy of the process, as mentioned in section 5.4.2, leading to the partnership with the University of Portsmouth (academic consultant). The other factor was related to doing things differently from the mother organisation CARE International UK, looking for a solution capable of producing credible results and more cost-effective when compared with the commission of the project to consultants – “the research we [CARE] do through consultants is vastly more expensive and it’s not necessarily as insightful as what we’re getting” (LWC1, p.c., June 4, 2018).
In practice, the decision to collaborate with the academic institution introduced an initial restriction to this process related to the availability of the academic consultant/evaluator as my participation was dependent on the timings of the PhD. This resulted in me joining the project when the basic design of the evaluation was already defined by the project leader (LWC) and the beginning of the implementation in the field was scheduled.

The initial stages of the LWC project contrast sharply with the preparation stages of the Imp-Act project which included a one-year grant for “the consultation and design process, involving a wide range of stakeholders across the microfinance industry” (Copestake and Simanowitz, 2005a).

This situation was important as it implied that the initial phase of negotiations, in which the main issues were the discussion of the design of the questionnaires and the constitution of the control group, was limited in time, unlike Barnes (2001, p.19) in which the design of the first wave of the survey in Zimbabwe took several months. It was also conducted at distance through the exchange of emails. As recognised by LWC2 in the final interview, this was the first challenge of the project, particularly to the evaluator (LWC2, p.c., June 4, 2018).

Beyond the constraints associated with the timings of the project, the proposal and subsequent selection of the outcomes to be evaluated, as well as the respective survey questions, were influenced by two other factors. First, the intended short duration of the survey interviews, planned to take 20 to 25 minutes. This was a requisite of the LWC team and the field partners in anticipation that the clients would not be receptive to participate in longer interviews. This condition limited the number of questions to be included in the survey and, thus, the number and degree of detail of the outcomes evaluated. In a study implementing a similar methodology (longitudinal survey) in Kenya, Hulme and Mosley (1996b) referred to interviews of 1 hour. Barnes and Sebstad (2000), although considering that the interviews should be short, suggested durations up to 45 minutes.

The other factor to consider in the selection of the outcomes is the partially goal-free approach (Scriven, 1991) described in chapter five. This “open” perspective towards evaluation adopted by the LWC team (LWC1, p.c., June 4, 2018) gave me autonomy to define the broad outcomes to include in the questionnaire, being these generally accepted by the partners. On the same note, the project was not conceived by LWC with a specific evaluation theory or model in mind, which fits with the conclusions by Christie (2003, p.33) and House (2003, p.56) referred to in chapter two. In practice, evaluators often adopt parts of a given theory according to their aims and the context, not embracing a particular theory.
Nonetheless, there was from the start an underlying awareness among the three partners in the LWC project that the level of involvement required from all parties would only make sense in a context of use, independently of what use might represent for each partner. Moreover, for me, given the lack of experience in the implementation of evaluation projects, it was important to define a theoretical framework to guide the work in the different stages of the project. These considerations complement the reasons presented in chapter five regarding the choice of the UFE model in the PhD research.

The application of the use-focused evaluation model was especially important for me as evaluator to focus on the relevant aspects of the assessment for the partners identified as the primary intended users of the evaluation. It was not so useful in the definition of the methodologies to be employed and other practical decisions regarding the implementation of the project. In this sense, Pawson and Tilley (1997) criticism regarding the lack of clarity on ‘how to’ implement an evaluation seems justified.

In regard to the design of the evaluation and the preparation of the household survey, probably the major challenge in both settings was the constitution of the control group. The challenge started with convincing the field partners of its importance to the credibility of the results, given that the evaluation methodology chosen by the LWC team was based mainly in the implementation of a quantitative household survey. Unfamiliarity with the evaluation process and anticipation of the logistic problems contributed to this situation, which was reinforced in Pakistan by the timings of the project and the impossibility of meeting with the local evaluation team prior to the first field trip. The non-familiarity with the research process and, particularly the requirements associated with the constitution and maintenance of the target and control groups was a challenge also highlighted in the context of the evaluations implemented in the ‘Microfinance for Decent Work’ project (ILO, 2015, p. 28).

8.2.1.2 Implementation

In the implementation of the project, one of the most important challenges in both settings was related with the training and monitoring of the interviewers, particularly because the recruitment of local researchers to supervise this process was not considered by LWC.

In Pakistan, the initial training was accompanied by the validation by me and the LWC team members of 81 interviews, with special attention given to the understanding of the questions and the completeness of the questionnaires. Still, AIM2 referred to incomplete interviews as one of the identified problems in the quality control of the first survey (AIM2, p.c., March 22, 2018). Adding to this, an analysis of the duration of the interviews showed that it varied depending on the interviewers and the timing of the interview. Some were as short as 10
minutes, which could not be sufficient to avoid errors and complete all the questions included in the questionnaire and the PPI form.

In Zimbabwe, the average duration of the interviews varied also with the interviewer. The analysis of the questionnaires indicated that the interviewers with shorter interview times (on average) were also those whose questionnaires presented more errors and missing values. The registration of this indicator allowed me to use it as a reference in the training of the second batch of interviewers, which together with the more planned monitoring from THR2 (social performance officer), helped improve significantly the quality of the data.

The quality of the work of the interviewers is a common problem in evaluations (Adams & Vogel, 2013; Gaile and Foster, 1996). The challenges identified in the two settings were in many aspects similar to those experienced by Bolnick and Nelson. The researchers anticipated these issues and gave particular attention to the training of the enumerators and their monitoring during the implementation of the survey. However, the measures implemented were insufficient to solve the problems. As a consequence, they had to conduct additional training and interviews, the data cleaning process took longer, and the sample size was reduced (Bolnick and Nelson, 1990, p. 306).

In the LWC project, the quality of the interviews and data collected could have most likely improved if there had been a longer presence of the evaluator in the field and/or there was the recruitment of local researchers to supervise the implementation of the survey. This would have allowed for extended periods for piloting the questionnaires, training interviewers and support the monitoring process. However, given the duration of the field visits, this was not possible, and the involvement of local researchers was discarded by LWC.

The initial period of validation was, in all cases, important to minimise differences in the interpretation of the questions and reinforce the message on issues related to the conditions of realisation and the ethical procedures related to the interviews.

Despite the referred limitations, and not neglecting the measurement errors and response bias potentially introduced by these limitations, the overall quality of the questionnaires was considered by me as good in the context of the LWC impact project, given the objectives of the partners and their perception of rigour. In Zimbabwe, the effort in the second round of the survey to collect additional qualitative information will allow for deepening the analysis in the next stage of the project.  

45 In Pakistan, the interviewers were encouraged to take note as well of additional information regarding the clients or their households. As the evaluator did not have access to the individual
In regard to the implementation challenges, it is also important to draw attention to the issues related to data entry, and the different situations observed in the two countries. In Pakistan this task was internalised, being performed by AIM2. This allowed for some cross-validation of client data with the MFI database, especially in the first questionnaire since the borrowers had concluded their loan application process immediately before being invited to participate in the survey.\textsuperscript{46} It allowed also to increase the consistency of data as AIM2 was familiar with the programme and the project, but it implied loss of independence and a potential source of bias associated with the interpretation of data by AIM2.

In turn, the process in Zimbabwe was independent of the MFI. In the baseline survey, an external company in the UK was hired by LWC, but as described in chapter seven, the results were not satisfactory. In the second round, one of the volunteers at LWC head office in London was in charge of the data entry with a close monitoring by one of LWV evaluation team members (LWC2). The problems encountered in the first round of the survey implied an additional validation of the questionnaires, which represented a delay in the project, similar to the experience above referred of Bolnick and Nelson (1990). This occurrence had, however, an advantage as it allowed me to familiarise with the dataset and collect complementary information from the open questions and notes of the interviewers (which was possible because the answers and notes were registered in English).

The discussion of the field implementation challenges in both settings have a common denominator in the lack of experience of the teams in the conduction of an evaluation. There was, thereby, a learning process and an experimental approach to many of the challenges faced, even if taking into account the recommendations and suggestions of evaluators and researchers expressed in the academic and practitioner literature.

In the reaction to the field challenges described above and in the previous chapters, the principles associated by Patton with an UFE model, namely situational responsiveness, multiple evaluator roles, methodological flexibility and creativity (Patton, 2008, p.28) were important.

\textbf{8.2.1.3 Use of Findings}

The value of an evaluation for the use-focused evaluators stems from its usage (Wilson and Merten, 2012). The usefulness of the LWC impact project, and consequently of the questionnaires, receiving the datasets after computerisation of the information, it is not possible at this stage to assess the quantity and quality of this data, which can be written in Urdu.\textsuperscript{46} This cross-validation of client information was limited to the information included in the MFI database, since as described in the case study on AIM (section 6.2.2), the processes at branch level are paper based, with only a part of the information being computerised at central level.
evaluations implemented, was also a main concern for the practitioners involved in the project (LWC and the MFIs) as illustrated by the focus of the AIM manager on the project report – “we want this report. It will be very useful for us” (AIM1, p.c., March 22, 2018) and by the quotations included in Table 5.1 (see page 100) which allude to the sharing of the evaluation findings with external stakeholders and to the potential changes at the level of the organisations and the programmes (some aspirational at the time of writing the thesis, and some in implementation). To contextualise the use of findings in the two evaluation settings, it should be taken into consideration the different results obtained: positive and providing a sense of confirmation/legitimisation to the AIM managers; disappointing, but not surprising given the macro-context during the period, for the THRIVE managers. Independently of this distinction, by the end of the first phase of the project, the managers from both institutions were motivated to disseminate the evaluation results, sharing the project reports with other programme stakeholders (internal and external), and to implement a third round of the survey. They have equally considered the participation of their MFIs in the project very important from a learning perspective, as documented in the respective chapters.

The lessons learnt from the project and the resulting capacity building regarding evaluation links to process use as advocated by use-focused evaluators (Fetterman, 2003; King, 2007; Patton, 2008). Patton (2008, p.193) suggested different types of process use, which can co-exist in the same project. Relevant in the context of the LWC impact project were promoting evaluative thinking within the organisation, enhancing shared understandings regarding evaluation, increasing engagement and ownership over the evaluation findings and contributing to the development of the programme and the institution.

Even though the LWC team and the MFI managers have declared they will expect the results of the third round of the survey to decide on substantial modifications in the microcredit programmes, the changes associated with the implementation of social performance management within the internal processes of the institutions were already in motion.

Building evaluation or impact monitoring capacities within the institutions does not necessarily mean the creation of a “large internal impact assessment unit within the MFI”, which was traditionally the perspective assumed by many donors. Instead, it can mean supporting the MFI in the development of its internal management and information system (MIS) in order to collect data immediately available or easily collected within the normal operations of the MFI (Hulme, 2000b). In the AIM case, this seems to be the perspective with the institution focusing on the internal strategies and procedures, including the development of a new MIS- “we are going to a paperless environment (...), what we will try to do is we just
try to implement all this set of questions into a paperless environment where we can get some kind of reports through our MIS” (AIM1, p.c., March 22, 2018).

With respect to the use of the evaluation findings, it is important to remember that the origins of the project were connected with the traditional view on evaluation associated with the provision of credible evidence to support the judgement of the merit of the programme, which Patton named as overall summative judgment within the UFE model (Patton, 2008). Patton (2008, p.140-141) suggested five other potential purposes for the evaluation, namely formative improvement and learning, accountability (assessment of the effectiveness of the programme), monitoring, developmental and knowledge generating.

For LWC and THRIVE, the learning and the summative purposes were complementary from the start of the project. In AIM, although the summative perspective was the main drive for the participation of the MFI throughout the project, at the late stages of the evaluation the development of the new MIS in the organisation and the intention to link it to an internal evaluation system also meant the consideration of a monitoring purpose.

In parallel, the research conducted within the PhD can be associated with the knowledge generating purpose, which was considered by Patton from an incremental and cumulative perspective, with the lessons learnt from the evaluation contributing to the general knowledge and practice (Patton, 2008, p.139).

The following section of the chapter focuses on the personal factor and the role of the evaluator, considered within the UFE model as fundamental to increase the usefulness of the findings for the primary intended users. It is important, although, to consider that use does not depend exclusively on these factors. Contandriopoulos and Brousselle (2012), as described in chapter two, drew attention to the cost-sharing structure of the evaluation project, and to the fact that the partial or total support of the evaluation costs by the field partners encourages the use of its findings. In the case of the LWC impact project, this appears to be a two-way relationship. The partner MFIs were willing to invest in the project anticipating the use of the evaluation results, supporting a small part of the direct costs of implementation (mainly AIM) and opportunity costs associated with the involvement of staff and management in the project. At the same time, this expenditure boosted the motivation of the LWC team, main funder of the evaluation leading them to “scaling up that side of our work” (LWC2, p.c., June 4, 2018). In addition, the University of Portsmouth funded me to become the academic consultant through a PhD bursary.

In the discussion of the use of findings within the UFE model, it is important to reflect as well on the focus of the project on the primary intended users. This happened logically in the LWC
impact project given the objectives of the leader LWC and the recognition of resource limitations to be able to promote the active involvement of other programme stakeholders, namely field staff and clients. Not entirely deliberately, this perspective had also to do with a gradual immersion in the evaluation process, considering that the two MFIs had no previous experience in evaluations.

These considerations do not exempt the project from the criticism regarding the exclusion of relevant stakeholders, including the intended beneficiaries of the programme. This exclusion can lead to bias in favour of the most powerful stakeholders and, in the limit, co-option of the evaluator by the PIU (Fetterman, 2003; House, 2003; Pawson and Tilley, 1997; Stake, 1983b; Stufflebeam, 2001).

In relation to the active involvement of clients and field staff, I have suggested to the PIU to complement the quantitative survey approach with the application of qualitative instruments, namely focus groups, but these suggestions were not received with much enthusiasm from the field partners and external events hindered the activities planned. This is one of the questions that can be re-thought in the perspective of continuation of the project, considering that THRIVE managers signalled their interest in deepening the analysis of the results.

The concern with the potential bias towards the decision-making stakeholders and the possibility of co-option, also put forward by Pawson and Tilley (1997), was always a concern for me as evaluator. Finding the balance between maintaining independence and encouraging the partners’ involvement was challenging, particularly when promoting or participating in the previously mentioned parallel activities.

There were, however, conditions in the project that helped minimising these risks. The initial open perspective of the LWC team towards the evaluation outcomes (and underlying theory) allowed me to suggest the outcomes to be studied and, equally, I have independently conducted the data analysis, which helped avoid eventual pressures from the partners.

In addition, in line with Chen’s (1990) suggestion, I have considered myself as one of the project stakeholders whose main aim was guaranteeing a high-quality and ethical evaluation. This position was challenging in the relationship with the field partners, requiring constant negotiation in the application of the evaluation methodologies and gradually improving my understanding of to which extent rigour could be pushed without jeopardising the trust relationship between partners. Finally, simultaneously playing the roles of evaluator and PhD researcher implied a constant tension between the two roles, with the latter keeping in check the actions of the first.
The final argument in this section addresses the “dialogues across difference and distance”, as described by Wadsworth (2001), and to which Patton (1985) refers as cross-cultural encounters, concepts introduced in section 2.2.4 (see p. 43). The two case studies had associated different demographic attributes of context diversity (ethnicity, language, gender, religion, age, etc.), as well as distinct contextual dimensions of power, economy, living situation and culture (Sengupta et al., 2004), both at the programme level and at the evaluation team level. These differences will most likely increase with the intended extension of the project to other countries.

The diversity of characteristics and cultural influences present in the evaluation teams “enriched the project” (LWC1, p.c., June 4, 2018), but represented also a great challenge, especially for me as evaluator in the first stage of the project when trying to understand the perceptions and intentions of the field partners regarding the evaluation. The evaluator’s self-examination of values, assumptions and the cultural context as advocated by Sengupta et al. (2004), among others, was therefore a constant during the project, representing at times a frustrating process. It was especially felt when preparing the drafts for the preliminary reports, implying several readings through the documents before sharing them with the partners. The process became easier as the project developed and there was more interaction between the partners, with the parallel activities being helpful in this process, especially in the Pakistani case.

8.2.2 The personal factor and the role of the evaluator

This section discusses the role of the evaluation teams and the connected personal factor, central elements of the UFE model. The personal factor of an evaluation, as defined by Patton (2008, p.66) corresponds to a group of people linked to the programme who personally care about the evaluation and its findings, enhancing evaluation use. In the LWC impact project, these group of people was formed by the LWC team members, me as evaluator, the managers of the MFIs actively participating in the project and the MFIs operational team members.

As referred in previous chapters, the evaluation teams in LWC and the two countries showed strong motivations and personal commitment throughout the project, guaranteeing that the personal factor was an essential factor in the success of the project.

Among the members of the evaluation teams, The LWC team and the MFI managers were identified as the primary intended users (PIU). These are the ones with decision-making
power over the programmes and whose interests were primarily addressed in the evaluations conducted.

In parallel with the focus on the PIU, the role of the evaluator is of utmost importance in the UFE model. In accordance with the model prescriptions, I have, on the one hand, encouraged the permanent engagement of the PIU, being attentive and responsive to their interests. And, on the other hand, in line with Chen (1990), I have taken on the role of one of the project stakeholders whose main interest is the development of a rigorous and ethical evaluation.

In the context of the LWC impact project, I was considered to be the academic ‘consultant’ who was commissioned to coordinate the implementation of the evaluations. This perspective fits well with the UFE model approach as asserted by Owen and Lambert (1998) for whom the evaluator should be more a consultant than an external and objective judge of merit. In the different stages of the evaluation, as long as execution was deemed possible by the LWC team, my recommendations were accepted. In some questions, such as the control group, it took some time for the partners to fully appreciate the suggestions, but eventually it became a major challenge for the field partners (as expressed in the final interviews – see Table 5.1) and the main concern for LWC in the extension of the project to other countries.

Also relevant in the discussion of the LWC impact project is the non-participation of the LWC and the MFIs evaluation teams in the analysis of the data collected in the two settings. As previously mentioned, data analysis as well as the preparation of the final reports, including the decisions on their structure and contents, were independently conducted by me, with the partners being called to comment a quasi-finalised version of the final report.

This independence in the analysis and interpretation of the results did not invalidate the consideration of the PIU interests at this stage. As described in the two previous chapters, in parallel with the comprehensive and independent final reports, summary reports focusing on specific issues identified by the PIU were also prepared.

In practice, this ‘active-reactive-adaptive’ posture of the evaluator in relation to the PIU interests (Patton, 2008, p. 213) could clash with the objective of achieving a high quality and ethical evaluation. The fact that resources are scarce, particularly in small socially focused MFIs, and the evaluations and programmes are influenced by contextual factors can lead to compromises at field level which are not compatible with the desired highest quality standards, even putting aside situations in which the evaluation partners create barriers of access to the field, as discussed in Siwale (2015). This was exemplified in the LWC project where the geographical dispersion of the clients in Zimbabwe, allied with time and cost
constraints, led to the less than optimal decision of conducting the client interviews in THRIVE offices.

One last note related to the role of the evaluator, which makes the link to the involvement of the PIU in the project, refers to the strategies used to stimulate their engagement in the different stages of the evaluations. This task was in some degree facilitated by the strong motivation of the field partners, but this motivation could have faded over time considering the different implementation challenges faced. In the description of the evaluations conducted, the main engaging strategies were described, including parallel activities, such as the co-authorship with the Pakistani evaluation team of a book chapter, and the organisation of workshops and conference sessions with the participation of different elements of the two MFIs. These parallel activities seemed to have had positive results, contributing to building the relationships between the team members, particularly in the Pakistani case.

There was, however, an engagement strategy which did not work as expected - the preparation and request of feedback on the preliminary reports, which were prepared by me after analysing the data from the baseline surveys. Hulme (2000b) suggested the discussion of interim findings with the MFIs staff as a way to achieve co-ownership of the evaluation and increase the possibilities of the findings being effectively used in the decision-making process. This argument is along the same lines of the involvement of the programme stakeholders as enabler of the buy-in of the evaluation findings and recommendations invoked by the use-focused evaluators introduced in chapter two (Christie, 2003; Fetterman, 2003).

Part of the problem in this failed attempt to engage the field partners was most likely related to the communication mechanism. Instead of the eight-page written reports sent to LWC and the MFIs, oral presentations and discussion would probably have been more effective, judging by the feedback obtained during the partner meetings, with both AIM and THRIVE. This option was, however, problematic since the evaluation teams in each case were in three different locations. Skype (or similar mechanisms) could have been a solution but this alternative was only truly explored in the latter stages of the project when there was already some familiarity between the different evaluation team members.

From a different perspective, in this section is also important to call attention to the role of the operational members of the two MFIs, AIM2 and THR2. They were both of undeniable importance in the field implementation of the chosen methods, with their relevance being enhanced by the conditions of the project, particularly the limited access to the field of the evaluator. The field trips in both waves of the survey were short (five to 10 days), and as
controlling the implementation of the surveys at distance was not possible, the operational members in the two countries became key pieces in the quality control of the questionnaires.

In both cases, they were young (20s), and considered as capable and promising staff members by the MFI managers, being involved in other activities in the institutions. Both had completed secondary education and were attending university level courses in accounting and finance during the period of the LWC project. They were, however, inexperienced in the evaluation process and the long distance support of the evaluator was not sufficient to avoid some problems. Examples of these problems were the repetition of interviews with the same clients in the baseline survey in Zimbabwe, and the reaction to the internal problems in Kahna Nau, which led to no interviews being conducted with non-clients from this area in the second round of the survey in Pakistan. As described in the respective case studies, both situations were managed and overcome but with costs in terms of time and quality of the surveys.

On the positive side, these two evaluation team members, in whom the respective MFI managers seem to be investing, were expected to stay in the organisation (and did stay during the LWC project), giving stability to the evaluation team and retaining the acquired knowledge and experience inside the institutions. The staff turnover at the MFI’s level can be a major challenge in evaluation projects, as in the case of the ILO ‘Microfinance for Decent Work’ project (see section 3.3, p.70). This is particularly relevant, as recognised by Patton who signalled the possibility of turnover within the evaluation teams as one of the potential weaknesses of the UFE model (Patton, 2008; Stufflebeam & Coryn, 2014).

Last but not the least, in regard to the personal factor in the LWC project, is the discussion of the role of the LWC team in the process and their relationship with the field partners. The active participation of LWC in the evaluations had positive and negative aspects. It was fundamental to ensure the necessary resources to implement the project, including financial and technical resources such as the access to the academic consultant. Its ‘umbrella’ perspective, being constantly involved in the process in both settings, has contributed to the consistency in the implementation processes, the motivation of the partners and the valorisation of the results, increasing the credibility of the findings and the possibilities of a wider dissemination of the project and the results of the evaluations.

However, this active participation of the LWC team also entails some potential weaknesses. One of them is related to the potential control over the process and, eventually over its results, which is linked to the concern explored in the previous section of the thesis of potential bias towards the decision-making stakeholders, as put forward by Pawson and
Tilley (1997) among other evaluation theorists with similar concerns (House, 2003; Stake, 1983a; Stufflebeam, 2001). In section 8.2.1 (p.199), the discussion included also factors of mitigation of this bias within the project.

Other major concern, which was not explored in-depth during the PhD research, relates to the power relations between LWC and the partner MFIs. The implementation of the project in two field partners with diametrically opposed positions in their relation with the leader of the project raised a number of questions. AIM was a large and mature microfinance institution, not dependent of the funding obtained from LWC (see section 6.2.1, p.122). Its participation in the LWC impact project can, thus, be seen as opportunistic in the sense that the MFI managers saw an opportunity to legitimate their work through an externally led evaluation conducted in collaboration with a University. This participation ended up representing also a chance to learn about the evaluation process in the context of the implementation of a new MIS in the institution.

The situation was very different for THRIVE, for which the funding obtained from the crowdfunding platform was a significant source of funds for the institution (see section 7.2.1, p. 158). As described in the respective case study, the institution was young and small, trying to achieve financial sustainability in a challenging economic and political environment, and in this sense, in need to present credible evidence of achieving positive results to maintain its existing funders (including LWC) and attract new funders. Moreover, the MFI did not have financial capacity to implement an impact project of this nature, being dependent on LWC budget and the decisions taken by LWC team related to the field implementation of the evaluation (even if the LWC team consulted the MFI managers for many of these decisions).

From my experience in the project, I would suggest that the alignment of objectives of the two partners (LWC and THRIVE) and the cultural proximity resulting from the presence of a British manager in the MFI mitigated this power imbalance. However, looking at the average dimension of the LWC partners (see Table 5.3) and taking into account the cultural differences, including potential language barriers, between the LWC team and some of the MFI managers, this concern should not be taken lightly.

The growth of the impact project and its potential extension to other field partners raises one final question associated with the role of LWC, which is related to the capacity of the LWC team to maintain the level of involvement (time, financial and human resources) and the characteristics that have made the project work in this initial phase.
8.3 A common project approach to evaluation in microfinance

The description and discussion of the individual case studies, as well as the cross-case synthesis developed up to this point are fundamental to tackle the idea of a common project approach to the evaluation of microfinance programmes. In this regard, the projects previously introduced in chapter three and funded by the ODA, USAID (AIMS), Ford Foundation (Imp-Act) and ILO, as in previous sections of the thesis, were the main comparison references.

The AIMS project represent a good benchmark for the LWC project as similarly, the three evaluations highlighted by Goldberg (2005) were implemented in urban areas located in different continents, and the methodologies applied appear to be generally similar (longitudinal data, control group formed by non-clients, before-after comparison). The most significant difference in the design of the evaluations between the two projects is related to the client sample. In the LWC project the clients participating in the first round of the survey were new-clients, which allowed for collecting baseline data, while in the AIMS project, they were existing clients in different loan cycles.

Equally a reference in this respect are the five evaluations presented by Hulme and Mosley (1996a; 1996b). As mentioned in chapter three, the case studies published did not detail the methodological choices made in all the settings, but they seem to have commonalities. It seems also that, although local research teams were recruited in each country, the evaluation teams were led by the two researchers in all the settings.

The LWC project seem to differ from the ILO and the Imp-Act projects, as in these two projects there were a more customised approach, with different methodologies being implemented in response to different research questions and cultural and institutional contexts. There were points of contact between the individual projects being implemented, as in the case of the adoption of similar AIMS tools in many of the Imp-Act projects, but there was not a common approach within the projects in this regard.

In the comparisons established between the LWC evaluation project and particularly the ILO and Imp-Act projects, it should be taken into account the differences in budget and dimension. Not only the number of participating MFIs was significantly higher (16 and 30, respectively), also the academics coordination team (in both cases located in the UK) and the local teams hired in the case of Imp-Act represented a level of costs that is not in any way comparable with the investment of LWC in this first phase of its project. This should not be surprising if considering the size of LWC as an institution (or rather said as a small department
within a large development organisation), and the exploratory approach to the impact evaluation and social performance management processes adopted.

8.3.1 Feasibility

Research Question 2.1: To which extent is it feasible to implement a common project approach to evaluation considering the diversity and complexity of the contexts in which microfinance programmes are implemented?

The answer to this question is, first of all, dependent on the adopted definition of common project approach and the amplitude of the concept. There is little doubt that the implementation of any evaluation is highly constrained by contextual factors (House, 2003; Patton, 2008), which condition all the stages of the process from the design of the evaluation to the use of its findings. Therefore, the “one-size-fits-all” approach, as often referred in the context of social policies and programmes, is not applicable to evaluation in microfinance and it does not correspond to the idea of common approach developed within the thesis.

In the LWC evaluation project, the common project approach represented a common overall research question, field partners with social objectives and evaluated programmes based on the same core product (productive credit) and a similar methodological approach (household survey using a purposely built questionnaire and a poverty scorecard; similar sampling strategy and analysis techniques). The evaluations also had a similar structure in terms of the stages of the project and composition of the evaluation teams. There were not, in any of the aspects mentioned, rigid propositions and throughout the project each of these was affected or affected the decisions taken in the two settings of the project.

One of the best examples of the influence of contextual factors and the required flexibility to deal with them relates to the utilisation of the Poverty Probability Index (PPI). The LWC team was keen on using PPI with the objective of assessing the poverty levels of the MFIs’ clients and track their changes over time. The expectation was that the application of an international recognised tool would increase the credibility of the results obtained and would function as a benchmark for comparisons with similar institutions in the respective country.

The choice of partners in this first stage of the project resulted, however, in findings that differed from those anticipated by the LWC team. On one hand, as described in chapter six, the poverty scorecard for Pakistan was clearly outdated, which contributed to a negative perception by the Pakistani managers relatively to the use of PPI. It also did not help that the unique features of the microcredit model implemented by AIM, based on the provision of *qard hasan* (interest-free) loans, has made the benchmarking possibilities associated with
the use of PPI, less appealing. On the other hand, there was no PPI developed for Zimbabwe, despite the enthusiasm towards the tool showed by THRIVE managers, which has led the institution to develop their own poverty scorecard.

Given the importance attributed by LWC to the use of PPI, two different, although based on similar methodologies, poverty assessment tools were employed in the project. It should be noted that this challenge regarding the utilisation of PPI would probably not have been discussed here if the project partners were from other countries. Zimbabwe was the only country among those in which LWC operates for which there was no PPI. For the other 10 countries, PPIs were in general updated (updates during 2018 for the Philippines and Peru), with the exceptions being the already mentioned Pakistan and Vietnam, both with poverty scorecards dated of 2009 (IPA, 2018). These considerations around the application of PPI link with the concern voiced by Bhat and Yamini (2012), and echoed in Boucher (2014), regarding the ownership of the tool and its consequences in terms of updating the index and adapting it to the local contexts.

There were two factors in the LWC project that were essential to the successful implementation of a common project approach from the partners’ perspective. First, the active participation in the two settings of the core LWC team and the evaluator, which was possible given the scale of the project at this stage. It compares with the common lead assumed by Hulme and Mosley in the ODA project, but contrasts with the approach followed in the AIMS project. In each of the three studies, the team of evaluators, including the lead evaluators, seem to have been different, despite interactions between the researchers within the global AIMS project (Barnes, 2001; Chen & Snodgrass, 2001; Dunn & Arbuckle Jr, 2001). It contrasts even more with the ILO and Imp-Act projects in which the MFIs had a much more prominent role in the design of the evaluations and its adaptation to their goals and contexts (Copestake & Simanowitz, 2005a; ILO, 2015). This increased ownership of the research projects was, as referred to in section 3.2.1, one of the main features of the Imp-Act project as intended by its promoter.

The other important factor was the experimental approach and learning attitude of the LWC team, the evaluator and the field partners in regard to the evaluation project, which reflected in the partial goal-free approach and the initial focus on broad categories of outcomes. If the strategy had been different, with the definition of specific theories of change for each microcredit programme and the discussion with the respective MFI managers of their outcomes of interest, it would have probably made more sense to think about individual evaluations and not a common project. The differences between the two institutions and programmes, summarised in the initial section of the chapter, would have implied the
definition of a specific set of indicators related to the outcomes of the programmes but also to the mechanisms of change leading to the observed outcomes in each case. Furthermore, a more in-depth analysis of the programme outcomes would also have required the application of different evaluation methods, including qualitative methodologies, as it was the case in the Imp-Act project, which suggests a different level of involvement from the MFIs. This was not, however, the objective of the two MFI managers, which was most likely linked to the lack of preparation of the institutions to adopt from the beginning more ambitious goals with respect to the evaluation.

From the exposed above, it can be concluded that the common project approach as a strategy to be considered by evaluation promoters or commissioners requires specific conditions related to the objectives and the characteristics of the institutions involved in the process.

8.3.2 Advantages and limitations

Research Question 2.2: What are the advantages and limitations of adopting a common project in the implementation of evaluations in different contexts?

The advantages of a common project approach to evaluation can be understood at different levels. From the perspective of the LWC team, the approach allowed for accessing additional information on the two MFIs, harmonised to a certain extent as similar indicators were used for the broader categories of outcomes, while equally achieving objectives in regard to building technical capacity within its partners (in this case focusing on evaluation and social performance assessment skills). The implementation of the impact project, and its potential extension to other LWC field partners, can therefore contribute to build communalities across the LWC portfolio. In addition, these advantages were achieved using common resources, including the academic consultancy, improving the cost-effectiveness associated with the evaluations.

To the MFIs participating in the project, there were two main advantages. First, they had the opportunity to participate in a multiple-country impact evaluation project led by LWC, an initiative of CARE International UK, a well-known and recognised developed NGO working in over 90 countries. In terms of the credibility of the results and the dissemination of the work conducted locally by the MFI, the project opened possibilities which were hardly achievable within an independent evaluation project.

Second, the application of common methodologies in the two settings allowed for an enriched learning process, integrating experiences and knowledge from different contexts.
and different perspectives, while pursuing common overall evaluation goals. The exchange of experiences between the evaluation teams during the implementation of the project, and with other LWC partner MFIs towards the end of the first phase, was one of the most positive effects of the project, independently of the results of the evaluations. It is important to note, however, that this process was successful in large part due to its active promotion by LWC, especially during the partners’ workshop in London (March 2018).

Another relevant aspect of the project was the promotion of a shared understanding of evaluation and the development of a common language among the partners, both contributing to gradually change the mindset of the MFI managers towards evaluation, which was briefly mentioned by LWC1 in the final interview when acknowledging that the partners were now starting to look at evaluation research more seriously (LWC1, p.c., June 4, 2018).

Finally, and as a cumulative result of these identified advantages, the common project approach had associated a motivational factor, enhanced by the partners shared overall aim of improving the lives of poor and more vulnerable populations in their respective countries. The positive perception regarding a common approach to evaluation, as mentioned in the previous section, was facilitated by the fact that a part of the evaluation team was the same in the two settings. It introduced, however, constraints in terms of the timings and availability of these common members, located in the UK, to work in the project and to travel to the two countries. The fact that I was a full-time PhD was helpful, as it allowed for more flexibility to adapt to the timings of the other partners.

Another aspect of the process that needs careful consideration relates to the power balance between the partners, similar to the referred for the individual evaluations regarding the risks of the active participation of LWC in the process. This is particularly important in what refers to the leader and main interlocutor between the institutions, but also in the interactions between the MFIs, which are very different in size, maturity, programmes implemented, etc.

Considering more practical limitations associated with the implementation of the evaluations, the different languages spoken in the LWC partner MFIs introduce constraints in the communications between the evaluation teams. Even if English is commonly spoken, it is not the native language for some of the evaluation members, including the evaluator. It also implies the adaptation and, eventually, translation of the survey questionnaires and other project materials (e.g. interviewer guides). Having in mind the need to adapt the questionnaires to the local contexts, and in the perspective of including more specific analysis as the project continues, the utilisation of a common questionnaire will be increasingly more
difficult. The solution may be the creation of a longer questionnaire template in English based on the initial experiences in Pakistan and Zimbabwe. This would allow for the different evaluation teams to choose the outcomes of interest and change the wording of the questions to better fit the local specificities (including the local languages). Ideally, this questionnaire template would be connected with the guidelines on outcomes management provided by SPTF (Sinha, 2017b).

In the adoption of a common project approach to evaluations in different settings, the communication of the project needs to be transparent in the processes used and the associated advantages and limitations, otherwise there will be a high probability of misunderstanding and misinterpretation of the evaluations conducted. Barnes and Sebstad (2000, p.2) refer to the “sceptical audiences or stakeholders who seek to discredit findings” often encountered by evaluators. In the implementation and communication of the evaluation processes, permanent attention should be given to the balance between harmonisation of the approach and customisation to the local contexts in order to address the expectations and motivations of the parties involved in the project.

### 8.3.3 Implications for the MFIs and the sector

*Research Question 2.3: How can the application of a common project approach to evaluation contribute to the valorisation of the evaluations and the research conducted in the microfinance sector regarding its impact?*

This is a more complex question at this stage of the project, but it seems important to offer initial considerations on the topic. The first comment is to say that the LWC project and, concretely, the strategy to implement a common project approach in the two settings cannot have much impact outside the project if it will be considered on its own. It should instead be perceived as one effort among other initiatives aiming to build a common understanding and a common language regarding evaluation, outcomes management and impact assessment, which will ultimately contribute to value the work of different actors.

This is nothing new. Barnes and Sebstad (2000) aimed to lay down standards for credible, useful and cost-effective impact assessments in microfinance. The proposed guidelines were developed within the AIMS project, taking in consideration that the academic standards for a rigorous impact assessment, which the authors associated with the implementation of costly experimental methodologies, were out of reach for the majority of the MFIs.

The AIMS guidelines and tools became relevant in the sector, and were the methodological basis used by most MFIs participating in the Imp-Act action-research project. This choice was
presented by Greeley (2005) as relating not only to the level of rigour associated with the tools, but also because they were considered to be “practical, cost-effective and relevant”. This was important since in many of the cases, the MFIs had privileged the learning agenda regarding social performance in comparison with external accountability objectives, which would imply focusing on attribution issues and methodologies capable of addressing selection bias and fungibility problems (pp.48-49).

This distinction of standards between a more academic approach (emphasizing rigour) and a more practitioner perspective (stressing usefulness) was also mentioned by LWC1 in the final interview. While academics were viewed as privileging maximum rigour, MFI managers and social investors were more flexible, willing to accept lower standards resulting from compromises at field level in the implementation of the evaluations as long as these were deemed necessary to guarantee the feasibility of the assessment, and an adequate level of quality was ensured (LWC1, p.c., June 4, 2018).

The argument seems reasonable but raises the difficult question of identifying what are necessary compromises in terms of implementation and what is the adequate level of quality. Attempts as the one mentioned in the end of section 3.2.3 within the ILO ‘Microfinance for Decent Work’ project to classify the overall credibility of the evaluations can be a good starting point in the way that they have established a number of factors to take into account when analysing the findings of an evaluation (ILO, 2015). However, two main problems reside. First, the definition of criteria for each of the factors, which is not totally clear in the project report. While criteria for sample sizes based on statistical criteria should be relatively easy to agree on, the task of defining good data quality seems to be much more complex. The second problem relates to the interpretation of the classification. Although, the report clarifies that a lower overall credibility of the evaluations does not mean that there are not results of significance, it should be expected that many readers would ignore this note or not get to read at all.

At this level, supranational organisations, in this case SPTF, can make more of a difference than isolated reports from projects such as the ODA, AIMS, Imp-Act or the ILO projects. The empirical knowledge and practical insights from these projects is certainly valuable on its own but they should also be absorbed and integrated in the common effort to develop harmonised standards and language for evaluation in the microfinance sector. This is obviously much easier said than done, and the history of SPTF, as of other organisations aiming to develop common standards (e.g. the SMART Campaign for client protection), illustrates well the challenges. On a positive side, SPTF seems to be moving decisively forward.
with the publication of the two mentioned reports on outcomes management (Sinha, 2017a; Sinha, 2017b).

Stepping down from the supranational to the level of the project, it should be noted that the LWC project has functioned as motivational booster for the LWC field partners which were not involved in the first phase of the project, but it has also the potential to be an example to explore by other social investors and donors working with socially focused MFIs in different contexts which are still lacking in terms of their social performance assessment.
9 Conclusions

The majority of the microfinance institutions, especially those which remain socially focused, “are not set up to rigorously evaluate” their work, as suggested by Hughes and Hutchings (2011, p.iii) with respect to development NGOs. Given the relevance of these institutions in the microfinance and financial inclusion sectors, the development of a widespread evaluation culture which adopts a broad definition of accountability and impact assessment in the terms defined in chapter two are crucial to improve the programmes implemented and their results at different levels.

9.1 Setting up the research

In chapter one, the merits of microfinance which contributed to the development and growth of the sector were briefly presented, highlighting the leading role of socially focused institutions, which are normally more constrained in the use of their resources. The importance for poor and vulnerable populations of accessing and properly using formal financial services is considered along with the emergence of delinquency crises in several microfinance markets across the world, and the questions raised by academic research using experimental methodologies with respect to the impact of the microfinance programmes on poverty reduction and other social outcomes. The conjugation of these factors makes the case for the evaluation of the outcomes and impact of microfinance, especially of microcredit programmes.

Despite the gradual growth in utilisation of other financial products, including savings, microinsurance and mobile payments, and the increasing consideration of the poor financial needs from the wider perspective of financial inclusion, microcredit is still the core product of many MFIs and the point of access to formal financial services for many low income entrepreneurs in the developing countries.

The focus on productive credit, exemplified by the two MFIs and the social investor involved in this research, derives largely from the expected outcomes of the programmes at different levels described in chapter three. In this chapter, the theoretical and empirical research on the topic is summarised, with particular attention being given to the methodological choices made in the assessment of microcredit programmes. The failure of many of the evaluations conducted in the sector to attribute (in some degree) the observed changes in the lives of the clients (positive and negative) to the microcredit interventions are frequently associated with methodological issues. These combined with the heterogeneity of the programmes and their outcomes, make it difficult to document microfinance impact (Adams and Vogel, 2013).
This discussion links to a topic introduced in chapter one, but running through the entire thesis, the different perspectives of the different actors involved in the microfinance sector. Among these actors, a common distinction is made between academics and practitioners with differences stemming from distinct interests, mindsets and capacities to interpret impact data, which reflects in the language used in regard to evaluation. The confusion regarding the language, and the different perceptions of credibility associated with different concepts, can be perceived in the academic literature on microfinance, namely in the ambiguous utilisation of the term ‘impact’ in many publications.

In the majority of the systematic reviews and sector reports presented in chapter three prevails a positivist approach, which associates rigour and credibility with the application of experimental methodologies (Duvendack et al., 2011; Goldberg, 2005; Odell, 2010, 2015; Vaessen et al., 2014; Van Royen et al., 202). In this dominant literature, the findings of evaluations using alternative methodologies, including quasi-experimental methodologies, seem to be somehow discredited.

It should be noted that many of these reports focus on formal impact assessment, valuing the studies which more effectively cope with selection bias and fungibility, the two main methodological issues in the assessment of social interventions according to the positivist approach (Adams & Vogel, 2013; Goldberg, 2005; Khandker, 1998; Sebstad & Chen, 1996). Moreover, some of the authors explicitly recognise that RCTs are not applicable in all circumstances and give attention to the role of qualitative methodologies, either on its own, as in the recent systematic review carried out by Peters et al. (2016), or most commonly, within a mixed methods approach (Duvendack et al., 2011; Odell, 2015).

There is no question that many of the denominated impact studies lack quality from an academic perspective, as concluded by the authors of the systematic reviews, and in some cases may not be serious in their purposes. In these cases, although a formal impact assessment may not be feasible, the evaluations can provide “descriptive indicators of the observed changes” as suggested in the ILO report (ILO, 2015), contributing to clarify the role of the microfinance programme in these changes and to build an ‘impact story’ for the institution over time though the combination and triangulation of different sources of information.

Moreover, the implementation of a systematic and flexible process of outcomes measurement and management can enhance the use of evaluation findings in the decision-making process of the institution, contributing also to increase the credibility and reputation of the institutions and programmes internally and externally.
The fact that the scenario just described is not prevalent in the microfinance sector (and probably in other development areas) can be attributed, at least partially, to the different perspectives regarding impact assessment and particularly the tension between rigour and usefulness associated with the design and implementation of an evaluation. The opportunity to participate in the LWC assessment project has given me as researcher the opportunity to contribute to this discussion with the analysis of a collaborative evaluation project.

9.2 The parallel with utilisation-focused evaluations

The complexity associated with the microfinances programmes and their evaluation is not confined to the sector, being identified by researchers from different fields in the development (e.g. Lewis and Mosse, 2006) and social (Guba & Lincoln, 1989; Pawson, 2014) sectors. It seemed, thus, sensitive to look upon the literature in some of these fields, namely education, from which many evaluation theories and models have been developed.

Chapter two starts with a brief description of the main evaluation theories and models, using as reference Alkin and Christie’s (2004; 2013) evaluation tree approach and Stufflebeam and Coryn’s (2014) classification of evaluation models and approaches. This introduction is meant to give an overview of the main theoretical issues with regard to evaluation theories and provide elements to justify my choice to use the utilisation-focused model in the PhD research. The UFE model is used in the establishment of the theoretical framework applied in the analysis of the LWC project, which is presented in chapter five. As described in this chapter, the focus on the UFE model resulted mainly from the adoption of a pragmatic approach to the evaluation project and the research, which is closer to the researcher preferences, and above all, fits better the conditions of the project. Furthermore, as mentioned previously, this approach contributed to ensure a smoother interaction between the partners with an easier acceptance of the idea of evaluation by the field partners.

From my perspective as researcher, the reference to the UFE model was particularly important to focus throughout the project on the relevant issues for the identified primary intended users of the evaluation and, simultaneously, remain attentive and promote high ethical standards in the implementation of the evaluations. The project dual objective of producing credible evidence for accountability purposes and building capacity on evaluation within the MFIs, which ultimately mirrors the distinction between summative and formative evaluation as put forward initially by Michael Scriven (Stufflebeam & Coryn, 2014, p.183), implied for me being aware and dealing with the trade-off between independence and participation, or in other words, rigour and usefulness. In this process, it was especially
important to alert the partners for the implications of the decisions taken in response to the different implementation challenges.

The LWC project, as many previous evaluations in the sector, positioned in between the two extremes - the ‘totally’ independent evaluation, privileging rigour, which is required to an external judgment of the merit of the programme, and the internal evaluation, favouring usefulness criteria and focusing on the improvement of the programmes and the organisations. LWC1 considered that “we reached actually a fairly good balance between the two” objectives (LWC1, p.c., June 4, 2018), a personal opinion which as such is subjective. From my perspective, as academic, the compromises at field level and the difficulties encountered regarding the control group in both settings raise questions in terms of data quality. For this reason, I consider that the objectives in terms of local capacity building and changes in the LWC team and the field partners’ attitudes regarding research, evaluation and social performance management were much more accomplished.

The parallel between the LWC project and an UFE allowed also to recognise some of the limitations associated with the project, which are referred to in the discussion developed in chapter eight (section 8.2). Among them, it is important to highlight the almost exclusive attention given to the programme decision-makers directly involved in the project (MFIs managers and the LWC team) when compared with other important stakeholders, particularly the clients and field staff of the institutions. Although the social mission of the MFIs was a permanent reference in the discussions of the partners, the potential stakeholder/primary intended users bias associated with the focus on the decision-makers of the programmes was not forgotten, in line with the recommendations of House (2003), Pawson and Tilley (1997) and Stufflebeam (2001). In the possibility of continuation of the project, this will continue to be one of the topics in the agenda of the evaluator/academic consultant.

Taking into account these concluding considerations regarding the use of the UFE model in the analysis of the LWC project, as well as the numerous references to the model in the thesis, especially in the analysis chapters, the answer to the third research question introduced in chapter one will have to be affirmative from the perspective of the researcher. This was not the only approach possible, but the use of the UFE model gave structure to the analysis and drew attention to aspects which may have been overlooked otherwise.

### 9.3 Contributions towards transparency and credibility

The research implemented within the thesis provided additional contributions towards the promotion of transparency in the microfinance sector and the identification of strategies and
tools aiming at increasing the credibility of the evaluations, and subsequently, the use of the findings and the lessons learnt from the process.

These contributions include the extensive descriptions of the evaluations processes in Pakistan and Zimbabwe in chapters six and seven. The chapters focused particularly on the field challenges encountered in the two settings, as these are frequently omitted or given little attention in the publications discussing evaluations and impact assessments conducted in the sector. Among the challenges identified were the coordination of the timings of the three partners, the constitution and maintenance of the control group, the translation of the questionnaires to the local languages, the training and monitoring of the recruited interviewers, the local of the interviews, the data entry process and data analysis. As previously referred, understanding the specificities of the problems faced and identifying possible solutions are important to increase transparency and trust in the evaluation results.

The thesis includes also the discussion of the common elements of the LWC project in the two settings and how this common project approach influenced the participation of the partners and motivated other LWC partners to consider more seriously the evaluation of their own programmes. This analysis was presented in chapter eight, being given particular attention to the specific conditions of the project which have made it successful in this valorisation of the evaluation process.

These conditions included the relation of LWC with its partners, giving LWC authority to conduct the process, and the association with an academic partner, and they seemed to have been working well with institutions that were willing, but still had not gathered the conditions (motivation, leadership, skills, funds) to implement social performance assessment in a more systematic and rigorous way.

The impact project implemented privileged general questions regarding the outcomes of the programmes at the client level, questions which are common to many evaluation projects, and which required some adaptation to the local conditions but still allowed for the application of a common set of questions in the two settings. A more in-depth exploration of the outcomes of the programmes or the analysis of the programme processes and its mechanisms of change, in line for example with the application of a comprehensive CIPP model (Stufflebeam, 1983), would imply a more tailored approach to the process. It would most likely also imply the development of the methodologies employed and the adoption of a mixed methods approach, with qualitative methodologies complementing the information collected via quantitative questionnaires.
Simanowitz (2003) attributed an important role to donors in the promotion of transparency in reporting outreach and in the support of research on impact, especially in MFIs with limited resources to conduct rigorous impact assessments. Traditionally, the role of donors in this process has been of financing the projects, hiring third parties, in many cases consultancy firms, to conduct overall summative judgment evaluations. The LWC project presents an alternative to this approach, similar to the Imp-Act project, in which the evaluation expert (academic, consultant) remains a central piece of the process but both the funder and the MFIs have also an active involvement in the evaluation. This approach loses in the independence of the findings, and thus, on its utility as a pure judgmental instrument, but gains (enormously) in the development of evaluative thinking within the institutions and in building internal capacity in regard to evaluation and social performance assessment and management.\textsuperscript{47}

Being closer to the interests and objectives of the microfinance institutions and adopting, even if implicitly, a use-based approach to evaluation, this approach can contribute to overcome the initial barriers faced by many MFIs, such as THRIVE (which have the interest but do not know where to start and do not have the necessary funds), and at the same time, minimise the evaluation anxiety and suspicious towards evaluation and evaluation experts, which to some extent happened in the case of AIM.

Finally, the thesis provides insights regarding the potential role of the Poverty Probability Index in increasing the credibility of the evaluations. Similar to the conclusion regarding the adoption of a common project approach, the application of the PPI in the LWC project showed that the materialisation of the expected benefits of the tool are dependent on a number of conditions, which in the project were largely related with the ownership of the tool. The fact that the development of the national PPI and its updates are controlled by the promoter of the tool implied that the managers of THRIVE who perceived the instrument as important to increase the credibility of the evidence they can provide to their sponsors, did not have access to it since there was no PPI developed for Zimbabwe. In turn, the poverty scorecard for Pakistan was clearly outdated which contributed strongly for the negative perception of AIM managers towards the instrument.

The analysis of the use of PPI was not as developed in the thesis as initially intended. This is a topic in which there were still a small number of independent empirical studies, and most

\textsuperscript{47} From the experience in the implementation of the LWC project, it seems that the dimension of the gains is closely linked to the motivation and commitment of the partners towards the evaluation, which was an element always present in the two MFIs involved in the project. For this reason, enormously is within brackets.
of them were based in the assessment of one-time applications of the tool (cross-sectional
data), which can actually be counterproductive and contribute as well to negative
perceptions on the usefulness of the instrument.

In the two evaluations described in the thesis, the application of PPI had associated
difficulties that for the majority of the other LWC field partners would not have been relevant
since the respective country poverty scorecards were updated recently. For these partners,
the application of PPI should be considered and evaluated as a continuous process. This
approach will allow the MFI managers to select the most suitable model of implementation
for the organization, and to adapt and complement the scorecard with other sources of
information, integrating in the analysis the local and organizational idiosyncrasies. In this
way, PPI might not be a mere accountability measure, but an important part of an outcomes
evaluation and management system based on self-assessment and promoting learning and
improvements within the institution.

9.4 Strengths and weaknesses of the research

Different issues are relevant in this analysis of the strengths and weaknesses of the research,
but the main questions to be raised are closely associated with the application of the
participant-observation methodology. The level of detail and the trust associated with the
information collected during the three and a half years of implementation of the project
would have been hard to achieve with any other methodology. The insider perspective, and
the relatively long time-span of the project, gave me the possibility to provide comprehensive
descriptions of the evaluations implemented, which was one of the main objectives of the
research. These two factors were also fundamental to gain a better understanding of the
contextual factors of the programmes and the evaluations, allowing for a more robust
discussion and analysis of the findings of the case studies.

Considering the complexity of the two evaluation settings, both in terms of the political and
socio-economic environment, and the institutional characteristics of the MFIs, it is important
to contextualise this better understanding of the contextual factors. Similar to Gravensteijn
(2014, p.74), this was most likely insufficient to “capture the full realities” of the institutions,
and particularly, of all the stakeholders of the organisations. It was, however, very helpful for
me in the implementation of the evaluations and the PhD research, given my limited initial
knowledge of the contexts and the reduced time spent in the two countries.

The dual role as evaluator (participant) and researcher, as described in the literature on
methodologies referred in chapter five (DeWalt & DeWalt, 2011; Yin, 2018), was highly
demanding. The balance of the time dedicated to each role was particularly challenging, with
the implementation of some of the stages of the process requiring more time than expected or being delayed, which reduced the time available to analyse and discuss the data collected.

This limitation is reflected in the chapters describing the two evaluations, in which the discussion around the first research question and sub-questions was not as developed as the first descriptive parts of the chapters. Equally, some of the topics emerging from the discussion of the research, such as the power relationships between the partners were identified, but not further detailed and analysed. These are, however, issues which I expect to address through future publications.

The role as participant in the project implied a permanent attention to be impartial and objective in the analysis of the evaluation processes and to not be influenced by the working relationship established with the LWC and the local evaluation teams, avoiding possibilities of co-option by the partners. This was particular important in the preparation of the final reports, prepared independently by me using neutral language and highlighting both positive and negative aspects of the evaluation findings in the two countries.

Self-reflexivity was also relevant to deal with the challenges associated with the cultural differences and geographical distance between the evaluation teams. In this regard, it is more difficult to self-assess my performance as evaluator. My perception in this regard is that the parallel activities described in chapter six contributed to gradually build a solid working relationship with the Pakistani partners. To this success, however, the involvement of the British-Pakistani member of the LWC team in these activities seem to have been fundamental. His role was also crucial in all the communications between the partners in the AIM evaluation. In the case of the THRIVE evaluation, this question was not as significant, with my relation with the Zimbabwe-based team strengthening over time without the need of intervention of a ‘gatekeeper’.

Other important question to analyse, as it implied as well advantages and disadvantages from the perspective of the impact evaluation and the PhD research, is my consultant role in a project commissioned by an external funder of the MFIs. This advisory role as thoroughly described throughout the thesis, meant that almost all decisions in regard to the project were ultimately taken by the commissioner, even if in many of these decisions the LWC team has followed my recommendations. The resources constraints, which were significant if the project is compared with others mentioned in the thesis, including the AIMS, Imp-Act and ILO projects, played a decisive role in the project decision-making process.
Beyond the considerations on the role of LWC as decision-maker, it should also not be forgotten that the active role of the LWC team in almost all stages of the project had positive and potentially negative implications as discussed in section 8.2.2 (p.203).

One of the main results of the implementation of the LWC impact project was the overall positive assessment of the project processes and methodologies by the LWC team. They are committed to the continuation of the project and its extension to other countries applying the same approach, with the adjustments arising from the lessons learnt during the project, as expressed by his most experienced member in the final interview (LWC1, p.c., June 4, 2018). A positive assessment of their participation in the project was transmitted also by the managers of the two MFIs in their interviews, who are also willing to continue the evaluations, and there were manifestations of interest of other LWC partners in initiating similar projects.

The PhD research, by identifying the positive aspects of the LWC impact assessment project and, simultaneously, highlighting the limitations associated with the approach followed in terms of its methodological weaknesses and the conditions required to the implementation of a common project approach, makes a contribution to what should be a broader effort within the microfinance sector to capitalise on all the experiences committed to conduct evaluations of the programmes which seriously attempt to balance principles of usefulness and rigour.

These evaluations and their results should be given credit, even if they will not achieve the quality standards associated with rigorous academic research. This valorisation will be especially welcomed in the expectation that the continuity of these processes will allow for the adoption in the following evaluation stages of other methodologies and approaches, and that it will encourage other MFIs who might have been initially absent from the process to consider the benefits of evaluation.

This process, which should not forget previous experiences such as the Imp-Act and AIMS projects mentioned above, can accelerate the development of social performance management and evaluative thinking in the microfinance sector, further pushing the work of organisations such as the Social Performance Task Force (SPTF).

Finally, within the strengths of the research is also important to refer to the establishment of bridges with other research fields, including evaluation and development studies, and how this broader perspective add value to the analysis of the impact evaluations implemented within the microfinance sector. In the thesis, the focus was on the utilisation-focused evaluation model and the utility of its pragmatic approach and learning and improvement
perspectives to look into the design and implementation decisions made within the LWC project. However, it is not difficult to perceive the attractiveness of evaluations models such as the Guba and Lincoln’s (1989) constructivist model or Pawson and Tilley’s (1997) realistic evaluation to conduct impact assessments of microfinance programmes in which it is privileged a more in-depth analysis of the programmes, evaluating the mechanisms of change associated with the interventions and taking a dynamic and complex approach to the social impact of the programmes at different levels. It is, although, essential to bear in mind that these approaches imply the implementation of different methodologies or combinations of methodologies (mixed methods) and require resources (financial, human, time) in much larger scale than the available within the LWC project.

9.5 Final considerations

The hype around microcredit and its potential to reduce poverty on large scale have significantly diminished, but this development may turn to be positive for the sector allowing for a more realistic definition of its overall goals and a recognition that microfinance programmes are designed, implemented, and should be judged, within broader social interventions, in which different private and public actors interact. These considerations have consequences for the assessment of the interventions implemented within the sector, requiring a pluralistic approach also in the evaluation of the programmes which promotes and valorises different types of evaluations, with distinct objectives, and involving evaluators/researchers and programme stakeholders in different degrees.

From the perspective of the LWC project, the implementation of a third wave in the two initial countries and the extension in the short-term of the project to one or two additional countries will allow for testing the hypothesis suggested by Schofield (2000) regarding the increase of the external validity of the findings as the project is replicated in multiple sites. This multiple site perspective will also enrich the knowledge of the sector by providing additional information on the microfinance programmes and their outcomes in different contexts, and the broader knowledge on evaluation by documenting the challenges and solutions implemented. It will also allow to deepen the analysis on the common project approach and its limitations and advantages in contributing to motivate more MFIs to implement social performance assessment and management systems, complementing in this way other initiatives in the sector.
References


231


Microcredit on Women’s Control over Household Spending in Developing Countries: A Systematic Review and Meta-analysis (Campbell Systematic Reviews 2014:8). DOI: 10.4073/csr.2014.8


Annexes

List of Annexes

Annex I – Transcripts Interviews Evaluation Team
   I.I Recorded Interview LWC1
   I.II Recorded Interview LWC2
   I.III Recorded Interview AIM1
   I.IV Recorded Interview AIM2
   I.V Written Interview THR1
   I.VI Written Interview THR2

Annex II - Questionnaires Household Survey (Baseline)
   II.I AIM Clients 2015
   II.II Pakistan PPI Look-up table
   II.III THRIVE Non-Clients 2016
   II.IV THRIVE-PAT Form

Annex III – Questionnaires Household Survey (Second Round)
   III.I AIM Non-Clients 2017
   III.II THRIVE Clients 2017

Annex IV – Evaluation Reports (2017-18)
   IV.I LWC/AIM Final Report
   IV.II THRIVE Final Report

Annex V – Ethics Forms
   V.I AIM Ethics Commission Letter of Approval
   V.II THRIVE Ethics Commission Letter of Approval
   V.III UPR 16

Note: The questionnaires for AIM Non-Clients 2015, THRIVE Clients 2016, AIM Clients 2017 and THRIVE Non-Clients 2017, as well as AIM and THRIVE Summary Reports were not included in the thesis annexes for space limitations. They are available upon request via email (joana.silvaafonso@myport.ac.uk)
I.I Recorded Interview LWC1

4 June 2018

Part 1

Interviewer (I): Good afternoon. This interview is with a member of the evaluation team from LWC. The request for a recorded interview as well as the list of questions were sent in advance to the respondent and he kindly accepted. Thank you for your collaboration. I just would like to reinforce that you have the right to not answer any of the questions or withdraw from the interview; that the recording will be stored in the University of Portsmouth server and only people related to the PhD have access to it; and that your name will not be used in any external publication. If everything is OK, can we start?

I: What was the origin of the LWC impact assessment and what were the initial steps to implement it?

LWC1: The origin of the LWC impact assessment... So, essentially the purpose of LWC is to improve the lives of poor people; as practitioners, we often are more involved in the actual implementation, but we don’t step back enough and see whether the work that we are doing is having positive impact or not. So, we had originally thought of doing the evaluation process ourselves, but we thought that we could potentially be criticised for not being impartial in terms of the evaluation that we did. So, what we decided to do was to look for academic collaboration, and that’s how the partnership with the University of Portsmouth, and particularly yourself ... that’s the reason. So, essentially the purpose was to see (if) what we are doing is having a positive impact or not in the lives of poor people. I mean that’s the origin, the reasons for the evaluation taking place.

We do a lot of ... it could be called focus groups, it could be called direct interviews. We do that already, we’ve done that since the project started in 2010. So, what we do is we use proxy data. We collect financial and social indicators on 18 different areas from each of our partners on a quarterly basis and we look at that data, and that gives us an insight into what scale, what outreach and we infer some sort of impact, whether they’re positive or negative from those 18 indicators. So, that’s a useful tool that we use already. The second thing is that we have a series of regular field visits to each of our partners. Each of our partners gets evaluated at least once a year; on some occasions, it’s even as frequent as twice a year. So, in these visits, from the outset, we speak with borrowers, we speak with programme staff,
we speak with the organisation and we get sort of an idea of what assistance our project is providing and whether it’s been beneficial or not for our partners. But, again, that’s a useful indicator for us, but perhaps one can say is not sufficient. So, we use already a couple of methods, and lately what we started doing as well, and will be doing this in a more regularly basis going forward is that we ask for ratings. Some of our partners already get quite comprehensive ratings undertaken by microfinance rating agencies. What we are doing more recently is that we are having social ratings undertaken as well by a qualified rating agency. So, we can have a good idea of where our partners strengths are, where their weaknesses are, whether the work we’ve been doing has a significant social impact or not.

So, these are all different methods that we are already using. So, but what we thought was that we didn’t have a specific study looking specifically at the people who we support with our funding and how their situation - economic, business, household, is improving or not over the course of our assistance. So, that’s where the assessment with the University of Portsmouth arose and that’s why we are doing it. I guess we have several different methods looking at the same sorts of questions, each giving us an answer which taken in context, we can probably triangulate a lot of the findings and say either we are having a good impact or not. But, certainly of all the methods that we have, we think this is the one that’s most comprehensive, most robust. So, that’s kind of the origins, the answer to the first question.

I: In the process, you have adopted a kind of “goal-free” approach. So, no programme theory was discussed before with the partners, it was a more ‘open’ evaluation. And you have decided to use a client household survey and use a poverty scorecard, specifically the PPI. What were the criteria for these or was there a specific reason to use this kind of approach?

LWC1: I think we say goal-free approach... I think in the back of our mind, maybe it has not been spelled out in the research, in our discussions but I think intrinsically we have a theory in our own minds of what we expect our assistance is providing. So, our programme theory, theory of change or whatever you may wish to call it - in our own minds is that we are helping microfinance institutions increase their scale of outreach, we’re enabling them to work with perhaps more vulnerable communities, we’re enabling them to work in more geographically isolated areas and through financial inclusion, often with a significant element of training as well and other services, we are hoping that the people we support - their income increased, their survival strategies are stronger, their assets have built up, etc. All the kind of positive things that you might associate with many development interventions. However, we didn’t spell it out, either in the discussions with the University of Portsmouth or in the design of the service. But, I think that as a development practitioner that’s what I’m hoping to see. However, we thought it’d be better not defining a lot of things beforehand. We wanted it to
be more open because eventually we want to extend this research to all our partners. We currently work with 14 partners. So, we started this research with just two of them. So, we thought that there would be a lot of things that might be learnt in this not pilot (because it isn’t a pilot), but in this initial collaboration that we could probably take on and adopt in future evaluations. So, you’re right we were more open in what we thought that might be seen, we could possibly see, we could probably adjust and adapt to as we go along.

I: Relatively to the poverty scorecards, was there any specific objective with that tool or was it a situation of being a tool recommended and used by many international agencies in many evaluation studies?

LWC1: Since it’s the industry standard for microfinance, we thought it would be useful to include that as well. So, it’s the industry standard and we wanted to be able to say where do our partners fall in terms of the industry standard as well. And also, it allow us to compare and certainly allow our partners to compare perhaps with peer organisations in the context of the countries where they work as well. One of the origins, going back to the motivations, is not just to get results for ourselves and is not just to get results for ... with standards for academic analysis as well, it’s also helping our partners to begin the path for themselves of starting to evaluate their work as well, whether it’s themselves or through third parties or through investors such as ourselves because neither of our partners really, the ones we begun with in Zimbabwe and Pakistan, neither really did anything, which is surprising. I think one of them is the largest Islamic microfinance organisation in the world, it has an outstanding loan portfolio of $135 million, more than 900,000 active clients as we speak so nothing really was being done in terms of evaluation. I think the other thing for us is that the research has a use in itself, the research has a use in terms of giving us results that we can disseminate to a wider audience, almost to justify our intervention. But the third thing is that the research begins a process within our partners as well. Which I think sometimes is kind of overlooked or not taken into account. I think our partners are now beginning to take the issue of evaluation research much more seriously as well.

I: So, that was among the expectations that you had for the project as well?

LWC1: Yeah, when we began LWC is just about helping poor people, is about helping our partners to increase their scale and outreach. One of the things that we were always concerned about is what sort of technical expertise can we provide and I think right at the beginning our partners said to us: look one of the areas that we are really weak is that we do not evaluate our work, we are probably weak in terms of social performance management, so for a number of years now we’ve always been thinking that as well as providing extra
capital to fund their operations, what other areas can we work in? So, there’s been a range
of areas that we’ve been working with them over the years and this is one of them.

I: In terms of the implementation, what were the main challenges of ‘putting’ the project into
the field and developing it?

LWC1: The main challenge for us has not been finance, has not been willingness of our
partners, has not been willingness of ourselves, the main challenge for us has been finding
the time. Our partners finding time, and particularly us as a team within LWC finding the
time. I think probably without the assistance of the University of Portsmouth, it simple
wouldn’t have been possible, without your assistance, Joana, it would not have been
possible. So, I guess that’s been the main challenge and I think as we go forward into the
future, I think possibly something that we’ll take away from this is that we need a specific
designated person to be responsible for this, within our team, if we really need, if we really
want to upscale the work. So, that’s probably been the main challenge. I think all the other
things, languages, timings, etc., I think they have been minor challenges and they have been
overcome. The main challenge has really been finding the time, us as a team and to a lesser
extent our partners in the field.

I: The project was implemented in 3 different continents, with an evaluation team that had
members with very different characteristics and cultural backgrounds, and professional
backgrounds as well. How do you think this diversity influenced the project itself?

LWC1: I think the diversity probably enriches the process if anything. I think we chose these
specific countries where we work for reasons to do with facilitating the research, so there
wasn’t a random selection of the countries or the partners that we work with. I think that
very specifically we’ve chosen them because English is widely spoken in both Pakistan and
Zimbabwe. We chose them because our partners were very, very willing to participate in the
research and helped us a lot. The diversity … I think was useful in terms of that we have one
interest-based microfinance organisation that works in very much the way that we expect
most microfinance institutions to work. I think the other partner, the fact that it charges no
interest and it’s an Islamic organisation … that in itself was very, very … interesting. I think
that’s a potential area for researchers in the future, the faith-based element in terms of the
fact that it’s not just age, gender, ethnicity, language, academic and professional experience.
I think between all of us, I think we probably got all of these areas covered. We got the
different faiths, we got the different genders, the languages, the proficiencies. I mean
probably, I might be able to have a particular insight into Pakistan for example, so essentially,
I am British and also Pakistani.
I: Do you think that was important in the process?

LWC1: I think so. I think is important particularly in Pakistan ... I think sometimes the MFIs find it difficult to communicate with researchers from outside. Once they identify the researchers have particular assets or insights, I think that facilitates the collaboration for them. I mean on loads of occasions, the team in Pakistan would say to me: look (LWC1) how do we translate this particular word, what is it that the researchers are trying to understand by this, you understand them because you resemble them, you are both at the same time. So, you understand our perspective and you understand their perspective as well. In terms of interpreting the results possibly, but also getting the buy-in from our partners. I think sometimes with research the difficulty is the buy-in from the partners in the field. I think in this occasion because of the unique relationship that LWC has with our partners, and also perhaps I have with at least one of these in Pakistan. I think that’s facilitated the research, probably the insight that we can get.

I: Well, that links with the following question. During the project, there was a big effort from all partners in terms of communication, there were several partner meetings and bilateral meetings between the partners. At the same time, there were, specially with Pakistan, a number of parallel activities like the visit of Dr. Amjad to Portsmouth, the chapter in the book or the session in the research conference. Do you think these parallel activities helped as well in building this relationship?

LWC1: I think that undoubtedly helped because you have the organisation say in Pakistan or in Zimbabwe whole heartedly committed themselves to the research, there’s been analysis and examination that’s been happening on a whole number of levels. It’s not just the level of your research that has been going on, I mean I’ve spoken at least at 2 international conferences in Dubai and in Istanbul on the research as well. I’ve got feedback from conference participants. We have regular meetings with the AIM Chief Executive and with the other partner from Zimbabwe here in London. So, I think that the personal factor has probably enriched the research greatly. I can’t think that it’s hindered or hampered or taken anything away from the research. I think only that it has added to it. So, we’ve got insight from a whole host of different levels, a whole host of different forums and methods and we’ve been evaluating, we’ve been writing, we’ve been presenting at conferences and that’s been happening in a whole range of levels.

I: How does LWC see the project findings and how do you expect the partners to use both the results from the surveys and the whole implementation process, what can be learnt from the implementation process?
LWC1: So, we got the results from Zimbabwe and from Pakistan. I think that looking at them both, I think on the face of it, both sets of results are positive. I think they were not as positive in Zimbabwe, but that’s probably largely due to the macroeconomic and political environment that’s been afflicting that country for the last few years. And, in fact, you could probably say, despite everything, the results in Zimbabwe are not negative.

First of all, how do we use the results... essentially, we use the results to disseminate to our funders. So, all the many tens of thousands of people that support LWC, the results are communicated to them because the question that they most often ask is: is our financial assistance having a positive effect on the people or not?. So, now we are in a position to answer that question. We can say look here are the results, done impartially, with a certain level of academic rigour and these are the results. So, we disseminated first and foremost to our supporters. Secondly, we have a smaller number of corporate donors and institutions that support us as well. We can disseminate the results to them. So, they look at the results as well. I think what has been reassuring to all of them is the fact that they’ve been done by a University researcher and they are more impartial. So, that’s the feedback that we’re getting. It’s well done, you didn’t just do it yourselves, you got somebody else to oversee the process and ... there’s more trust in them because they’ve been done in that particular way. So, that’s how we use the results.

How do our partners use them? Again, they disseminate it to their investors and supporters. I already know that they’re disseminating these results. But the second thing is that we want to see what processes, what possibilities are there for both, us and for our partners, to integrate the way that we have been doing the research into their normal operations. So, I think that’s been really useful as well. I think the research generally has been of huge practical use for our partners and for us as well. What is possible, what isn’t possible, what’s feasible, what isn’t feasible? What are the challenges, what are the weaknesses? And how can we incorporate this into our day to day routine work? So, that’s what we are trying to learn from this process as well. I’m not sure we are at the stage where they’re looking at the results in such detail and saying look we’re going to have to tweak our programmes because they’re not probably reaching the people that they should reach. Certainly, they’re looking at the poverty scorecards and seeing where they stand compared to peer organisations. But I think that hopefully will be a step further down the line where they look at the findings and start tweaking their own programmes accordingly.

The other thing is that we’re hoping to extend the research to other countries now. So, we’ve looked at the initial experience in a couple of countries and we’re looking to extend it. So, we wouldn’t have got involved with the research if there wasn’t a practical use for it for us or for
our partners. Once we were clear from the beginning that this is of practical use to all parties involved, it’s not just an isolated piece of academic research, that’s why we got involved. So, it’s not the first request that we’ve had from universities, and I’m sure it won’t be the last from researchers. But we have to be clear that there has to be real benefit for our partners and for ourselves before we get involved, and in this research there has been.

I: So, the main strengths and main limitations or challenges?

LWC1: The main weaknesses first of all is the fact that we think that the control groups could have been larger and we should have made more effort to chase them for the second round of interviews. That is probably something we need to address when we extend this to the next countries. The main strength is that there was a lot of buy-in from everybody, it’s stuff that we can use to disseminate to our partners and provide evidence on the impact of our programmes or not. I guess that’s the main strength. So, we can turn around to anybody - donors, researchers, whoever, and say look our work has been independently evaluated to a degree, this is the result, so it’s providing robust evidence.

I: So, In the process, or in the way the evaluation has been implemented and designed, the focus has been on the decision-makers, in terms of the interests and the objectives of the decision-makers, which are LWC in one side and the MFIs in the other, but there are other stakeholders of the microcredit programmes that were involved, namely field staff and the clients that were part of the household survey. What do you think is the role of these stakeholders in the process?

LWC1: I think that ultimately we do our work to benefit those who are the clients. That’s the ultimate goal of our work. I think if the results of this research just rest with us, the decision-makers, it wouldn’t consider this a success. So, I think what we’re trying to do is to say well, either we’re reaching our objective of helping poor people or not. I think the fact is we’re trying to communicate through our partners to ensure that ultimately these people benefit. I mean, our partners are well aware of this as well. I think, they do this not for an end in itself but to improve the lives of poor people. So, I think from discussions certainly in Pakistan, the feedback that I’ve been getting is that almost for the first time the field staff and a lot of the clients were saying “this is actually the first time that anybody has asked us about this question”. So, they are involved, and they feel that they’ve been listened to. I guess the situation is a little bit different in Zimbabwe where they have more communication and focus groups with a lot of the people. So, I see that their role is certainly that they’ve been listened to and hopefully we can communicate to them the fact that their wishes and their feedback has been acted upon.
I: This kind of project and any kind of evaluation always implies a trade-off between being completely external, independent and accountability driven, and the other side that is more focused on the internal use of the process and the improvement of the programmes, and with a much higher degree of involvement of the partners in the field. And, of course, all the resources and the limitations of resources always play a role in the balance between those two. How does LWC see this trade-off in the context of the project and the decisions that were made throughout the project in terms of implementation?

LWC1: yeah, I think we have a fairly good balance. I think in the beginning we wanted to do two things. In the beginning we said look we need to provide evidence to our lenders, our supporters that LWC is having a positive impact or a negative impact, what is the evidence? We were prepared to accept negative evidence as well. We didn’t go into it with preconceived ideas. We thought we are running a risk, if the research come back and say look your work is not particularly good, I mean that was a result that we were ready to accept as well. So, we first of all wanted to show evidence to our supporters and lenders. At the same time, we have this mission of providing technical assistance. So, LWC is not simply a crowdfunding platform. When someone asks me to define what is that LWC do, I say it’s support our partner microfinance organisations with extra funding and capital, so they can extend their outreach to vulnerable communities and geographically isolated communities, but also we provide technical assistance to our partners to help them reach this goal. So, this assistance could be in terms of social performance management, could be in terms of providing access to markets overseas, could be... recently we had a case where we’re supporting one of our partners with the purchase of new microfinance software, so we’re paying for that ... We do HR, we provide legal assistance, so there’s all different areas that we provide them with assistance. So, this research was also in terms of helping them to integrate better social performance management procedures into their routine operations. So, there were these two goals that were behind this research. So, when we say we had a trade-off, I’m not so sure of a trade-off, we just had two goals and we thought this way that we’ve done it we reached actually a fairly good balance between the two. We’ve not gone more for one or for the other. We’ve kind have done both.

And I think the other thing is that for our level of provision of evidence ... because we have so many different areas we’re looking at in the performance, it’s not just this. So, if we said look this is the only method that we’re using to assess our impact, maybe we could have said ok we will increase the sample size, maybe we will increase the control group, maybe we will go for some sort of randomised control sample, maybe we could have done a whole host of things, but it’s not the only method that we use. We have several other methods that also
provide us with insight and feedback of whether we are achieving our mission. So, we use this as one of the methods, maybe it’s the one that in the context of Pakistan and Zimbabwe has greater weight, but nevertheless is just one of the methods. So, we look at the other methods as well. I think that’s important for us as well. If it was just one, the only method, maybe we would have done a little bit more. But I think that within the parameters, within the restrictions, I think we did incredibly well, and that’s largely due to you in terms of the results that we’ve got.

But our audience is different as well, our audience is not an academic audience, our audience requires more than 50% probability. So, if on the basis of likelihood is more likely than not that your intervention is having a positive impact, then we are happy to continue investing. If it was perhaps somebody else that we might be thinking we require greater than 3/4, greater than 90%. That’s not the parameters to which we work if that makes sense. But our investors, a lot of them actually are just content with individual case stories, that might be sufficient for them, but we want to take it a level further. So, the scrutiny, I think maybe that’s the right word, the level of scrutiny of our investors is much less than perhaps an academic audience.

I: Do you think that for the MFIs that probably want to work with the financial investors, it’s the same order of scrutiny?

**LWC1:** I think by and large, it’s the same level of scrutiny. Particularly for Pakistan, for example, where they were by and large the Islamic investors … are not looking for a return. So, that’s an important difference as well. I think that’s the same level of scrutiny.

I: And in terms of the degree of independency, is there also a different level of standards … is the idea of impartiality important?

**LWC1:** It’s hugely important. I think for our partners, a good possibility, a good method would have been to evaluate the projects themselves and come up with the results; and the investors would have looked at it as good but not impartial. The next level up is for one of their funders, another investor to do the research as would be the case for LWC doing the research; this would have been viewed as better but not necessarily the best because an investor has a certain level of commitment to the programme, a level of wanting to see positive results. I think the third level, the best would be to have an academic, University or researcher, doing the work. That would be the best option… for them to be able to point out – look an independent researcher from a university in the UK has done this work for us. Is better than them doing the work, is better than another investor doing the work. So, they can now present these findings to other investors and say look we had our work analysed by
an impartial academic researcher and these are the results. So, for them, it’s something that they’ll value greatly.

I: There was a similar approach in terms of evaluation in the two countries that worked in the project - Pakistan and Zimbabwe - very different countries, very different contexts, very different MFIs and programmes that they are implementing. How do you think this common approach worked? There were naturally adjustments in terms of the implementation, but is it something that is worthwhile and is replicable in other MFIs and contexts or would it be better to just think individually what could be the best approach in terms of the evaluation?

LWC1: I think it’s always best to have a tailored approach depending on the context of the country, the culture, etc. So, you always have to adjust to a very specific context, to very specific factors. But, generally the model that we have of sampling new borrowers and a control group, and then revisiting them a couple of years later, we found that is pretty good, and we’d like to sort of replicate that to other countries and contexts. As I said the only real, significant difference that we’re thinking about doing for the future, once we extend the approach to other countries, is to look at the control group, to increase it significantly and to make sure that the drop-out rates in the second time are not as high as we experienced certainly in Pakistan. Possibly, maybe reward the people for their time as well, in terms of a small payment, particularly those from the control group. That would be the only two things that we would do differently. But I think the basic model that we have looking at new borrowers, comparing them to a control group, works well, has worked well and we think will work well in other circumstances.

I: Thanks, (LWC1). Do you have any question that you want to ask or are you satisfied with the interview?

LWC1: No, that’s it....

Part 2

I: So, coming back ...

LWC1: Yes, I just want to do an extra point here. Obviously CARE is a quite large international development organisation, with programmes in more than 90 countries. We do a lot of evaluation of all our programmes across the different fields and many thousands of staff that we employ. I think the feedback that we’ve been getting from colleagues in different areas, not just in microfinance, but in other areas as education, health, water and sanitation, and other areas, is they’re saying that this is a quite good way that you’ve collaborated with the university. A lot of our colleagues in other CARE offices, either it’s in the US, Australia or some
of the others, they’ve actually asked me for copies of the reports that you wrote and saying this a quite good collaboration that you’ve got and it’s probably a better way, slightly more impartial than the traditional way that we’ve been using. We usually get consultants to do it. By and large, if CARE has a project and it comes to the stage of evaluation, they’ll commission consultants to do the evaluation, and the feedback that we’re getting is that the consultants’ reports are not as rigorous, not as comprehensive, and not necessarily got the same insight that we’re getting from our collaboration. I don’t know whether it’s because the team is slightly different, or we have the insights, or the fact that we’ve collaborated on several different levels, at conferences, in books, etc. has given us extra insight. But, certainly, CARE was looking at it and saying actually this might be a better way to do it.

I: But there were other collaborations with universities before? RCTs, eventually?

LWC1: No, because as far as I know, all of this has been done by and large with consultants. It’s not been with universities. So, the fact that we’re getting such good research in terms of its rigour and be able to (...) analyse throughout our collaboration very cost-effectively. The other thing I think we need to stress is the financial side. We have done this at a very low cost, in terms of the research we do generally in CARE. As far as I am aware, the research that we do through consultants is vastly more expensive and it’s not necessarily as insightful as what we’re getting. Probably in the eyes of many outside people, whether it’s investors or our partners, it doesn’t have the same credibility. So, I think the other thing is that we have a lot of credibility of the research, the way it was done, who’s done it. The consultant’s report doesn’t have the credibility to the same level. Whether rightly or wrongly, I’m not going into methodology or techniques, I’m just saying that who is doing ... for a lot of people, particularly our partners, it seems a lot more credible. That’s an important point that I had forgot to mention.

Notes:

Recording duration: Part 1 - 40:23; Part 2 – 4:24
Interview with LWC Evaluation Team Member took place on 4 June 2018 at CARE office in London
Clean Transcript – no names, no major pauses or repetitions of words
I. Recorded Interview LWC 2

4 June 2018

I: Good afternoon. My name is Joana Afonso and I am a PhD candidate at the University of Portsmouth. My research focus on the challenges of evaluation in the microfinance sector, and the research question has been addressed through two case studies in Pakistan and Zimbabwe where an evaluation project led by LENDWITHCARE has been conducted since 2014-15. This interview is with a member of the evaluation team from LWC and the request for a recorded interview was kindly accepted. So, I would just like to reinforce that you have the right to not answer any of the questions and to withdraw from the interview at any time; that the recording is going to be stored in the University database and only the persons related to the PhD have access to it; and that your name will not be used in any external publication. If everything is OK, can we start?

LWC2: Yes, that’s fine.

I: The first question has to do with the origins of the project. What were your expectations, as member of the LWC team, related to the implementation of this project?

LWC2: So, you are probably aware, for a long time we’ve been kind of wanting to find a much more rigorous way to look at the impact of LWC. LWC has been facilitating loans for nearly 8 years, by the time we came to the partnership with the University of Portsmouth, that was 5 years and we really wanted … it’s a project that landed in our lap but something that we really wanted to do for a while... where we could instead of just collecting case studies on a bit of an ad hoc way, this was an opportunity for us to in a much more rigorous way collect data on the impact. So, for the team it was what is the impact of our loans on the poverty levels of the clients and their households, something that we hoped to address and expected to kind of achieve through this project.

I: In the implementation, you have worked with two different countries and two different MFIs, what were the major challenges for you, as part of the evaluation team, in the implementation. And at the same time, what were also the strengths, the ‘good points’, the advantages that the project had?

LWC2: I’ll start with the strengths because I think that both of the partners, which was probably part of the reason why we worked with them, were both very keen to take part in this project and though very different in the way they module their microfinance modules, the fact that our partners were keen made it quite easy to an extent. They were very accessible, so both of the contact points in country were very quick to reply, keen to take
part so that part of the process working from the UK was a positive experience because we found them eager to take part and responsive to our requests.

But then again on the flip side of that is that... all the development of the questionnaire was, for example, happening between three different parties, all placed in different parts and different countries. I guess that was probably a little bit of a challenge, trying to understand, more for the University than for us, try to put together those initial questionnaires that also kind of met the requirements of different people, so obviously LWC had a kind of objectives, the microfinance institution also had objectives that need to be put together and the University of Portsmouth had objectives as well, so just kind of marrying that all together.

Other challenges around the implementation, I mean, there has been a bit of a challenge around, more so in Pakistan from my perspective, around ensuring that there was enough emphasis on the impact on economic empowerment, women’s economic empowerment. That was a slight challenge I think, more internally, and that is something that we’ve learned over the process. As CARE International, women’s economic empowerment is a huge part of our work, so ensuring that the work that you did around the impact measurement looked at that enough was a challenge. And then, the one that everybody knows, obviously, the control group. The control group was really hard to find, to work out how to find them, where to find them, whether we pay for them to take part or we don’t pay for them, and then obviously in the subsequent round of interviews tracking them down was a big challenge.

Personally ... another sort of issue ... my personal experience was inputting the data once we’ve got it, so I was working here to try to find the best way to input the data once we’ve collected it. For the first time we use, no the second time we used like an agency which then was a bit of a challenge in terms of analysing the data because the way they input it wasn’t necessarily ... they kind of missed some of the nuances in the answers. The first time around, the microfinance institution was inputting the data so obviously a few kind of concerns around that I guess in terms of the objectivity of entering the data themselves, but overall I found actually it was a very positive experience and I think that is largely due to the will of everybody who was taking part for it to work.

I: You had a hiatus of time that you weren’t working in the project because you had another personal project. You were here at the beginning when everything started, the first implementations and then you came back one year after. Did you find differences in the process, in the partners ... in the flow of the project?

LWC2: Not really. I’m trying to think when ... I came back when we were inputting the data for the second round from Zimbabwe. I entirely missed the second round for Pakistan, so I
only have been around for the results ... for me, there’s no noticeable change. And then Zimbabwe ... No, I haven’t notice anything. I think when I came back we were learning the lessons from the first time that we entered the data so, no, I can’t really think of anything that was completely different.

I: And, in regard to the results and the way that LWC and the MFIs are seeing the results and planning to use them, what is your perception on the findings and what we’ve got so far?

LWC2: My perception is that the findings have been really interesting. I think we were all pleasantly surprised and pleased by some of the changes we seem to be seeing in Pakistan over a relatively short period of time, and it’s been something that we have been sharing internally at CARE, with our monitoring and evaluation advisers and experts. And, then again, Zimbabwe has been really useful insight to see potentially not so positive changes and potential reasons why that might be happening. I think for us the next round is really, really, important to see I guess if it was just an anomaly - why did that happen, or is it continuing to happen in the same fashion, or are we seeing something completely different this time? Are we able to find the control group at all this time? And I think probably after that round, then it is really the time for us to start asking questions about what we want to do with these learnings. Obviously, we want to continue seeing it and is still a short period of time, but for CARE we need to think about, especially in Zimbabwe, if we are seeing potentially not positive changes happening, what does that mean for CARE as an investor in these microfinance institutions. And so, for us, I think that’s the learnings that we take from the results of the study and we need to work, I think we’re going to have to work closely with the team, the monitoring and evaluation team, to work out what’s the best route. And also, I’d really like to take the opportunity to kind of dig a bit deeper into some of the things that come out of the impact study, so have the opportunity for focus groups possibly, the opportunity to delve into that question of women’s economic empowerment a bit further, especially in Pakistan. So, I think those are things we really need to start focusing on. We got a good kind of first look of some of the things that might be happening and now we need to think what we’ll be doing with that learning and the actions that come as result of it.

I: Is it (evaluation project) something ... do you find it will be replicable in other partners of LWC or was it too ‘scary’ and kind of put you off ...?

LWC2: No, definitely I think it is replicable. I know that some of our partners are already very keen and have learnt from the ones that have already done it what’s involved, and I do think we probably... I mean the issue of the control group is always going to be there, but we might have to think about that a bit more deeply if we do replicate it in another partner. Because
that’s something that has come up here quite a few times when we were sharing it internally; is about the control group and the strength of the comparison group basically. So, I think yes, it’s replicable and I think, I know that people want to do it, it’s just kind of … making sure that we strengthen it each time, not just simply replicating it. And, for me, strengthening it would be more incorporation around the women’s economic empowerment and this issue of the control group.

I: Last question, personal experience. Was it a positive experience in terms of your future career and your development to participate and be part of the team?

LWC2: Yeah, absolutely. I haven’t had much experience, if any at all, about doing this sort of thing before. We kind of adjust … there was no strict way of monitoring social impact for us, we would have interviews and we collected all that information over the years but there was no kind of structure to how we were collecting it and what we were looking for, so I have learned loads and I am actually going to do a monitoring and evaluation course in July with Bond to try and learn a bit more practical skill around that. And as a team, we are investing more money into impact, measuring and monitoring impact, scaling up that side of our work. So, definitely it’s been a useful experience for everybody. I’ve learnt a lot.

I: do you have any question?

LWC2: No.

I: Thank you again for your collaboration.

Notes:
Recording duration: 13:27
Interview with LWC Evaluation Team Member took place on 4 June 2018 at CARE office in London
Clean Transcript – no names, no major pauses or repetitions of words
I.III Recorded Interview AIM1

22 March 2018

Interviewer (I): Good afternoon. My name is Joana Afonso and I am a PhD candidate at the University of Portsmouth. My research focus on the challenges of evaluation in the microfinance sector, and the research question has been addressed through two case studies in Pakistan and Zimbabwe where an evaluation project led by LENDWITHCARE has been developed since 2015. This interview is with a member of the evaluation team involved in the project representative of the institution’s management. The request for a recorded interview as well as the list of questions to be asked were sent in advance to the respondent and he kindly accepted.

Thank you for your collaboration. I would like to reinforce that you have the right to not answer any of the questions or withdraw from the interview at any moment; that the interview recording will be stored in the University of Portsmouth server and only those related to the PhD, namely the supervisors and the examiners will have access to the data; and that your name will not be used in any external publication. Finally, I would like to add that I will send you the relevant chapter of the thesis before the submission for your review. If everything is OK with you, can we start the interview?

AIM1: Yes, I do not have any objection. You can proceed.

I: Starting by the expectations and the objectives of the institution. What were the main objectives of Akhuwat by participating in LWC project?

AIM1: Actually, this is really interesting. Akhuwat is a unique model where we provide interest-free loans all over the Pakistan. We just started in 2001 and now we are sitting in 2018. So, people need some evidence. They always talk about the financial sustainability of this kind of interest-free microfinance model. Some people from different schools of thought like academia, researchers, PhD scholars, microfinance professionals all over the world, they come to us and they always ask the same question: either it is a financial sustainable model? So, we said yes, this is a sustainable model, but they don’t believe. They say that OK, but this model is only suitable for the Islamic countries and you cannot apply it in other parts of the world. So, we said OK we need some evidence, but that evidence should be produced by some external party instead of Akhuwat itself. So, that’s why this is a key thing for us, that we should need some information which is prepared by reliable and professional people, so that the world could understand the importance and the existence of Akhuwat.
I: Did the objectives, or the way that you were expecting to use the evaluation, changed during these 3 years?

**AIM1:** That is both actually. Because we also need to understand the real impact (of what) we are doing for the last 18 years. And we need this evidence, not only for the researchers and external parties, but for the staff also and for the board of directors - that we are here on a very strong social mission. So, whether we are achieving our social mission or not, whether we have been distracted, whether we are going on the same way, the same path which we thought at the initial stage of this organisation. So, this is a mixture of both things. This is useful for our understanding firstly that this will be a motivational factor for me if I say that we are going in the right direction as for the other persons who are external parties of Akhuwat. So, this was a mixture of that.

I: The results were quite positive overall for the institution. Did these results correspond to what you have expected? Were there surprises in the results?

**AIM1:** Definitely, we are thinking that the results are quite satisfactory right now, but we want more improvement than that. So, we are waiting for the end of this research. Once this research will end, then we will be in a better position to understand that either we are close to our target or not. So far, the results are quite satisfactory for us.

I: And, in terms of implementation of the project, what were the main challenges for the institution? Were there costs that you were not expecting (monetary or not-monetary)?

**AIM1:** We were not worried about costs on the first instance. We were just worried about the implementation. How we will implement this impact evaluation because we have to hire some external people for the questionnaires, we had to interact with the professional people like you, the universities, we have to get data from the clients and we have to deal with the control group also. These were all challenges that we faced during the impact assessment, but I am so much happy that everything was settled, we overcome all these challenges, some people helped us, and we have been able to get this project implemented.

I: Did it took more time from your internal staff than you were expecting? Or was it something that you were already prepared to?

**AIM1:** We were not worried about the time and the cost. Because we had some mission. So, we want this report. It will be very useful for us. So, we decided that no matter either human resources are involved with it. No matter, either the cost that will be there. No matter a particular guy will be recruited only for this purpose. So, we initially decided that we have to do this, and we will go to any place for this.
I: And, in terms of the implementation, what do you think could have been done differently, or could have been done with more support in the different stages of the process?

AIM1: I think that there is only one problem which we faced when we were dealing with the control group. We should have some understanding about the control group, we should have some data and record for the control group. If they migrate from one place to another place without telling us, so what would happen? How do we trace them? So, this was the challenge. As for us remaining things are concerned for this impact assessment, that was really fantastically designed by the University of Portsmouth professionals and Lendwithcare team. They were always so much helpful and supportive, and they guided us, and took our suggestions all the way, so it was a mutual understanding for all of us. I do not think that at any stage of time we thought that was going to be difficult for Akhuwat to implement this kind of research. It was not like that.

I: In terms of reports, there was an initial report in 2015 and then a second report (basically an improved version of the first) and now the final report. Was this OK for you in terms of the information that you wanted? Or do you think it should be more reports, less reports...

AIM1: It depends upon the set of questions that was developed before the start of this impact assessment. So, we were expecting the same things, we were not expecting anything else. Whatever you put in it you will get the same results, so we were definitely prepared for that. We are asking these set of questions from the clients and this will be the results in these types of set of questions. So, we were ready for that. We were not surprised to it.

I: So, in terms of the results, now that we have (the results) from the second survey, how do you expect the institution is going to share the survey and use these results?

AIM1: Definitely, our main objective was to get evidence and show to the people, so we would love to share as soon as possible with government officials because we have so much dealing and so much funding from the government department, so will share these results with the government, we will share these results of the impact assessment with international forum. We will share these results with academia, researchers, scholars, PhD students. If anybody will come to us and ask the same thing and the same question, we definitely share these things with them.

I: Were there already some decisions in terms of your organisation that you have taken also because of the results or because of what you have learned from the project?

AIM1: Actually, as far as the results of this impact assessment are concerned, we did not think to change any system, we have not taken any decision, because the results are so much
satisfactory for us right now. What we will do is just try to wait for the end results, the final report and after the final report, we will present to the board. And then we will be in a better position to understand also where we should go. So, we will wait for that results.

I: One of the evaluation tools that was used in the project was PPI. I think this was not used in the institution before. How did the institution see the instrument? Is it something that you are going to continue to apply in the future? Is it a useful instrument?

AIM1: OK, let me tell you very frankly that in Akhuwat in 2011 and 2012 we used the PPI form. So, PPI form is something which has been obsolete because we do not know who prepared the PPI, we do not know how they get the data for preparing that set types of questions. So, what we are trying to do is that right now we are not using PPI. And we believe that these questions have been obsolete. If someone has a washing machine at his home, does it mean that we should not give them a loan? If someone has a motorcycle at his home, if someone has a toilet at his home, the score goes up.

So it is really difficult for Akhuwat to implement the PPI form, but what we try to do is to develop our own mechanism for evaluation of poverty level, so we are doing this for many years and I do not think the PPI would be suitable for our environment, because you know all over the world the living standards have changed, people are so much familiar with technology, people are so much familiar with other things which can improve their lives, so people are getting better and better, so PPI form has been done in 2008 and we are sitting in 2018, so almost 10 years have been gone. So how you can use that form? And nobody bothered to edit it, so that is why we have decided that we will have our own methodologies, we have our appraisal mechanism to evaluate the poverty assessment of any particular family.

I: Mechanism that you are developing internally with your own evaluation team?

AIM1: Yes, exactly. But you know it is not formally designed, we did not do like a PPI questionnaire that gives us some scorecard. Our appraisal form just gives an understanding about the family structure, our appraisal form just tells either he is deserving or not. We did not evaluate in terms of any score. That is why we are not using PPI.

I: In general terms, was the participation in the project valuable for you personally and for the institution?

AIM1: Definitely, it was so much useful for Akhuwat and myself, and the staff. Because we were also very keen to see the results of where we are going.
I: And for the future, is this process, you have already mentioned some of these things, but the idea of evaluating impact and outcomes is something that is going to be part of your daily management?

AIM1: Yes, definitely. We talked that we are going to a paperless environment, and with the help of this impact assessment, if something comes to our knowledge and if it is useful for Akhuwat, what we will try to do is we just try to implement all this set of questions into a paperless environment where we can get some kind of reports through our MIS. Right now, what we are doing is try to capture all this information into a hard sheet, so it is very difficult for us to get different kinds of reports because we have to enter .... So right now, we are not using such kind of information, although we are keeping all this information into hard sheet, but in the future with the help of such kind of report and for the social impact, we will try to enter all types of information like what is it earned? How it will increase (earnings)? What is the education level of your children? And all these different questions. We will try to enter all such information into the management information system so when there is a subsequent loan we can see the impact of our loan.

I: Thank you for your collaboration. If you have question or if you have any comment that you want to do about the project

AIM1: Thank you so much for this interview, I am so much happy, and I am really thankful to Lendwithcare and, especially, the University of Portsmouth for extending their support to Akhuwat because this is something very incredible not only for Akhuwat, but I believe for the whole world because they will understand the real message of Islam and this interest-free loan concept. I believe you people are definitely helping us, and we are swimming against the tides, so I believe it will be helpful for many people in the future.

Notes:
Interview with Shahzad Akram, Chief Credit Officer of Akhuwat Islamic Microfinance. The interview took place on 22 March 2018 at Hogan Lovells office in London.
Duration interview: 16:00
Clean Transcript – no major pauses or repetitions of words. The name of the interviewee was added at the end of the transcript in response to his request (email 6 July 2018).
22 March 2018

Interviewer (I): Good afternoon. My name is Joana Afonso and I am a PhD candidate at the University of Portsmouth. My research focus on the challenges of evaluation in the microfinance sector, and the research question has been addressed through two case studies in Pakistan and Zimbabwe where an evaluation project led by LENDWITHCARE has been developed since 2015. This interview is with a member of the evaluation team involved in the project at the operational level. The request for a recorded interview, as well as the list of questions, were sent in advance to the respondent and he kindly accepted to do the interview.

Thank you for your collaboration. I would like to reinforce that you have the right to not answer any of the questions or withdraw from the interview at any moment; that the interview recording will be stored in the University of Portsmouth server and only those on the PhD team will have access to the data and finally, I would like to add that I will send you the chapter with your questions before submitting the thesis. If everything is OK and you do not have any other questions, can we start the interview?

AIM2: Yes, we can start.

I: The first question has to do with your role in the project. What were your main responsibilities and tasks during the project?

AIM2: Hello. I am working as Project Manager. My main responsibility was to make all the arrangements for you and Lendwithcare team to visit Pakistan, to arrange all the clients, to communicate to my field staff to set a timing of each client for interview, to talk with the universities to send us some volunteers, to discuss with the volunteers, to guide them about the institution, about Lendwithcare, about University of Portsmouth, about what is impact assessment, why we are doing impact assessment, what are the benefits for all the stakeholders involved in this research and to give them a little detail about how our clients react, what is the qualification of our clients, how to interact with the clients in local language. Like, normally we use English and Urdu languages but to interview the clients we need to be more specific with their language which is normally Punjabi, so this is what I was doing.

I: And, besides the project, what were your all other main responsibilities. You are a Project Manager, right?
AIM2: Yes, I am the Lendwithcare coordinator at the institution. I need to communicate with all the 4 branches of Lendwithcare, to train them, to guide them how to write a better profile, how to take pictures, how to record videos, how to send the details of the clients to me through a proper format, then I need to translate that profiles in English and upload on the website of Lendwithcare.

I had also so many things to do. I assist my Chief Credit Officer in different tasks assigned by our directors, I am working with the social development department, we are developing the MIS of the institution. We are migrating all 17 years data from an Excel based MIS to a web-based MIS and we started a donations campaign. We have developed a mobile app for a quick reporting to senior management and we are introducing cashless system in the institution.

I: Do you have an idea of how much time during these 3 years, how much of your working time was dedicated to the impact assessment?

AIM2: Normally when the project of impact assessment was in progress I needed to put half of my day to this project; like in the first phase we have worked for 2 months, so 1 month was for the impact assessment. Similarly, in the second phase, we have worked for 3 months, so half of my time was spent on that project. And when I put the data into an Excel sheet, it took more time. I needed to dedicate my all days for that project.

I: During the evaluation, we worked with clients from 4 branches. What was the role of the branch managers and the field staff? Did they help in some way in the project?

AIM2: No. They have only had to select the clients and to set a timing of interviews. Like we made groups and our interviewers will visit them. One group will stay at home at 9 am, second at 10, third at 11. They have just facilitated the interviewers who conducted the interviews.

I: So, who did the first contact with the clients to ask them if they were willing to participate?

AIM2: Branch staff.

I: OK. And the reaction of the clients was normally of …

AIM2: For the clients this is not a big issue because before the impact assessment started, we have communicated that we have selected you people to involve in a research, so we will come towards you for the next 3 to 5 years, so they have accepted that at that time. So, we have no issue regarding the interviews of clients as much. But some of the clients were reluctant in the second phase. If they have delayed the payments and we have not given them loans, then such clients were reluctant to give the interview in the second phase.
I: You were also responsible for the interviewers during the surveys and to receive the questionnaires from the interviewers. What were the main challenges to control the quality of the interviews?

AIM2: The big question was to interpret what they have written and what they mean to write. Because there were some questionnaires they have returned and phoned after saying this is wrong, then they are cutting and writing next. And there are some technical mistakes, like earning and expenses are not matched. When I started putting data, 2 years ago they mention I am 40 years old, now the interviewers have written (that) she is 50 years or 60, so these differences create problem for me.

I: But did you talk with the interviewers about these issues or it was something that you noticed after, when introducing the data?

AIM2: After.

I: OK. And you also had to contact non-clients. How did it work in the first contact? How was the process? What were the challenges there?

AIM2: Because in Pakistani culture people don’t share their information to any person at any moment, we tried to contact our clients to please introduce some of your friends and family who are working the same, in same atmosphere, with same earnings to give us an interview. They introduced, and we have interviewed. But in the second phase, I guess about 30 to 40% have taken loans. Because they were not aware about the institution, but when they have given the interview as non-clients, they just started asking, oh please give us a loan, so we had to give them loans. But other persons, which are our non-clients still, they don’t want to give answers to any question, they just ask them to go, they are harsh, they use abusive words, so this was the problem with the non-clients.

I: OK. And this was more on the second phase?

AIM2: Yes. On first phase, there were also problems because when we had collected 100 interviews with non-clients, there were 4 to 5 files which had no information. Then, I had to visit for completing the remaining file and then I have seen messy problems for taking interviews, the people are not serious, they are not willing to share their personal information because they have no benefit, we need to motivate them and even then, we are not 100% sure they are giving true information.

I: And with the clients, were there any issues when you repeated the survey?

AIM2: About the clients, it was not such issue, but as I earlier said, the issue was only with those clients we have not given a second loan because they had not repaid their first loan
quite well. So, they were the problem, they don’t want to answer because they were angry with our staff, with our branch staff and when someone from the institution visited them, they did not want to give any interview, so this was the problem.

I: OK. And, introducing the questions, I imagine it was a very time-consuming task, what were the main challenges there? Were you working alone in that phase?

AIM2: Yeah, it was time-consuming, it took a lot of time and I needed to review every aspect of the questionnaire. If there is anything wrong, they have missed anything, I then needed to call the interviewers who have missed that thing - please kindly can you guide me in what you have written? And, if something is missing in the questionnaire, like the number of families, I need to ask the interviewers and the branch staff why is this missing. And some interpretations were not as explanatory as they should be, so I need to ask the interviewers again and again, and it took a lot of time to enter the data.

I: Was it different the second time?

AIM2: I guess on both times it took a lot of time because everyone have its own perception. Like one interviewer is asking the question and writing the same answer, the other interviewer has another perception, he is getting the answers on different aspects, so this was the problem to manage and maintain all aspects of the questionnaire.

I: OK. Looking back to these 3 years of the project, what do you think were the aspects that could have be done in a different way or could have been done better from all the partners?

AIM2: For the next phase, the biggest problem will be where the non-clients are. We haven’t written their proper address. We have written just their name and their contact numbers. We should write their CNIC. If in future, they have become our clients, we should know! But we haven’t written any CNIC, so we are not aware if they are our clients or not. If the branch staff can recognise them, then we can understand if they are our clients; if the branch staff is not recognising, then this is an issue. We should write their unique ID which is used in our MIS, this is one thing. (Also) proper address and nearest landmark of their house, so we can identify the person; any reference from our client to non-clients, so in future if we want to take the interviews, we should be aware where the clients are! So, this was I guess the most important part.

And similarly, in the case of clients, we should have some type of 1-page agreement - please if you leave, move to another city, kindly share with us, because we have selected you for this big impact; because if you want to leave, you can go anywhere, it’s your right! But if you
can communicate us and share your latest contact number, so we can contact you through telephone and can go through the research till the 5 years. Yeah, this was the problem.

I: And, for you, what was the most valuable or what was useful for you and/or for the institution?

AIM2: I have got an international exposure to talk with you, with the CARE team and I had also the exposure visiting the University of Portsmouth, meeting with the professors, then discussing in a class, so it give a little good exposure to own self and also to the institution’s staff. And, overall, it is giving a good opportunity to present the institution model to all the world.

I: How do you think you will use this experience in the future?

AIM2: We have learned a lot from this exercise and what we have decided is that we are going to a paperless organisation and we will try to use some of the questions of the questionnaire in our MIS to tell the impact. Because before this impact assessment, we have not as comprehensive impact assessment, so now we are going paperless, some of the things we have in hard forms, like - how many family members? How many earnings earned? How much household expenses? How much earnings? How much studying? - we haven’t in the system, so we were not able to analyse any results. But now we are going on paperless, we will introduce some questions from questionnaire and we will analyse this data in the future.

I: Are you going to use, as well, the PPI?

AIM2: We are going to develop our own PPI. We are going to use our own credit ratings, we are going to work on different aspects to measure the impact of our loans in quite better way. This research is the motivation for us to induce the latest techniques of measuring impact at our level, instead of waiting for universities ... We should also be aware, it may not be as correct as research of an independent researcher, but we must have an idea of where we are moving.

I: OK. Thank you for your collaboration. Do you have anything else that you want to say about the project or the questions?

AIM2: Yeah. That’s fine. Thank you so much Joana for selecting the institution and giving us (the) opportunity to do this research. This was wonderful. Hopefully we will complete this next phase next year and we will see good results.

Notes: Interview took place on 22 March 2018 at Hogan Lovells office in London.
Duration interview: 17:20
Clean Transcript – no names, major pauses or repetitions of words.
I.V Written Interview THR1

Received by e-mail on 18 May 2018

Part I – Expectations and Results

1. What were THRIVE objectives by participating in LWC assessment project?

THR1: Thrive wanted to understand what characterises poverty in the context of its operating catchment area and broadly Zimbabwe. Furthermore, in the absence of a standard poverty assessment tool, Thrive wanted the assessment to reveal the extent to which training and micro-credit were impacting livelihoods as an intervention.

2. Did the initial objectives and potential uses of the results changed during the 3 years of implementation of the project?

THR1: No, these remained the same.

3. The survey results are mixed, with signs of deterioration in living standards but also positive signs regarding savings and client resilience. How do you see these results? Did they correspond to the expectations of THRIVE managers?

THR1: We are disappointed with the overall result, but we are not surprised.

Remember, the operating environment in which the test was conducted was very dynamic during the course of the assessment. The baseline was established prior to the introduction of the surrogate currency and the cash crisis that ensued afterwards. Back then, the macro-economic conditions were tough but relatively better than they are now. After the currency anxieties, the difficulties in making international payments, the challenges in making imports, the informal devaluation of the surrogate currency, things got really for most of our borrowers who even at baseline stage where just better than hand to mouth cases. The assessment completed during the heat of the cash crisis during which our borrowers were paying huge premiums to eventually access local cash let alone hard currency to facilitate their business activities. The cash crisis meant the average person on the street was highly illiquid and had minimal disposable income. Most of our borrowers trade on a cash basis and can only thrive if a ready medium of exchange is in the possession of potential buyers. This was far from the case, purchases mainly being done via point of sale machines (which our borrowers could never dream of having) or mobile money which for a while remained unpopular.

Summary-the macro-economic difficulties during the test period set a very low upper limit in terms of what any intervention could do to improve livelihoods.
Part II – Implementation

4. What were the main challenges in the implementation of the project for THRIVE? Were there non-expected costs (monetary and non-monetary) associated with the project?

THR2: The major challenge was in locating the non-clients and the clients that were or are in default.

5. One of the evaluation tools used in the project was the poverty scorecard internally developed. During the project, THRIVE has changed the use given to the tool. Can you describe what changed and why were these changes introduced?

THR2: When the project started the Poverty Assessment Tool was administered to sampled clients. With time Thrive realised that the scores would also assist in assessing the impact the loan had on the clients’ living conditions, though we know that the changes (positive or negative) that happen to clients cannot be attributed to Thrive only as a number of factors play a role. So now the PAT is now being administered to all clients and the data capturing is done on a random sampling of 30% of the clients disbursed at any given period. During the project that is when the Rural PAT was developed.

6. What, from THRIVE perspective, should have been done more, or differently, by the project partners (at all stages of the project, including the flow of communication and reporting)?

THR2: One thing that could have been done more differently was the use of the same interviewers in both phases of the project. This would have made it easier to locate the non-clients. Also, even the interviewers would also notice the change that would have taken place between the initial interview and the follow up interview.

Part III – Use of Implementation Lessons and Evaluation Results

7. How do you expect THRIVE to share and use the survey results and the lessons learnt from the implementation process?

THR1: The simplest is for Lend with Care & the University of Portsmouth to document the findings. The said duo comprises of a reputable development partner and a academic/research institution, your findings are obviously authoritative and can be accepted as credible. The research report can then be shared as widely as possible and the report can refer to Thrive as openly as necessary in so far as it shouldn’t unnecessarily affect reputation and business prospects.
8. In your opinion, was the participation in the project valuable/useful for THRIVE? Why?

THR1: Absolutely, Thrive is a learning organisation, the findings of the study should go a long way in making Thrive more effective we hope by broadening our understanding of poverty. They should also contribute to the general knowledge body, reveal areas of future study and improve future programming by Thrive and anyone else with access to the study.

Thrive’s SPM unit would have benefitted from this exercise as well in terms of handling big researches etc.

9. How does THRIVE Management see the perspective of development of a PPI for Zimbabwe in the near future and its implications in the use of the internally developed instrument (PAT)?

THR1: Thrive is desperate to see the development of a standard PPI that fits the global reporting template. The internally developed PAT is useful but has its own shortcomings. Once a country specific PPI is in place, the internal PAT will be relegated to the Thrive museum.

10. What is your perspective for the future in terms of evaluation of outcomes and its use in the management process at THRIVE?

THR1: For dynamic environments such as those prevalent in Zimbabwe and broadly the poorer regions of the world, evaluation of outcomes should be a continual process not a time-lined event. The politics and as a result the economics are always changing, poor livelihoods respond sharply to these changes and so do outcomes. Continual assessments will keep development practitioners informed about the impact (if any) they are making and whether they should put more effort/resources or re-think interventions.

Notes:
According to the email received from THR1, the questions 4 to 6 were answered by THR2.
Part I – Role in the project

1. What were your main responsibilities and tasks during the evaluation project?

THR2: My main tasks and responsibilities included the following;
- Took part in the recruitment of Interviewers
- Assisted in the training and monitoring of the interviewers
- Conducted the quality check for the questionnaires
- Approving interviewee token of appreciation, Interviewers bus fares and airtime and managing the budget.
- Sending the questionnaires to the UK

2. What other main responsibilities and tasks did you have as Social Performance Officer? Did this change over the 3 years of the project?

THR2: Generally, my main tasks were:
- Data capturing of the PAT, client satisfaction, exit data and complaints compilation and data analysis and report writing.
- Chairing the Steering Committee
- Organising non-financial trainings
- PAT verification
- Training evaluation
- Training new staff on Client protection Principles and social performance

3. Can you estimate the share of your working time dedicated to the project in its different stages?

THR2: It is difficult to say but:
- During the training of the Interviewers I would dedicate 75% of my work time to the study.
- During the first week of the interviews 60% of my working time was also dedicated to the study.
- This however reduced to between 40-30% as the Interviewers would have grasped the questionnaire and making fewer mistakes.

Part II – Evaluation Implementation

4. The household surveys were implemented with a sample of clients from 2 THRIVE branches. How were the branch managers and other branch staff involved in the evaluation?
THR2: The Branch Managers and the Thrive teams were very supportive to the Impact study team especially in the second phase, they provided directions to the clients’ houses or workplaces as well as times that some of the clients are available.

5. In the selection and initial contact with the new clients, who did the first contact and how was the general reaction of the clients to the invitation?

THR2: I did the initial contact and then later trained the staff members in inviting clients for the interview. The clients were very interested in the exercise because they felt that it was a good thing to be part of an international study.

6. In the first survey, there was a hiatus of time between the interviewers training and the implementation of the questionnaire by some of the interviewers. Did you feel that this hiatus had an impact on the motivation of the interviewers or the quality of their work?

THR2: The hiatus of time between training and actual questionnaire administration; I believe did not affect quality of work negatively because the Interviewers shared their experiences as the interviews were rolling and also I was available to give guidance. What it affected really was interviewing of the same clients twice in some instances.

7. Being one of your roles the coordination of the interviewers during the surveys, what were the main challenges in controlling the quality of the questionnaires in the two surveys?

THR2: In the first phase quality control took place but was not as scheduled like the second phase. Probably the improvement was due to the first phase’s experience.

8. The evaluation included interviewing a comparison group of non-clients with similar characteristics to the institution clients. What were the main challenges to recruit and interview the non-clients in 2015?

THR2: Interviewing clients was easy because they had information about the study which was given to them usually after during their last training session. This was a bit different for non-clients which was not as easy because of the following reasons:

- Some of the non-clients were sceptical about giving details about their family members.
- Other were motivated to the extent of asking to be interviewed when they realised that the respondents were getting a token of appreciation. (though these were avoided because of fear of giving wrong information in the bid to get money only)
A few refused to take the token of appreciation due to their belief system that do not allow one to receive money from strangers.

9. The household survey was repeated in 2017. What were the main issues at the operational level in this repetition (both with clients and non-clients)?

THR2:

- Clients who were in good standing with their payments were readily available for the interviews and some before the interview dates were scheduled would enquire on when the next set of interviews would be.
- Most of the first interviews for clients were conducted at the branches and took less time to get all the interviews done, but the second phase involved visiting some of the interviewees of which not all were readily available or stayed close to each other to reduce the travelling.
- Most of the clients who were in default would not avail themselves for the interviews because some thought they were being called pertaining their non-performing loan.
- With non-clients if clients phones were not available or poorly captured these could not be interviewed unless in few instances where one interviewee knew about them. This was for those clients who were interviewed at shopping centres.

Part III – Overall questions

10. Looking back to these 3 years of project, what from your perspective could have been done more or differently by all the partners?

THR2: Using the same interviewers in all the phases of the research as it would have made it easier in identifying non-clients.

11. In your opinion, was the participation in the project valuable/useful, both for you personally and for THRIVE? Why?

THR2:

- The project was useful to Thrive as it enables Thrive to know the impact that they are having in the lives of their clients.
- The study also showed Thrive that their partner LWC is concerned with the operations of Thrive and would want to assist in furthering the impact it has on its clients and their families.
- As Thrive it has also given us a hope or aspiration to conduct similar studies in the future so that the organisation can measure its impact.
- On a personal level it has widened my perspective on how a big study can be done in practice.
- The participation has also inspired me to pursue my studies as well.

Note: Written answers received by email.
# II.I Questionnaire AIM Clients 2015

## Impact Assessment Form for Lendwithcare

**Date:** 
**Gender:** 
**Age:** 
**Questionnaire No.:** 

**Name of Borrower:** 
**Name of Interviewer:** 

**Sanction No.:** 
**Care ID:** 
**Branch Name:** 

**Number of People in Household:** 
**Contact No.:** 

### Details of Household Members

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Name</th>
<th>Age</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Instructions for Interviewer:

- The questionnaire should be answered by Akhuwat’s client. Were there other people in the room during the interview?
  - a. [ ] No  
  - b. [ ] Yes  
  
  If yes, please identify their relationship with the client ____________________________

- Did they answer questions directed to the client?
  - Yes [ ]  
  - No [ ]

- Length of interview ____________________________

- Is the client willing to be contacted or visited again to complete the questionnaire?
  - Yes [ ]  
  - No [ ]

**Q.1** What is your business and for how long have you owned this business?  
(state in years or months if less than one year)

**Q.2** How did you fund your activities prior to involvement with Akhuwat?  
(respondent can indicate more than one reason, please rank 1, 2, etc. according to level of importance)

  - a. [ ] Another MFI  
  - b. [ ] Another Bank  
  - c. [ ] Money Lender  
  - d. [ ] None

  - d. [ ] Family and Friends  
  - e. [ ] Own Funds  
  - f. [ ] Other ____________________________

**Q.3** At the moment are you repaying any other loan(s)?

  - a. [ ] No  
  - b. [ ] Yes  
  
  If yes, from which source?

    - i. [ ] Another MFI  
    - ii. [ ] Another Bank  
    - iii. [ ] Money Lender

    - iv. [ ] Family and Friends  
    - v. [ ] Supplier or Employer  
    - vi. [ ] Other ____________________________

**Q.4** How did you hear of Akhuwat?

  - a. [ ] Family and Friends  
  - b. [ ] Advertisement  
  - c. [ ] Television

  - d. [ ] Internet  
  - e. [ ] Sign Board  
  - f. [ ] Other ____________________________
II.1 Questionnaire AIM Clients 2015 (cont.)

Q.5 For what purpose did you seek the loan?

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

Q.6 Why did you approach specifically Akhuwat for the loan?
(respondent can indicate more than one reason, Please rank answers according to importance, give 1 to the most important reason, 2 to the second most important, etc.)

a. ☐ It is a loan that is compliant with Islamic teachings.
b. ☐ It is cheaper than other providers
c. ☐ The application and approval procedure is straightforward and relatively easy
d. ☐ Other reason (please elaborate)

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

Q.7 How satisfied are you with Akhuwat’s service?

a. ☐ Very Satisfied       b. ☐ Satisfied       c. ☐ Neither Satisfied or Dissatisfied
d. ☐ Unsatisfied          e. ☐ Very Unsatisfied

Q.8 For the last month, can you estimate the income that you earned personally?

a. ☐ No
b. ☐ Yes if yes, how much __________________________

Q.9 For the last month, can you estimate the total income earned by all members of the household?

a. ☐ No
b. ☐ Yes if yes, how much __________________________

Q.10 For the last year, in terms of sales (revenues) has the business been

a. ☐ Growing
b. ☐ Stable
   c. ☐ Decreasing

What are the main constraints to your business growing?

________________________________________________________________________

________________________________________________________________________

Page 2 of 3
II.1 Questionnaire AIM Clients 2015 (cont.)

<table>
<thead>
<tr>
<th>Q.11</th>
<th>For the last month, can you estimate the total expenses of the household?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>❑ No</td>
</tr>
<tr>
<td>b.</td>
<td>❑ Yes if yes, how much</td>
</tr>
<tr>
<td></td>
<td>How much of this was spent on:</td>
</tr>
<tr>
<td>a.</td>
<td>Food</td>
</tr>
<tr>
<td>b.</td>
<td>Education</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q.12</th>
<th>For the last year, did your household save?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>❑ Never</td>
</tr>
<tr>
<td>b.</td>
<td>❑ Occasionally</td>
</tr>
<tr>
<td>c.</td>
<td>❑ Regularly</td>
</tr>
<tr>
<td>d.</td>
<td>❑ Other (please specific)</td>
</tr>
<tr>
<td></td>
<td>If the answer is b or c, how did they save?</td>
</tr>
<tr>
<td>a.</td>
<td>❑ Commercial bank</td>
</tr>
<tr>
<td>b.</td>
<td>❑ MFI or other savings group</td>
</tr>
<tr>
<td>c.</td>
<td>❑ At home</td>
</tr>
<tr>
<td>d.</td>
<td>❑ Other (please specific)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q.13</th>
<th>Is your house?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>❑ Owned yourself (property title)</td>
</tr>
<tr>
<td>b.</td>
<td>❑ Owned yourself (without property title)</td>
</tr>
<tr>
<td>c.</td>
<td>❑ Rented</td>
</tr>
<tr>
<td>d.</td>
<td>❑ Owned by family</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q.14</th>
<th>In the previous year did you make any improvements to your house?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>❑ No</td>
</tr>
<tr>
<td>b.</td>
<td>❑ Yes if yes, which</td>
</tr>
<tr>
<td></td>
<td>Buy any new household furniture?</td>
</tr>
<tr>
<td>a.</td>
<td>❑ No</td>
</tr>
<tr>
<td>b.</td>
<td>❑ Yes if yes, which</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q.15</th>
<th>Are there serious health problems with any member of the household?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>❑ No</td>
</tr>
<tr>
<td>b.</td>
<td>❑ Yes if yes, do they effect</td>
</tr>
<tr>
<td>i.</td>
<td>❑ Adults</td>
</tr>
<tr>
<td>ii.</td>
<td>❑ Children</td>
</tr>
<tr>
<td>iii.</td>
<td>❑ Both</td>
</tr>
<tr>
<td></td>
<td>Please explain nature of illness:</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Q.16</th>
<th>Was there any event in the last year that had a strong impact on the household or the business (such as marriage, birth, funeral, accident or illness)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>❑ No</td>
</tr>
<tr>
<td>b.</td>
<td>❑ Yes if Yes, what was the event and how did it affect the household or business?</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: English version of the questionnaire. The questionnaire used in the interviews was the version in Urdu.
### II. II PPI Form for Pakistan - Summary Look-up Table

<table>
<thead>
<tr>
<th>PPI Score</th>
<th>National Line (per adult equivalent)</th>
<th>USAID 'Extreme'</th>
<th>Intl. 2005 PPP (per person)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50%</td>
<td>75%</td>
<td>100%</td>
</tr>
<tr>
<td>0-4</td>
<td>0.0</td>
<td>73.8</td>
<td>95.4</td>
</tr>
<tr>
<td>5-9</td>
<td>2.4</td>
<td>71.2</td>
<td>95.1</td>
</tr>
<tr>
<td>10-14</td>
<td>9.9</td>
<td>39.7</td>
<td>84.1</td>
</tr>
<tr>
<td>15-19</td>
<td>5.8</td>
<td>38.9</td>
<td>68.0</td>
</tr>
<tr>
<td>20-24</td>
<td>2.6</td>
<td>18.9</td>
<td>57.6</td>
</tr>
<tr>
<td>25-29</td>
<td>1.0</td>
<td>12.7</td>
<td>47.1</td>
</tr>
<tr>
<td>30-34</td>
<td>0.4</td>
<td>8.7</td>
<td>39.5</td>
</tr>
<tr>
<td>35-39</td>
<td>0.1</td>
<td>5.8</td>
<td>29.8</td>
</tr>
<tr>
<td>40-44</td>
<td>0.0</td>
<td>2.6</td>
<td>17.4</td>
</tr>
<tr>
<td>45-49</td>
<td>0.9</td>
<td>4.6</td>
<td>16.9</td>
</tr>
<tr>
<td>50-54</td>
<td>0.0</td>
<td>1.2</td>
<td>10.7</td>
</tr>
<tr>
<td>55-59</td>
<td>0.2</td>
<td>0.3</td>
<td>7.4</td>
</tr>
<tr>
<td>60-64</td>
<td>0.0</td>
<td>0.3</td>
<td>5.1</td>
</tr>
<tr>
<td>65-69</td>
<td>0.0</td>
<td>0.0</td>
<td>0.7</td>
</tr>
<tr>
<td>70-74</td>
<td>0.0</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>75-79</td>
<td>0.0</td>
<td>0.0</td>
<td>1.1</td>
</tr>
<tr>
<td>80-84</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>85-89</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>90-94</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>95-100</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table shows the estimated poverty likelihoods associated with scores.  
Source: Schreiner (2010, p. 335)

Note: The PPI Form for Pakistan in its English version can be found in section 4.3 (p.79)
II.III Questionnaire THRIVE Non-Clients 2016

Impact Assessment Form for Lendwithcare (Control Group)

Date: ___________  Gender: ___________  Age: ___________  Questionnaire No. ___________

Name: ___________  Residence Area: ___________

Educational Level: ___________  Marital Status: ___________  Interviewer: ___________

Details of Household Members

<table>
<thead>
<tr>
<th>Name</th>
<th>Age</th>
<th>Relation with client</th>
<th>Occupation</th>
<th>Age when left school</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of dependants: ___________

To be filled by the interviewer:

Length of interview: ___________  Contact No. ___________

Is the entrepreneur willing to be contacted in the future to repeat the questionnaire?  Yes  No

Q.1 What is your (main) business and how long have you owned this business? (state in years or months if less than one year)

Q.2 On average, how many hours in a day and how many days in a week do you spend on your business?

Q.3 Is your business seasonal?
   a. No  b. Yes (please elaborate) ___________

Q.4 Do you work alone in the business (number should be included in the box)?
   a. No (unpaid family/friends)  b. No (paid employees)  c. Yes

Q.5 How do you fund your activities? (respondent can indicate more than one option, please rank 1,2, etc. according to level of importance)
   a. MFI  b. Bank  c. Money Lender
   d. Family and Friends  e. Own Funds  f. Other ___________
Q.6 Why have you not applied for a loan to fund your business? (only if answer to Q5 is not a. or b.)
   a. [ ] Did not need the loan (enough own funds or family support)
   b. [ ] Loans are too expensive (business profit is not enough to repay the loan)
   c. [ ] Process is too complicated or too slow (too many documents, many questions, too much time)
   d. [ ] Did not fulfil the conditions to have a loan. Why? ________________________
   e. [ ] Other reason (if other, please elaborate)

Q.7 Are you repaying any loan for other purposes (consumption, education, housing, etc.)?
   a. [ ] No  b. [ ] Yes
      If Yes, from which provider?
   i. [ ] MFI  ii. [ ] Bank  iii. [ ] Money Lender
   iv. [ ] Family and Friends  v. [ ] Supplier or Employer  vi. [ ] Other

Q.8. For the last month, how much was your personal income?
   a. From the business(es): ____________  b. From other sources: ____________

Q.9 For the last month, how much was the total income of the other members of the household?
   a. From businesses: ____________  b. From salaries: ____________  c. Other sources: ____________

Q.10 Are you the household breadwinner?
   a. [ ] Yes  b. [ ] No
      If No, what is the breadwinner’s income source?
   i. [ ] Casual/Temporary Worker  ii. [ ] Self Employed
   iii. [ ] Other  iv. [ ] Permanent Paid Worker

Q.11 For the last year, in terms of sales (revenues) has the business been
   a. [ ] Growing  b. [ ] Stable  c. [ ] Decreasing
      What do you think is missing (you need) for your business to grow?
      ____________________________________________________________
      ____________________________________________________________

Q.12 For the last month, how much were the expenses of the household on:
   a. Food: ____________  b. Education: ____________
   c. Health: ____________  d. Other expenses: ____________
      (rent, electricity, transport, phone, etc.)
Q.13 For the last year, did your household save?
   a. Never   b. Occasionally   c. Regularly
   If the answer is b or c, how did they save?
   i. Commercial bank   ii. ROSCA   iii. Mobile Banking
   iv. At home   v. Other (please specify) ________________

Q.14 Is your house?
   a. Owned yourself (property title)   b. Owned yourself (without property title)
   c. Rented   d. Owned by the spouse
   e. Owned by family   f. Other (please specify) ________________

Q.15 What is the number of rooms in the house (excluding bathrooms and toilets)? ________________

Q.16 Do you share a toilet and bathroom with any other family?
   a. No   b. With 1 other family
   c. With 2 or more families

Q.17 In the previous year did you make any improvements to your house?
   a. No   b. Yes If Yes, what? ________________
   Buy any new household furniture or asset (example vehicle)?
   a. No   b. Yes If Yes, which? ________________

Q.18 Do you own any of the following assets? (introduce the number inside the box; 0.5 if battered and old)
   d. Fridge   e. Home theatre   f. DVD player
   g. Heater   h. Sofa set   i. Car

Q.19 Do you participate in any community organisations?
   a. No   b. Yes If Yes, which? ________________
Q.20 Who generally makes the decisions about:

a. Household income (e.g. loans, spending, asset acquisition)
   i. ☐ The entrepreneur
   ii. ☐ The entrepreneur and other household member(s)
   iii. ☐ Other household member(s) (state who)

b. Household activities (e.g. cooking & cleaning)
   i. ☐ The entrepreneur
   ii. ☐ Other household member(s) (state who)

cc. Child related issues (e.g. schooling, health, well-being)
   i. ☐ The entrepreneur
   ii. ☐ Other household member(s) (state who)

Q.21 Are there serious health problems with any member of the household?

   a. ☐ No
   b. ☐ Yes

   If Yes, who has the condition?
   i. ☐ Adults
   ii. ☐ Children
   iii. ☐ Both

   Please explain nature of illness:
   ______________________________________________________________

Q.22 Was there any event in the last year that had a strong impact on the household or the business (such as marriage, birth, funeral, accident or illness)?

   a. ☐ No
   b. ☐ Yes

   If Yes, what was the event and how did it affect the household or business?
   ______________________________________________________________
II.IV THRIVE-PAT Form

POVERTY ASSESSMENT TOOL - Zimbabwe
Interviewer: Text in bold should be read aloud. Text in italics are instructions and should not be read aloud.

You should use probing questions if necessary to elicit responses to all questions. Ensure all the sections are filled in clearly.

<table>
<thead>
<tr>
<th>Fill out the information below before the survey begins. Do not ask the respondent for this information.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Name</td>
</tr>
<tr>
<td>Individual Name</td>
</tr>
<tr>
<td>Interviewer</td>
</tr>
<tr>
<td>Branch</td>
</tr>
<tr>
<td>Date of Interview</td>
</tr>
<tr>
<td>Client Location</td>
</tr>
<tr>
<td>Interview Location</td>
</tr>
<tr>
<td>Loan Cycle</td>
</tr>
<tr>
<td>Business Activity</td>
</tr>
</tbody>
</table>

1 Is this household male or female headed?  
1. Male  
2. Widowed  
3. Female  
4. Never Married  
Score:  

2 What is the breadwinner’s income source?  
1. Casual/Temporary worker  
2. Self employed  
3. Other  
4. Permanent paid worker  
Score:  

3 What is the person to room ratio at this household?  
(ie number of rooms excluding bathrooms and toilets / number of people at the household inc guests)  
Score:  

<table>
<thead>
<tr>
<th>Number of rooms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of people</td>
<td></td>
</tr>
<tr>
<td>1. Grandparents</td>
<td>&gt;/4</td>
</tr>
<tr>
<td>2. Parents</td>
<td>&gt;/3</td>
</tr>
<tr>
<td>3. Children</td>
<td>&gt;/2</td>
</tr>
<tr>
<td>4. Grandchildren</td>
<td>&gt;/1</td>
</tr>
<tr>
<td>5. Others</td>
<td>&gt;/=0.5</td>
</tr>
<tr>
<td>Total</td>
<td>less than 0.5</td>
</tr>
</tbody>
</table>

Ratio:  

4 Do you share a toilet and bathroom with any other family?  
Score:  

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>2</td>
</tr>
<tr>
<td>With 1 other family</td>
<td>-2</td>
</tr>
<tr>
<td>With 2 or more families</td>
<td>-5</td>
</tr>
</tbody>
</table>
II.VI THRIVE-PAT Form (cont.)

Survey number __________

5 What is the household dependency ratio?

<table>
<thead>
<tr>
<th>Number of dependants at household</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;0.35</td>
<td>3</td>
</tr>
<tr>
<td>0.35 &lt;= x &lt;= 0.5</td>
<td>2</td>
</tr>
<tr>
<td>0.51 &lt;= x &lt;= 0.75</td>
<td>1</td>
</tr>
<tr>
<td>&gt;0.75</td>
<td>-2</td>
</tr>
</tbody>
</table>

6 What is the household size?

<table>
<thead>
<tr>
<th>&lt;3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
</tr>
</tbody>
</table>

7 Do you own any of the following? (Half the score if its battered and old)

<table>
<thead>
<tr>
<th>A</th>
<th>Basic Item</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Electric hotplate</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Bed</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Electric 3/4 plate stove</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Fridge</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>Luxury item</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Home theatre</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>DVD player</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Heater</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>Sofa set</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>Car</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
</tr>
</tbody>
</table>

8 What is your household tenure status?

<table>
<thead>
<tr>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

9 Total score

<table>
<thead>
<tr>
<th>Yes</th>
<th>Inactive</th>
<th>Active</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Are you a bank account holder?

<table>
<thead>
<tr>
<th>No</th>
</tr>
</thead>
</table>

These are all the questions I need to ask you today. Thank you for your time and effort in completing this survey.
III.1 Questionnaire AIM Non-Clients 2017

Impact Assessment Form for Lendwithcare

<table>
<thead>
<tr>
<th>Questionnaire No.</th>
<th>Name Interviewer</th>
<th>Branch:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date:</td>
<td>Name Entrepreneur:</td>
<td></td>
</tr>
<tr>
<td>Gender:</td>
<td>Age:</td>
<td>Marital Status:</td>
</tr>
<tr>
<td>Education Level:</td>
<td>Contact No.</td>
<td></td>
</tr>
</tbody>
</table>

Details of Household Members

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Relation with client</th>
<th>Age</th>
<th>Occupation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total Number of People in the Household: ____________________

Instructions for Interviewer:
Were there other people in the room during the interview?
Yes ☐  No ☐  If Yes, please identify their relationship with the clients ____________________________
Did they answer questions directed to the client?  Yes ☐  No ☐

Length of Interview: ____________________ Is the client willing to be contacted again to complete the questionnaire? Yes ☐  No ☐

Q.1 What is your business?

__________________________________________

For how long have you owned this business? (state in years or in months, if less than one year) ____________________

Q.2 On average, how many hours in a day and days in a week do you dedicate to your business?

__________________________________________

Q.3 Do you work alone? (if there are regular employees, indicate the number; if partial time, write 0.5)

Now:  a. ☐ No  b. ☐ Yes  If Yes, Unpaid Employees: ________  Paid Employees: ________

Start first loan:  a. ☐ No  b. ☐ Yes  If Yes, Unpaid Employees: ________  Paid Employees: ________

Q.4 a) When making relevant business decisions, do you decide alone or jointly with other household members?

a. ☐ Alone  b. ☐ With other(s) household member __________________________

Q.4b) When making relevant decisions on how to spend the money in the household, do you decide alone or jointly with other household members?

a. ☐ Alone  b. ☐ With other(s) household member __________________________
III.I Questionnaire AIM Non-Clients 2017 (cont.)

Q.5 How do you fund your business? (respondent can indicate more than one option, please rank 1, 2, ... according to level of importance)
   a. □ MFI  b. □ Bank  c. □ Money Lender
   d. □ Family and Friends  e. □ Own Funds  f. □ Other ___________________

Q.6 Have your business funding sources changed compared to the time of the last survey?
   a. □ No  b. □ Yes

If Yes, why? __________________________________________________________

Q.7a) If you have now applied for a loan, what was the purpose of the loan? (if respondent answered a, b or c in Q.5)

______________________________________________________________

Q.7b) What was the main reason to choose the financial provider? (respondent can indicate more than one option, please rank 1, 2, ... according to level of importance)
   a. □ Cheaper cost of loan  b. □ Simple process/requirements  c. □ Recommendation of others
   d. □ Proximity to home/business  e. □ Compliance with Islamic principles  f. □ Other ___________________

Q.8 For the last month, can you estimate the income that you earned personally?
   a. □ No  b. □ Yes  If Yes, how much ____________________

Q.9 For the last month, can you estimate the total income earned by all members of the household?
   a. □ No  b. □ Yes  If Yes, how much ____________________

Notes on Income Values: _______________________________________________________________________

Q.10a) Over the past 12 months, have business revenues (sales) been:
   a. □ Growing  b. □ Stable  c. □ Decreasing

Q.10b) Compared with the time of the first survey, business revenues (sales) are:
   a. □ Better  b. □ Similar  c. □ Worse

Q.11 What difficulties does your business face?

__________________________________________________________

Q.12 For the last month, can you estimate the total expenses of the household?
   a. □ No  b. □ Yes  If Yes, how much ____________________

Notes on Expenses Values: _____________________________________________________________________
III.I Questionnaire AIM Non-Clients 2017 (cont.)

Q.13 Over the past 12 months, did your household save?
   a. ☐ Never           b. ☐ Occasionally        c. ☐ Regularly
   If the answer is b or c, how did you save?
   i. ☐ Commercial Bank   ii. ☐ MFI          iii. ☐ At home
   iv. ☐ Committee/Savings group  v. ☐ Other

Q.14 Is your house?
   a. ☐ Owned yourself (property title)       b. ☐ Owned yourself (without property title)
   c. ☐ Rented (rent: ______________________)  d. ☐ Owned by family

Q.15a) Over the past 12 months did you make any improvements to your house?
   a. ☐ No       b. ☐ Yes
   If Yes, which ________________________________

Q.15b) Over the past 12 months did you buy any new household furniture or other household asset?
   a. ☐ No       b. ☐ Yes
   If Yes, which ________________________________

Q.16 Are there serious health problems with any member of the household?
   a. ☐ No       b. ☐ Yes
   If Yes, do they affect: i. ☐ Adults         ii. ☐ Children    iii. ☐ Both

Q.17 Was there any event in the past 12 months that had a strong impact on the household or the business (such as marriage, birth, funeral, accident or illness)?
   a. ☐ No       b. ☐ Yes
   If Yes, what was the event and how did it affect the household or business? ________________________________

Q.18 What has changed during the last two years (since the first survey)?
   In the business (such as type of activity/clients, sales, profits, working conditions, assets, employment, etc.):
   ________________________________
   ________________________________

   In the household/personal life (such as family, income, consumption, food security, schooling, quality of life/satisfaction, etc.):
   ________________________________
   ________________________________

Q.19 Do you think your life compared with the time of the first interview is overall is:
   a. ☐ Better       b. ☐ Similar       c. ☐ Worse
## III.II Questionnaire THRIVE Clients 2017

**Impact Assessment Form for Lendwithcare - Clients**

<table>
<thead>
<tr>
<th>Date:</th>
<th>Interviewer:</th>
<th>Questionnaire No.:</th>
<th>Branch:</th>
</tr>
</thead>
</table>

**Name of Borrower:**

**Group:**

**Educational Level:**

**Marital Status:**

**Age:**

**Contact:**

**Gender:** F

### Details of Household Members

<table>
<thead>
<tr>
<th>Sr. No</th>
<th>Relation with Client</th>
<th>Age</th>
<th>Occupation</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Instructions for Interviewer:

The questionnaire should be answered by Thrive’s client. Were there other people in the room during the interview?

a. [ ] No  b. [ ] Yes  If yes, please identify their relationship with the client __________________________

Did they answer questions directed to the client?  Yes [ ]  No [ ]  Length of interview: __________________________

Is the client willing to be contacted or visited again to complete the questionnaire?  Yes [ ]  No [ ]

### Questions

**Q.1** What is your (main) business and how long have you owned this business? (state in years, or months if less than one year)

**Q.2** On average, how many hours in a day and how many days in a week do you spend on your business?

**Q.3** Is your business seasonal?

a. [ ] No  b. [ ] Yes (please elaborate) __________________________

**Q.4** Do you work alone in the business?

a. [ ] No, I work with _________ (number) unpaid family/friends.

b. [ ] No, I have _________ (number) paid employees.

c. [ ] Yes
III. III Questionnaire THRIVE Clients 2017 (cont.)

Q.5 At the moment are you repaying any other loan(s)?
   a. □ No                                b. □ Yes                  If yes, from which source?
   i. □ Another MFI                           ii. □ Bank                          iii. □ Money Lender
   iv. □ Family and Friends                    v. □ Supplier or Employer           vi. □ Other

Q.6 How did you use the first loan?

Q.7 Did you have any problems making repayments on your loan(s)?
   a. □ No                                b. □ Yes                  If yes, please explain why?

Q.8a) Thrive Loan History

<table>
<thead>
<tr>
<th>Loan Nr.</th>
<th>Amount</th>
<th>Duration</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Q.8b) Are you planning to apply for future loans at Thrive?
   a. □ No                                b. □ Yes                  Why?

Q.9 Have you received further training from Thrive (after the initial training)?
   a. □ No
   b. □ Yes, if so, what type of training did you receive (topic, duration)?

Q.10 Has the training been useful?
   a. □ Very Useful                     b. □ Useful                      c. □ Neither useful nor useful
   d. □ Unuseful                        e. □ Very Unuseful

Please provide an explanation:
III.III Questionnaire THRIVE Clients 2017 (cont.)

Q.11 For the last month, how much was your personal income?
   a. From the business(es): _____________
   b. From other sources: _______________

Q.12 For the last month, how much was the total income of the other members of the household?
   a. From businesses: _____________
   b. From salaries: _____________
   c. Other sources: _______________

Q.13 Are you the household breadwinner?
   a. Yes
   b. No

   If No, what is the breadwinner’s income source?
   i. Casual/Temporary Worker
   ii. Self Employed
   iii. Permanent Paid Worker
   iv. Other

Q.14a Over the past 12 months, in terms of sales (revenues) has the business been:
   a. Growing
   b. Stable
   c. Decreasing

Q.14b What do you think is missing (you need) for your business to grow?

Q.15 For the last month, how much were the expenses of the household on:
   a. Food _____________
   b. Education _____________
   c. Health _____________
   d. Other expenses _____________

   Total Expenses: _____________

Q.16 For the last year, did your household save?
   a. Never
   b. Occasionally
   c. Regularly

   If the answer is b or c, how did they save?
   i. Commercial bank
   ii. ROSCA
   iii. Mobile Banking
   iv. At home
   v. Other (please specific)

Q.17 Is your house?
   a. Owned yourself (property title)
   b. Owned yourself (without property title)
   c. Rented _____________
   d. Owned by the spouse
   e. Owned by family
   f. Other (please specify)

Q.18 How many rooms do you have in your house (excluding bathrooms and toilets)? ________

294
### III.III Questionnaire THRIVE Clients 2017 (cont.)

Q.19 Do you share a toilet and bathroom with another family?
- a. [ ] No
- b. [ ] With 1 other family
- c. [ ] With 2 or more families

Q.20 Do you own any of the following assets? (introduce the number inside the box; 0.5 if battered and old)
- a. [ ] Electrical hotplate
- b. [ ] Bed
- c. [ ] Electric 3/4 plate stove
- d. [ ] Fridge
- e. [ ] Home theatre
- f. [ ] DVD player
- g. [ ] Heater
- h. [ ] Sofa set
- i. [ ] Car

Q.21a Over the past 12 months, did you make any improvements to your house?
- a. [ ] No
- b. [ ] Yes if yes, what?

Q.21b Over the past 12 months, did you buy any household furniture or asset (example vehicle)?
- a. [ ] No
- b. [ ] Yes if yes, which?

Q.22a Do you currently participate in any community organisations?
- a. [ ] No
- b. [ ] Yes if yes, which?

Q.22b In the last 12 months, have you had a leadership role [such as president, secretary, treasurer or coordinator] in the community organisation?
- a. [ ] No
- b. [ ] Yes if yes, what was your role?

Q.23 Who generally makes the decisions about:
- a. Household income (e.g. loans, spending, asset acquisition)
  - i. [ ] The borrower
  - ii. [ ] The borrower and other household member(s)
  - iii. [ ] Other household member(s) (state who)

- b. Household activities (e.g. cooking & cleaning)
  - i. [ ] The borrower
  - ii. [ ] The borrower and other household member(s)
  - iii. [ ] Other household member(s) (state who)

- c. Child related issues (e.g. schooling, health, well-being)
  - i. [ ] The borrower
  - ii. [ ] The borrower and other household member(s)
  - iii. [ ] Other household member(s) (state who)

Q.24 Are there serious health problems with any member of the household?
- a. [ ] No
- b. [ ] Yes
  - If Yes, who has the condition?
    - i. [ ] Adults
    - ii. [ ] Children
    - iii. [ ] Both
  - Please explain nature of illness:
<table>
<thead>
<tr>
<th>Question</th>
<th>Description</th>
</tr>
</thead>
</table>
| Q.25 | Was there any event in the last year that had a strong impact on the household or the business (such as marriage, birth, funeral, accident or illness)?  
  a. No  
  b. Yes  
  If Yes, what was the event and how did it affect the household or business? |
| Q.26a | Can you describe any changes to your business in the last 12 months?  
  (examples: type of activity/clients, sales, profits, working conditions, assets, employment ...) |
| Q.26b | Can you describe any changes to your household/personal life in the last 12 months?  
  (examples: family composition, consumption, schooling, food security ...) |
| Q.27a | Compared with the time of the first interview, your life overall is:  
  a. Better  
  b. Similar  
  c. Worse  
  Q.27b | What are the most important factors that have contributed to these positive or negative changes (if applicable)? |
Table of Contents

Table of contents ........................................................................................................................................ 2
Summary .................................................................................................................................................... 3
Introduction ............................................................................................................................................... 4
Part I – The clients and their businesses ................................................................................................. 5
Part II – The changes after the first loan cycle ......................................................................................... 8
  The business outcomes .......................................................................................................................... 8
  The household outcomes ....................................................................................................................... 10
Part III – The comparison with non-clients ............................................................................................ 14
Conclusion ............................................................................................................................................... 16
References .............................................................................................................................................. 16
Annexes .................................................................................................................................................. 47
  List of variables ..................................................................................................................................... 17
  Non-parametric tests .............................................................................................................................. 18
  Regression results ................................................................................................................................. 21
Summary

This report is part of the Lendwithcare (LWC) assessment project and focuses on the evaluation of LWC Pakistani partner, the microfinance institution AIM Islamic Microfinance (AIM). The report was prepared by the University of Portsmouth (UoP), partner in the project, after a second wave of a household survey to a sample of AIM clients who have been supported by the LWC crowdfunding platform.

The study sample included 500 new AIM clients and 100 non-clients, first interviewed in 2015 by a team of independent interviewers recruited from local universities. The second wave of interviews took place in 2017, after all the clients had completed repaying their first loan (20 to 22 months later).

This report offers an initial snapshot of how the lives of LWC supported entrepreneurs have changed since they became AIM clients. Considering the nature of the methodology used, the analysis does not give definite explanations for the identified changes but provides critical data on the microcredit clients’ lives and the existence of correlations between the factors studied. In this sense, it can become the foundation of a more in-depth analysis and an important incentive for the establishment of continuous evaluation processes in the microfinance institution involved in the project.

The results show that AIM/LWC clients are mainly male (63%), married (90%), with low levels of literacy (48% illiterate), living in family houses (58%) in households of 6 members. The average age is 39 years old, and on average they have significant business experience (10 years). 54% work alone in the business and they report working an average of 64 hours per week. Their first loan was, on average, 20,000 PKR (approximately $190) and was destined mostly to fund working capital. By the time of the second interview, 63% of the clients had successfully applied for a second loan and 13% were considering applying.

The data collected in both waves of the survey point towards an improved situation for most of the clients’ businesses, with a significant share of them reporting increased business revenues and less constraints in the access to capital for their businesses. Positive changes were also identified in the employment indicators: a considerably lower number of entrepreneurs were working alone in their business in 2017, and a reported net increase of 191 jobs during the period. It is important to note, however, that this increase in employment was concentrated in a relatively small group of entrepreneurs, and that these created ‘jobs’, even if paid, were mostly informal.

The analysis of the changes at the household level was based mainly on the variation of the PPI, Progress out of Poverty Index (renamed in October 2017 as Poverty Probability Index). Most clients (55%) have seen their PPI total score increase between 2015 and 2017, anticipating an improvement in the economic circumstances of their households. Further analysis of the changes in the different components of the index will be necessary to identify the reasons for the direction and amplitude of the variations, including for those who have experienced a deterioration of the score.

Among the findings, it is interesting to note that two of the clients’ segments that were found to be comparatively poorer at the baseline survey - female and illiterate clients - registered slightly larger increases in their average PPI scores, suggesting a reduction in the poverty gaps for these sub-groups of clients. Furthermore, the comparative analysis of clients and non-clients identified that the access to the loan was one of the statistically significant factors associated with the changes in the PPI score for those respondents who have experienced positive variations. Simultaneously, for respondents who experienced negative variations the loan was not statistically significant. Given the limitations of the study, especially the difference between clients and non-clients sample sizes, this result requires further corroborations.

A final remark would be to call attention to the overall positive evaluation made by the clients regarding the changes in their lives - 68% declared having a better quality of life in 2017. It is also interesting to note that 93% of these clients have since become AIM donors, with 66% reporting making regular donations to the institution.
Introduction

The Lendwithcare assessment project started in 2014 with the main objective of assessing the experience of borrowers supported by the crowdfunding platform. By the end of 2014, AIM in Pakistan had been chosen as the first field partner to participate in the evaluation, and the University of Portsmouth also became a partner in the project.

AIM is a Pakistani microfinance institution (MFI) with its headquarters in Lahore (Punjab). It is a NGO funded mainly by national and international donations, which has enabled the institution to finance thousands of micro entrepreneurs through gird hossan or interest free loans, compliant with the principles of Islam. AIM’s microcredit programme is unique in some of its features, including the family-loan approach and the public signature of loan contracts in religious sites. The use of mosques, and on occasions churches, in addition to strengthening the client’s relationship with the institution, is part of a strategy to reduce fixed costs. The physical structure composed of simple and small branches and the support of volunteers are also vital elements of this strategy.

The institution started operations in 2001. Today it is the largest microfinance institution in Pakistan in terms of number of active borrowers, with more than 855,000 clients, 700 branches spread throughout the country provinces and 5,600 staff. AIM has been LWC’s partner in Pakistan since February 2013. As of 30th September 2017, more than 11,300 loans, with an average amount of $236, have been funded by LWC.

The borrowers supported by LWC lenders come from 4 branches within the Lahore area: Badami Bagh, Kot Khawaja Saeed and Kahna Nau are located within the city, while Kasur is situated around 30 miles from Lahore. During the months of April and May 2015, a sample of 500 AIM new clients selected from the four branches and 100 non-clients were interviewed by a team of independent interviewers recruited from a local university. The interviewers received training from the UoP/LWC team to apply the questionnaire, which was prepared by UoP with contributions from LWC and AIM. The baseline survey allowed for the characterisation of the respondents and established a basis for comparative analysis in 2017, after the implementation of a second wave of the survey.

The second questionnaire was to be applied to the same clients and non-clients, after the clients had completed the first loan cycle. It included some questions identical to the first questionnaire, new questions about the changes in a selection of outcomes at business and household level, and two open questions aiming to capture the perception of the interviewees regarding the observed changes. In both surveys, the PPI form was also completed.

Longitudinal studies present several challenges, one of the most significant being attrition. As expected, two years after the baseline survey not all participants were located or were available to be interviewed again. This was the case for a significant number of non-clients, particularly in the branch of Kahna Nau. The final dataset included valid data for the two periods for 447 clients and 52 non-clients. The reasons for non-participation by the clients in the second survey include death (3), moving to other country or city (21), not located by the interviewer (26) and incomplete interviews (3).

---

1 All figures and tables are based on survey data.
2 Data by 30th September 2017 available in MicroWatch Issue 45 (http://www.pmn.org.pk/publications/category/MicroWatch)
3 PPI is a poverty measurement tool initially developed by the Grameen Foundation under the name of Progress out of Poverty Index. It has been renamed in October 2017 as Poverty Probability Index. Further information on PPI can be found at https://www.povertyindex.org/
4 Interviews took place mainly between February and May 2017; non-clients from Kahna Nau were interviewed in November 2017.
The geographical distribution of the sample for both ‘clients’ and ‘non-clients’ in the two surveys is shown in Table 1.

<table>
<thead>
<tr>
<th>Branch</th>
<th>Clients 2015</th>
<th>Clients 2017</th>
<th>Non-clients 2015</th>
<th>Non-clients 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Badami Bagh</td>
<td>79</td>
<td>74</td>
<td>28</td>
<td>14</td>
</tr>
<tr>
<td>Kot Khawaja Saeed</td>
<td>105</td>
<td>91</td>
<td>21</td>
<td>16</td>
</tr>
<tr>
<td>Kahna Nau</td>
<td>132</td>
<td>107</td>
<td>32</td>
<td>10</td>
</tr>
<tr>
<td>Kasur</td>
<td>184</td>
<td>175</td>
<td>19</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>500</td>
<td>447</td>
<td>100</td>
<td>52</td>
</tr>
</tbody>
</table>

Part I - The clients and their businesses

AIM clients are mainly male. The distribution of clients is unbalanced in terms of gender (Figure 1), but it is overall in line with the gender distribution of the microcredit program for the institution as a whole (40% of female borrowers in 2015, when the sample clients were first interviewed).³

[Figure 1 - Gender Sample Distribution]

Female participation is lower in the more urban and economically developed areas of Badami Bagh and Kot Khawaja Saeed. Considering that during the period of the baseline survey, all new clients from the four branches were invited to participate in the study, and the level of the microfinance on offer was similar in all areas (15 main microfinance service providers have branches in these four areas, with 4 further providers present in all but Kasur), it can be conjectured that women are more likely to look for MFI funding for their businesses in the poorer areas.

The average age of the clients in the sample is 39 years. Figure 2 illustrates the distribution of the clients by age groups and the concentration on clients between 30 and 50 years (70%). The focus is, therefore, on more experienced entrepreneurs, which also reflects the average business time (10 years for those with active businesses in 2017). Female clients are, on average, older than male clients (41 years vs 37 years) but less experienced in the business (9.1 years vs 10.5 years).

³ See Khan et al. (2017)
A strong characteristic of the sample clients is their low level of formal education. Almost half of the clients are illiterate (48.4%) and a further 16% received only primary education (grades 1 to 5). As might be anticipated, this is particularly notable for female clients – 67% have received no formal education.

Another common feature of most clients in the sample is their marital status: 89.7% are married, 9.4% single and an almost insignificant percentage are widowed and divorced (0.45% for each). Household size averages 6.39, a value slightly above the same sample average in 2015 (6.18) and the data for the Punjab province in the HIES – Household Integrated Economic Survey 2015-16 (6.04).7

Clients in the sample live mainly in houses owned by other family members (58%) or their own houses (33%). The number of clients living in rented accommodation was initially low (43) and has slightly decreased in the second survey (39). An intriguing result is that the share of clients stating owning their own house decreased compared with the baseline (40%). It would be interesting to understand if this variation results from inconsistencies in the clients’ responses or if it actually represents a change in home ownership.

Decision making regarding how household money is spent is most frequently shared (less than one quarter of the clients reported taking these decisions on their own). Similarly, for almost two thirds of the female clients, business decisions are made jointly with other household members (often the husband but in many cases also the parents, siblings and children). For male respondents, the percentage of clients managing the businesses on their own is significantly higher (69%).8

---


8 The question on decision making was included only in the second questionnaire, hence no comparison can be made with the baseline situation.
AIM's clients run different types of businesses. In the analysis, sewing/tailoring is singled out as the most common activity reported in the sample, and it is managed by both female and male clients. The remaining businesses were classified under a broad typology of trade (mainly retail); production and construction; and services and education (including hairdressers, recycling and transport).

Figure 4 – Business Activities in 2017

It is relevant to note, however, significant differences at branch level. Trade is the most common activity in the three branches located in Lahore, representing between 36% and 39% of the reported client activities, while in Kasur the main business is sewing/tailoring (39%) followed by production (32%). Services assume a higher importance in Kot Khawaja Saeed where the number of businesses is similar to sewing/tailoring (21-22%). There were 102 clients (23%) whose main activity changed between the two survey periods.

54% of clients stated working alone in the business in 2017. Those who have employees report providing paid jobs to 336 persons and occupation to 67 additional unpaid employees. In the 2017 survey, data was also collected on the working time of the entrepreneurs. For the 349 AIM clients who replied to the question, the average working load was 64 hours/week. Only 24 entrepreneurs declared working less than 40 hours/week, which suggests that part-time dedication to the business is not strong in the sample, even for women who are expected to divide their time between business and household activities.

By the time of the second survey, all clients had finished repaying their first loan (the loan duration varied from 12 to 20 months). 63% of these clients had applied for a second loan that had already been approved or was being evaluated, while 13% were thinking of applying but had not initiated the process. The drop-out rate was 24%. There are three main reasons given by the clients for not taking up a second loan, displayed in Figure 5.

Figure 5 – Reasons for Drop-Out

UoP – PORTSMOUTH BUSINESS SCHOOL

7
LWC – AIM REPORT 2018

The reasons for loan refusal by the organisation are specified in only a small number of the cases and they relate to the client’s failure to comply with application requirements (formation of a group, documentation). According to the clients’ responses, the indebtedness level is low (95% of the interviewees declare having no other loans from any source, formal or informal, at the time of the interview). This is similar to the situation encountered two years before.

Looking at the declared use of the loan (which many microfinance evaluation studies acknowledge as relevant to the programs’ outcomes), it looks as if most clients (80%) have used the loan as stated in their loan application and the baseline survey: mainly to fund working capital needs. There are, however, 84 cases in which the use of the loan reported in 2017 is different from the stated purpose in the 2015 survey. Of these, as shown in Table 2, only 8 clients admit non-productive uses of the loan - mainly payment of household expenses and/or repayment of other debts. This result should be considered with a ‘pinch of salt’ (at this stage) as there was no further validation of the information given by the clients.

<table>
<thead>
<tr>
<th>Purpose of Loan 2015</th>
<th>Use of the Loan 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Working Capital</td>
</tr>
<tr>
<td>Working Capital</td>
<td>-</td>
</tr>
<tr>
<td>Fixed Assets</td>
<td>34</td>
</tr>
</tbody>
</table>

Part II – Changes after the first loan cycle

From a theoretical perspective, links can be established between microcredit and a wide range of socio-economic and well-being indicators. However, considering the characteristics of the population and the logistic challenges associated with the project, the partners established a time limitation for the questionnaires, and consequently, only a few of the potential outcomes were selected to be studied. The selection was made considering relevant impact studies in the sector and comments and suggestions from the field partner. The questionnaires included both business and household level outcomes.

In the interpretation of the results, it is important to keep in mind that a longitudinal survey of this nature does not give definite explanations for the identified changes but provides critical data on the microcredit clients’ lives and the existence of correlations between the factors studied. In this sense, it can become the foundation of a more in-depth analysis and an important incentive to establish continuous evaluation processes in the MFIs involved in the project.

**The business outcomes**

Microcredit clients run small informal businesses, usually with no organised accounting, which makes it more challenging to collect quantitative data on their businesses. The decision was to include a qualitative appreciation of sales growth for the previous 12 months (in both surveys) and a comparison of the sales/revenues level at the time of the two interviews.

Figure 6 illustrates the different scenarios encountered by AIM new clients and the overall positive evolution of most of the businesses during the period. 74% (329 clients) reported higher sales in 2017 compared with the baseline survey. This better performance was achieved by clients with different starting points at the baseline. As the figure shows, there were clients whose business was stable (192) or growing (197) when they applied for the

---

5 For further development see Duvendack et al. (2011), Odell (2010) and Banerjee et al. (2015).
first loan, but also clients with businesses going through difficulties (56). For the clients in a worse position in 2017 (5% of the total sample), the situation was mainly associated with business failure.

Figure 6 - Comparison Sales Revenues in 2017 vs Baseline (No. Clients)

There was also a positive evaluation in two other business areas: access to credit and employment. In both questionnaires, the clients were asked about the main constraints they face in their businesses. There was a significant change in their replies. The percentage of clients identifying lack of capital as the main limitation decreased from 76.5% to 20% in the second questionnaire. Female clients, similarly to the first survey, seemed to experience more difficulties accessing capital (27%) than their male counterparts (16%). The results suggest, however, that access to the microcredit loan did address, at least partially, previous credit constraints. The second most common constraint was electricity outages, especially for Kasur clients.

The comparison of the number of workers reported to be involved in the businesses at the time of the two interviews shows also a positive impact on local employment. The majority of the clients still worked on their own (54%) but this represented a significant decrease compared to the initial situation when 69% declared working alone. The total number of workers (client + unpaid employees + paid employees) for the whole sample had a net increase of 191 from 2015 to 2017. These employment figures are strongly influenced by the performance of one small group of clients, mainly in the sewing/tailoring and production activities, who are responsible for hiring 160 paid and 7 unpaid employees in 2017.

The 2017 questionnaire also included an open question aimed at understanding the perception of clients on the business level changes, and how they valued them. The answers were categorised according to the options shown in Table 3. Increase in revenues and assets were the two most reported business outcomes. Employment creation was less stated, which is partially explained by the fact that the major changes in employment were concentrated in a small number of entrepreneurs.

<table>
<thead>
<tr>
<th>Change in Business</th>
<th>No. Respondents</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increase in sales/profit</td>
<td>329</td>
<td>74%</td>
</tr>
<tr>
<td>Increase in business assets</td>
<td>150</td>
<td>34%</td>
</tr>
<tr>
<td>Increase in clients/orders</td>
<td>80</td>
<td>18%</td>
</tr>
<tr>
<td>Increase in employment</td>
<td>37</td>
<td>8%</td>
</tr>
<tr>
<td>No change</td>
<td>46</td>
<td>10%</td>
</tr>
<tr>
<td>Business shut down/no business</td>
<td>29</td>
<td>6.5%</td>
</tr>
</tbody>
</table>
The household outcomes

PPI (formerly the Progress out of Poverty Index and now the Poverty Probability Index) was the tool chosen to primarily track changes in the poverty level of the clients’ households. Looking at the average values for the PPI sample variation, the results point towards a reduction in the poverty levels of the clients (on average) with a lower likelihood of their households to be considered poor. PPI results, similarly to the questionnaires, do not allow for exploring the causes of the variations, but they raise questions about the dynamics of these changes, which require further qualitative analysis.

Brief methodological note:

PPI is a poverty scorecard developed exclusively for each country by a team of experts led by Mark Schreiner. It includes 10 multiple-choice questions about household characteristics and assets ownership. The answers to these questions have associated scores and their sum gives the PPI score for the household. It ranges from 1 to 100 and translates into a likelihood of the household being poor according to the selected poverty line.10 The higher the score, the lower the probability of the household being considered poor.

The PPI for Pakistan was developed in 2009, and has not been updated since, which could raise some questions on the instrument’s ability to accurately identify the poor in the current environment. It is also important to note that PPI is an instrument developed to assess national averages, whereas all the clients in the sample were located in Punjab, and therefore scored the maximum for the location indicator, given that Punjab performs economically better than most other provinces in the country.

These considerations were taken into account in two ways. The use of the tool focuses on tracking the changes within the sample: how the clients’ poverty level may have improved, with no definite assertions being made on their absolute poverty level. Moreover, the poverty line chosen was the $2.50/day 2005 PPP (Purchasing Power Parity), which is closer to the recently revised international extreme poverty line of $1.90/day.

The PPI survey was conducted in both the first and second phase of interviews for both clients and non-clients, and the index scores and their variations have been computed. Results were analysed using $2.50/day as the reference poverty line and, to better visualise the differences within the sample, the clients were classified in 4 categories:

- ‘Most Likely Poor’ (probability of the household income being below $2.50/day higher than 75%);
- ‘Probably Poor’ (probability between 50% and 75%);
- ‘Vulnerable’ (probability between 25% and 50%);
- ‘Better Off’ (probability lower than 25%).

Figure 7 shows the sample composition considering these categories in 2015 (baseline) and 2017. The overall economic condition of the households seems to have globally improved with a share of households with a very high probability of being considered poor significantly lower in 2017.

---

10 PPI offers, as possible reference poverty lines, a national poverty line and international poverty lines, being most commonly used the $1.25/day and $2.50/day 2015 PPP (https://www.povertyindex.org/).
The aggregate picture conceals, however, a diversity of situations that are best identified when looking at the variations of the individual scores. A majority of clients (55%) have seen their PPI increase between 2015 and 2017, anticipating an improvement in their economic circumstances, but there is also 6% of the sample for whom there was no change and 38% who experienced a deterioration of their score. An important follow-up on this report will be the analysis of which factors (questions in PPI) contributed the most to the variations in both directions, as well as the amplitude of variation.

The existence of differences among clients regarding PPI scores, and respective variation, becomes more evident when applying statistical tests to compare client sub-groups (according to some defined characteristic). The main results of this analysis include:

- In the 2015 survey there were significant differences between the average PPI scores (and poverty levels) for female and male clients. This gap seems to have slightly decreased after the first loan cycle, as the positive variation of the PPI score was stronger, on average, for the female clients (12 percentage points compared to 8 percentage points).
- There were also significant differences in the average PPI scores for illiterate clients compared to those with formal education. Similarly, the poverty gap between these two client segments appears to have reduced by the 2017 survey, with a larger positive variation of the scores for the illiterate clients (13 percentage points compared to 6 percentage points).
- PPI average scores, both in 2015 and 2017, varied considerably with location. Kot Khawaja Saeed is the ‘wealthiest’ branch, with its clients scoring an average PPI of more than 75 points in both surveys. A distinct portrait emerges when looking at the much lower average scores for the Kanha Nau and Kasur branches (60.7 and 63.5, respectively). In all branches, the average PPI scores had improved by the second survey and no statistically significant differences in PPI variation were found between branches.
- The average PPI scores in both surveys suggest that clients who own their houses are economically better than those living with family or renting.

Aiming to complement the analysis based on PPI, both surveys included questions on Income and expenditure. These questions are sensitive and prone to error, especially in informal contexts, and as expected, several clients...
did not know or did not want to share information about income and expenses. Data was collected on personal income (net income from the business as well as other individual sources), household total income and household total expenses (at current values).

A first look at the personal income monthly values reported by the clients in 2017 shows that the sample is heterogeneous:\(^{12}\)

- Female clients reported earning, on average, 7,187PKR less a month than male clients.
- Clients with no formal education also performed worse, with average monthly incomes 3,984PKR lower.
- Clients in Kasur declared incomes, on average, 3,128PKR lower than the other three branches while, inversely, the monthly personal income in Kot Khawaja Saeed was, on average, 5,548PKR higher.
- Entrepreneurs developing sewing/tailoring activities reported earning, on average, 4,765PKR less than those with other activities (trade, services and production).
- Home-owners declared monthly personal incomes, on average, 4,934PKR higher than those clients living with family or renting.

The differences in terms of gender and type of activity fade when looking at the household income and expenditure. The profile is similar when analysing branches, educational level and house ownership sub-groups - household income and expenses are, on average, higher for clients in Kot Khawaja Saeed, clients with higher education and home-owners, and they are comparatively lower for those clients in Kasur and who are illiterate.\(^{13}\)

Table 4 summarizes the results for the 270 households with available data for the three indicators in the two surveys.\(^{14}\)

<table>
<thead>
<tr>
<th></th>
<th>Personal Monthly Income</th>
<th>Household Monthly Income</th>
<th>Household Monthly Expenses</th>
</tr>
</thead>
<tbody>
<tr>
<td>2015</td>
<td>13,380 PKR</td>
<td>24,479 PKR</td>
<td>17,007 PKR</td>
</tr>
<tr>
<td>2017</td>
<td>16,041 PKR</td>
<td>29,377 PKR</td>
<td>17,434 PKR</td>
</tr>
<tr>
<td>Real Variation 2015-17 (%)</td>
<td>+ 19.9%</td>
<td>+ 20%</td>
<td>+ 2.5%</td>
</tr>
</tbody>
</table>

The analysis for the sub-sample indicates an overall improvement of the economic situation of the clients’ households, given that the real variations of average personal and household incomes are significantly higher than of the average household expenditure. Yet again, a more in-depth analysis of the numbers shows diversity of individual situations, including a share of clients within this sub-sample (30%) who appear to have experienced a decrease on their real personal income during the period.

An interesting result when comparing sub-groups of clients is that the variation of personal income in real terms (2015 values) was considerably higher for female clients (100% compared to 36%), albeit the declared earnings continue to be significantly higher for the male clients. This result suggests a step forward to reduce the income gender gap.

\(^{12}\) Differences between the sub-groups were found to be statistically significant at a significance level of 5%.

\(^{13}\) Further details on Annex II.

\(^{14}\) The amounts declared for the indicators in 2017 were deflated to the values of 2015 to eliminate the inflation effect and calculate real variations of the variables.
The expenditure analysis shows a small increase in the household monthly expenditure (in real values), which seems to be accompanied by the growth in the purchase of household assets. 103 clients (23%) reported having bought fixed assets for the household during the previous 12 months, which compares to approximately half of this number (52) for the same question in 2015.

Inversely, savings seem to have generally diminished, at least in terms of frequency. Although most clients save regularly (49.2%) or occasionally (15.4%), the number of clients stating ‘never saving’ increased from 80 to 158, representing 35.3% of the whole sample. For those reporting regular or occasional savings, community savings groups are the main savings channel (77%), with the remaining choosing to save at home or, in a small number of cases, in commercial banks.

Beyond consumption, asset acquisition and savings, 93% of the clients have declared making donations, regularly or occasionally, to AIM. This fits in with the institution’s philosophy of ‘brotherhood’ and the creation of solidarity ties between the clients. Donating to help others in a worse position is a moral obligation for Muslim clients but can also be seen as a ‘guarantee’ that help will be available if an emergency occurs, working in this way as a form of ‘social insurance’.

Parallel to the business outcomes question, in 2017 clients were also asked how the household had changed since the first interview two years previously. Table 5 presents the most common changes explicitly mentioned by the clients.

<table>
<thead>
<tr>
<th>Table 5 – Clients’ Perception on Household Changes</th>
<th>No Respondents</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improvement of living standards</td>
<td>205</td>
<td>46%</td>
</tr>
<tr>
<td>Increase in satisfaction</td>
<td>139</td>
<td>31%</td>
</tr>
<tr>
<td>Increase in income</td>
<td>146</td>
<td>33%</td>
</tr>
<tr>
<td>Increase in consumption</td>
<td>116</td>
<td>26%</td>
</tr>
<tr>
<td>Improvement in education</td>
<td>32</td>
<td>7%</td>
</tr>
<tr>
<td>Increased expenses</td>
<td>33</td>
<td>7%</td>
</tr>
<tr>
<td>No change/‘hand to mouth’ living</td>
<td>25</td>
<td>5%</td>
</tr>
<tr>
<td>Lower income</td>
<td>6</td>
<td>1%</td>
</tr>
</tbody>
</table>

The mentioned changes seem to corroborate other results suggesting (overall) a better situation and an increase in consumption. They also support the idea that not every client has seen their situation changed or improved and, in a small number of cases, normally associated with business failure, the overall situation of the households is worse.

In their answers to the open question some of the clients went beyond generic ideas on living standards and expenses coverage and referred to specific areas like savings (1), house improvements (2), food quality (1), financial independence from donations (1), and financial capacity to marry a child/relative (4).
Part III – Comparison with non-clients

One of the objectives of the project was to evaluate the ‘impact’ of the loans funded by LWC lenders. For this purpose, it was considered critically important to set up a comparison group of non-clients, even acknowledging the challenges associated with its field implementation.

In 2015, with the logistical support of AIM, 100 non-clients with similar businesses and located in the same neighbourhoods as the clients were interviewed. At baseline, the comparative analysis of personal and business indicators between the larger group of clients and the non-clients revealed no significant differences.

The analysis after the second survey is hampered by a low response rate among non-clients, exacerbating the difference in sample size. By the end of May 2017, there were 39 valid interviews with non-clients, none of them from the branch of Kahna Nau. For this reason, a second phase of interviews was conducted in November 2017, focusing on this area. As a result, data from 10 non-clients from Kahna Nau and 3 from Kasur were added to the database.

Mann-Whitney tests were applied to the two sub-groups of clients and non-clients (for the group of variables identified in Annex I). No statistically significant differences were identified regarding most of the variables, and only PPI scores in 2015 were statistically significant at a significance level of 1%. This difference between poverty levels at baseline did not exist when considering the initial group of 100 non-clients. There are, however, no differences between the PPI scores in 2017, or its variation during the period.

Looking at the results for the clients and non-clients available in the second survey, the average PPI scores suggest that AIM clients were more likely to be poor at baseline than the sub-group of non-clients. Furthermore, while, on average, the economic situation of the clients seemed to have improved (as the PPI score increases during the period), the reverse is observed for non-clients.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Client</th>
<th>Average Non-client</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPI2015</td>
<td>63.6</td>
<td>70.5</td>
</tr>
<tr>
<td>PPI2017</td>
<td>66.8</td>
<td>70.0</td>
</tr>
</tbody>
</table>

To further investigate the factors contributing to the changes in the poverty levels (measured by PPI), regression analysis, namely a quantile regression model, was implemented, using the change in poverty index scores as the dependent (explained) variable.\textsuperscript{15} The explanatory variables included age, gender, marital status, household size, location (branch), educational level, working hours/week, PPI score in 2015, business time, type of activity, and loan (access to a AIM business loan).

To apply the quantile regression model, the sample was divided in 4 equal groups based on the PPI scores variation and the correspondent quartiles 1 to 3 were computed (Table 7). The quartiles start from the sub-group of clients with lower PPI variation scores (negative variations) to those who have experienced the stronger positive variations. (Note that higher PPI scores are associated with a lower likelihood of the client being poor, and so, an increase in the PPI score implies a better economic situation).

\textsuperscript{15} The variables in the study do not follow a normal distribution, which requires the implementation of regression models other than the traditional OLS (that relies heavily on the assumption of normality). In this case, quantile regression was applied, allowing not only to overcome the issues regarding the variable behaviour but also to get some further insight on the sample heterogeneity.
The regression results can be consulted in Annex III. They show that, for all quartiles, the variables ‘age’, ‘PPI2015’ (in other words, the poverty level at baseline) and ‘working hours per week’ are able to explain significant variations in the change in PPI scores observed over time, and that the signs of the coefficients are consistent throughout the quartiles (positive for age and working hours, and negative for PPI2015). This means that, on average:

- an increase in the age of the entrepreneur is associated with a higher PPI score, thus a lower probability of being poor;
- an increase in the number of working hours per week has a positive effect on the change in PPI scores during the period;
- there is a negative relationship between the initial poverty level of the entrepreneur and the change in PPI observed over the sample period, with poorer entrepreneurs achieving the greatest increases in PPI scores, potentially reducing the poverty gap.

Interestingly, personal characteristics such as household size, marital status, two of dummy variables for location (Badami Bagh and Kot Khawaja Saeed) and one for education level (secondary school) were found to explain a negative and statistically significant change in the PPI scores for the respondents in the lower quartile. While access to the loan was statistically significant for the respondents in quartile 3 (upper quartile), indicating that the loan at least partly explains the increase in PPI scores observed between 2015 and 2017 for these entrepreneurs.

Gender, business time and type of activity were not statistically significant, at a 5% significance level, for all quartiles.
Conclusion

This report represents a key moment in the LWC evaluation project, which started in 2014-15. It gives an insight into the changes on the lives of AIM clients after they have completed their first microcredit loan. It also represents a first step on what is meant to become a continuous evaluation process aiming to benefit all stakeholders involved – LWC will have evidence from the field to provide to their different supporters; AIM will have relevant information on their clients to inform its management decisions and improve their microcredit programme, and ultimately the clients will benefit from an improved programme that better suits their needs. Moreover, the lessons learnt over the past two years have the potential to be useful not only for the group of AIM clients supported by LWC lenders, but to all clients in similar contexts (which includes the majority of the institution’s branches located in urban/peri-urban areas) as well as to other LWC partner MFIs.

Although two years is a relatively short period, which some authors suggest being insufficient to capture changes in poverty levels,36 the findings from the two waves of the household survey show improvements for a significant percentage of the new AIM clients, both at business and household level. This is reinforced by the initial results of the comparison between clients and non-clients. Despite being a first exploration of the data with methodological limitations associated with the unbalanced sample sizes, the analysis suggests that the microcredit loan is one of the factors explaining the change in PPI scores over time for those who have experienced stronger positive variations (and not being significant for those with negative variations). Further testing and in-depth analysis will be necessary to corroborate these findings and provide explanations on how the participation in the microcredit programme contributed for the outcomes.

It is also important to stress that these positive outcomes should not be taken for granted. There was a share of the interviewed clients who have experienced a deterioration of their economic situation; hence, understanding the causes for their declining performance would be important for AIM to better identify risk factors and minimize both the volume of these situations and their consequences.

A final note would be to highlight the clients’ perception on the changes they experienced over this period. Even given the interviewees’ time restrictions, not allowing for the full development of answers, the general feeling after completing the first loan was of satisfaction, and it is particularly interesting that, at least a part of these clients, acknowledge they have overcome some of the barriers in their access to finance.

References


36 See Chen et al. (2010)
Annex I

List of quantitative variables tested using Mann-Whitney Tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Age in 2017 (estimated from survey 2015)</td>
<td>No.</td>
</tr>
<tr>
<td>Business Time</td>
<td>Years owning the business in 2017 (survey 2017)</td>
<td>No.</td>
</tr>
<tr>
<td>Working Hours</td>
<td>Average hours/week in 2017 from original data on ‘working hours/day’ and ‘working days/week’ (survey 2017)</td>
<td>No.</td>
</tr>
<tr>
<td>Loan Amount</td>
<td>First loan amount (survey 2015)</td>
<td>PKR</td>
</tr>
<tr>
<td>PPI_2015</td>
<td>PPI score at baseline (survey 2015)</td>
<td>No.</td>
</tr>
<tr>
<td>PPI_2017</td>
<td>PPI score in 2017 (survey 2017)</td>
<td>No.</td>
</tr>
<tr>
<td>PInc_2017</td>
<td>Monthly personal income in 2017, current values (survey 2017)</td>
<td>PKR</td>
</tr>
<tr>
<td>HHInc_2017</td>
<td>Monthly household total income in 2017, current values (survey 2017)</td>
<td>PKR</td>
</tr>
<tr>
<td>HHExp_2017</td>
<td>Monthly household total expenses in 2017, current values (survey 2017)</td>
<td>PKR</td>
</tr>
<tr>
<td>Weight PI/HHI_2017</td>
<td>Proportion of personal income in the household total income in 2017, current values (survey 2017)</td>
<td>No.</td>
</tr>
<tr>
<td>PPI Variation</td>
<td>Percent difference between PPI scores in 2015 and 2017 (survey 2015 and survey 2017)</td>
<td>%</td>
</tr>
<tr>
<td>Total Employment</td>
<td>Absolute difference between total employment in 2015 and 2017 (survey 2017)</td>
<td>No.</td>
</tr>
<tr>
<td>Absolute Variation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personal Income Variation</td>
<td>Percent difference between personal income in 2015 and 2017, both at 2015 values (survey 2015 and survey 2017)</td>
<td>%</td>
</tr>
<tr>
<td>Household Income Variation</td>
<td>Percent difference between household income in 2015 and 2017, both at 2015 values (survey 2015 and survey 2017)</td>
<td>%</td>
</tr>
<tr>
<td>Household Expenses Real Variation</td>
<td>Percent difference between household expenses in 2015 and 2017, both at 2015 values (survey 2015 and survey 2017)</td>
<td>%</td>
</tr>
</tbody>
</table>

PKR: Pakistani Rupees
The analysis was carried out for all the variables included in Annex I. Mann-Whitney tests were used to test the null hypothesis of the distribution of the selected variables being equal for defined sub-groups. The sub-groups are associated with clients’ characteristics and were chosen considering the results from the baseline survey and qualitative information collected throughout the project. In order to apply the tests, dummy variables were constructed for:

- Gender (female)
- Location (Badami Bagh, Kot K. Saeed, Kahna Nau, Kasur)
- Educational level (illiterate – no formal education; primary school – grades 1 to 5; secondary school – middle and matric levels; higher education – intermediate level and university)
- Type of activity (trade; production and construction; services and education; sewing/tailoring)
- House ownership (home owner – with or without property title)
- Programme membership (client).

The table below shows the results statistically significant at a significance level of 5% (***) or 1% (**), demonstrating the existence of differences between the sub-groups for the studied variable. As guidance in reading the table, here is an example regarding location:

- Clients from Badami Bagh branch were found to be different from clients in the other three branches in terms of the loan amount with clients in Badami Bagh accessing lower amounts, on average (19,122 PKR compared with 20,155 PKR).

### Gender

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Female</th>
<th>Average Male</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPI Variation</td>
<td>0.12</td>
<td>0.08</td>
<td>20,336 **</td>
</tr>
<tr>
<td>Real PI Variation</td>
<td>1.01</td>
<td>0.36</td>
<td>13,415 ***</td>
</tr>
<tr>
<td>Loan Amount</td>
<td>19,750</td>
<td>20,120</td>
<td>20,013.5 ***</td>
</tr>
<tr>
<td>Age</td>
<td>41.5</td>
<td>37.8</td>
<td>17,571 ***</td>
</tr>
<tr>
<td>Working Hours/Week</td>
<td>60</td>
<td>67</td>
<td>11,118 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>14,971</td>
<td>22,158</td>
<td>10,061 ***</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>60.1</td>
<td>65.7</td>
<td>18,538.5 ***</td>
</tr>
</tbody>
</table>

### Location

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Badami Bagh</th>
<th>Average Other branches</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Employment Var.</td>
<td>0.07</td>
<td>0.53</td>
<td>10,477.5 ***</td>
</tr>
<tr>
<td>HH Income 2017</td>
<td>38,739</td>
<td>30,495</td>
<td>3,651 ***</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>23,085</td>
<td>18,039</td>
<td>8,405.5 ***</td>
</tr>
<tr>
<td>Loan Amount</td>
<td>19,122</td>
<td>20,155</td>
<td>10,735 ***</td>
</tr>
<tr>
<td>Working Hours/Week</td>
<td>57</td>
<td>66</td>
<td>6,561.5 ***</td>
</tr>
</tbody>
</table>
### LWC – AIM REPORT 2018

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Kot K. Saeed</th>
<th>Average Other branches</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Employment Var.</td>
<td>0.94</td>
<td>0.32</td>
<td>10,869.5 ***</td>
</tr>
<tr>
<td>HH Income Real Var.</td>
<td>0.76</td>
<td>0.29</td>
<td>4,178.5 **</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>75.3</td>
<td>60.7</td>
<td>8,213.5 ***</td>
</tr>
<tr>
<td>PPI 2017</td>
<td>76.2</td>
<td>64.5</td>
<td>9,165 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>23,944</td>
<td>18,396</td>
<td>10,849 ***</td>
</tr>
<tr>
<td>HH Income 2017</td>
<td>47,622</td>
<td>29,120</td>
<td>2,968 ***</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>22,242</td>
<td>17,969</td>
<td>11,904.5 ***</td>
</tr>
<tr>
<td>Variable</td>
<td>Average Kahna Nau</td>
<td>Average Other branches</td>
<td>U statistic</td>
</tr>
<tr>
<td>Total employment Var.</td>
<td>-0.01</td>
<td>0.57</td>
<td>12,968 ***</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>58.3</td>
<td>65.3</td>
<td>13,827 ***</td>
</tr>
<tr>
<td>PPI 2017</td>
<td>60.7</td>
<td>68.8</td>
<td>12,848 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>18,240</td>
<td>19,959</td>
<td>13,964 **</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>20,683</td>
<td>18,296</td>
<td>12,828.5 ***</td>
</tr>
<tr>
<td>Variable</td>
<td>Average Kasur</td>
<td>Average Other branches</td>
<td>U statistic</td>
</tr>
<tr>
<td>HH Income Real Var.</td>
<td>0.23</td>
<td>0.46</td>
<td>9,838.5 ***</td>
</tr>
<tr>
<td>HH Expenses Real Var.</td>
<td>0.06</td>
<td>0.22</td>
<td>16,098.5 ***</td>
</tr>
<tr>
<td>Business Time</td>
<td>10.2</td>
<td>9.8</td>
<td>20,474.5 **</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>59.5</td>
<td>66.3</td>
<td>18,423 ***</td>
</tr>
<tr>
<td>PPI 2017</td>
<td>63.5</td>
<td>69.0</td>
<td>19,321 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>17,694</td>
<td>20,822</td>
<td>17,927.5 ***</td>
</tr>
<tr>
<td>HH Income 2017</td>
<td>25,908</td>
<td>36,430</td>
<td>8,561 ***</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>14,254</td>
<td>21,862</td>
<td>9,991.5 ***</td>
</tr>
</tbody>
</table>

### Educational Level

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Illiterate</th>
<th>Average Other Education</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPI Variation</td>
<td>0.13</td>
<td>0.06</td>
<td>21,641 **</td>
</tr>
<tr>
<td>Age</td>
<td>41.2</td>
<td>37.2</td>
<td>18,332.5 ***</td>
</tr>
<tr>
<td>Business Time</td>
<td>10.4</td>
<td>9.5</td>
<td>21,838.5 **</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>57.9</td>
<td>68.8</td>
<td>15,586.5 ***</td>
</tr>
<tr>
<td>PPI 2017</td>
<td>62.5</td>
<td>70.7</td>
<td>17,657.5 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>17,435</td>
<td>21,469</td>
<td>16,886 ***</td>
</tr>
<tr>
<td>HH Income 2017</td>
<td>28,535</td>
<td>34,815</td>
<td>10,504.5 ***</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>16,989</td>
<td>20,578</td>
<td>19,459.5 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Secondary School</th>
<th>Average Other Education</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPI Variation</td>
<td>0.03</td>
<td>0.13</td>
<td>17,116 ***</td>
</tr>
<tr>
<td>HH Expenses Real Var.</td>
<td>0.29</td>
<td>0.09</td>
<td>14,826.5 ***</td>
</tr>
<tr>
<td>Age</td>
<td>36.7</td>
<td>40.5</td>
<td>15,799 ***</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>70.9</td>
<td>60.0</td>
<td>13,058.5 ***</td>
</tr>
<tr>
<td>PPI 2017</td>
<td>71.2</td>
<td>64.7</td>
<td>15,775.5 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>22,755</td>
<td>18,007</td>
<td>14,425 ***</td>
</tr>
<tr>
<td>HH Income 2017</td>
<td>36,349</td>
<td>29,635</td>
<td>8,525 **</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>21,487</td>
<td>17,684</td>
<td>14,744.5 **</td>
</tr>
</tbody>
</table>
## LWC – AIM REPORT 2018

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Higher Education</th>
<th>Average Other Education</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH Inc. Real Variation</td>
<td>1.08</td>
<td>0.33</td>
<td>700 ***</td>
</tr>
<tr>
<td>Loan Amount</td>
<td>20,933</td>
<td>19,946</td>
<td>2,381 **</td>
</tr>
<tr>
<td>Age</td>
<td>32.7</td>
<td>39.5</td>
<td>1,903.5 ***</td>
</tr>
<tr>
<td>HH Size</td>
<td>5.2</td>
<td>6.5</td>
<td>2,052 **</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>80.9</td>
<td>62.9</td>
<td>1,154 ***</td>
</tr>
<tr>
<td>PPI 2017</td>
<td>82.6</td>
<td>66.2</td>
<td>1,376 ***</td>
</tr>
<tr>
<td>HH Income 2017</td>
<td>45,000</td>
<td>31,002</td>
<td>805 ***</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>25,400</td>
<td>18,665</td>
<td>21,205 **</td>
</tr>
</tbody>
</table>

### Type of Business

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Services</th>
<th>Average Other Activity</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working Hours/Week</td>
<td>70</td>
<td>63</td>
<td>5,854 **</td>
</tr>
</tbody>
</table>

### Average Tailoring

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Sewing/Tailoring</th>
<th>Average Other Activity</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Employment Var.</td>
<td>1.1</td>
<td>0.2</td>
<td>13,764 ***</td>
</tr>
<tr>
<td>HH Expenses Real Var.</td>
<td>0.03</td>
<td>0.21</td>
<td>14,685 **</td>
</tr>
<tr>
<td>Loan Amount</td>
<td>19,792</td>
<td>20,044</td>
<td>17,815.5 **</td>
</tr>
<tr>
<td>Working Hours/Week</td>
<td>60</td>
<td>66</td>
<td>10,002.5 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>16,085</td>
<td>20,850</td>
<td>12,189 ***</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>16,896</td>
<td>19,618</td>
<td>16,031.5 ***</td>
</tr>
</tbody>
</table>

### Average Trade

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Trade</th>
<th>Average Other Activity</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>HH Expenses Real Var.</td>
<td>0.33</td>
<td>0.09</td>
<td>14,149 **</td>
</tr>
<tr>
<td>Working Hours/ Week</td>
<td>68</td>
<td>63</td>
<td>10,007 **</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>20,902</td>
<td>18,958</td>
<td>15,083 ***</td>
</tr>
<tr>
<td>HH Income 2017</td>
<td>34,718</td>
<td>30,243</td>
<td>8,669 **</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>20,113</td>
<td>18,334</td>
<td>16,788.5 **</td>
</tr>
</tbody>
</table>

### Average Production

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Production</th>
<th>Average Other Activity</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loan Amount</td>
<td>20,486</td>
<td>19,801</td>
<td>15,646 ***</td>
</tr>
<tr>
<td>Business Time</td>
<td>11.3</td>
<td>9.4</td>
<td>14,854 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>21,908</td>
<td>20,850</td>
<td>14,730 **</td>
</tr>
</tbody>
</table>

### House Ownership

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average House Owner</th>
<th>Average Non-house owner</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Employment Var.</td>
<td>0.7</td>
<td>0.3</td>
<td>18,185 **</td>
</tr>
<tr>
<td>HH Income Real Var.</td>
<td>0.59</td>
<td>0.27</td>
<td>6,878.5 ***</td>
</tr>
<tr>
<td>HH Expenses Real Var.</td>
<td>0.24</td>
<td>0.12</td>
<td>16,107 **</td>
</tr>
<tr>
<td>Age</td>
<td>40.6</td>
<td>38.4</td>
<td>15,217.5 **</td>
</tr>
<tr>
<td>Business Time</td>
<td>9.8</td>
<td>10.2</td>
<td>18,870.5 **</td>
</tr>
<tr>
<td>PPI 2015</td>
<td>69.0</td>
<td>61.1</td>
<td>16,019.5 ***</td>
</tr>
<tr>
<td>PPI 2017</td>
<td>71.9</td>
<td>64.4</td>
<td>15,666.5 ***</td>
</tr>
<tr>
<td>Personal Income 2017</td>
<td>22,892</td>
<td>17,956</td>
<td>14,253.5 ***</td>
</tr>
<tr>
<td>HH Income 2017</td>
<td>38,103</td>
<td>29,320</td>
<td>7,313 ***</td>
</tr>
<tr>
<td>HH Expenses 2017</td>
<td>21,198</td>
<td>17,722</td>
<td>16,374.5 ***</td>
</tr>
</tbody>
</table>
### Annex III – Quantile Regression

Dependent Variable: PPI Variation  
Method: Quantile Regression  
Sample: 499  
Included observations: 383

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Type</th>
<th>Quartile 1 (0.25)</th>
<th>Quartile 2 (0.50)</th>
<th>Quartile 3 (0.75)</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (2017)</td>
<td>Q</td>
<td>0.006200 ***</td>
<td>0.006694 ***</td>
<td>0.009906 ***</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>D</td>
<td>0.053218</td>
<td>0.006750</td>
<td>0.031250</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>D</td>
<td>-0.110948 **</td>
<td>-0.041616</td>
<td>-0.093870 *</td>
<td></td>
</tr>
<tr>
<td>Badami Bagh Branch</td>
<td>D</td>
<td>0.089220 ***</td>
<td>0.046550</td>
<td>0.100315 **</td>
<td></td>
</tr>
<tr>
<td>Kahna Nau Branch</td>
<td>D</td>
<td>0.012617</td>
<td>-0.006522</td>
<td>0.017976</td>
<td></td>
</tr>
<tr>
<td>Kot Khawaja Saeed</td>
<td>D</td>
<td>0.157151 ***</td>
<td>0.080715 *</td>
<td>0.074495</td>
<td></td>
</tr>
<tr>
<td>Trade</td>
<td>D</td>
<td>0.028370</td>
<td>0.037280</td>
<td>0.020908</td>
<td></td>
</tr>
<tr>
<td>Services and Education</td>
<td>D</td>
<td>-0.016557</td>
<td>-0.038621</td>
<td>0.011931</td>
<td></td>
</tr>
<tr>
<td>Production and Construction</td>
<td>D</td>
<td>0.034542</td>
<td>0.040377</td>
<td>0.002536</td>
<td></td>
</tr>
<tr>
<td>Primary School</td>
<td>D</td>
<td>0.067060 *</td>
<td>0.025416</td>
<td>0.032394</td>
<td></td>
</tr>
<tr>
<td>Secondary School</td>
<td>D</td>
<td>0.076846 **</td>
<td>0.042123</td>
<td>0.000686</td>
<td></td>
</tr>
<tr>
<td>Higher Education</td>
<td>D</td>
<td>0.046482</td>
<td>-0.000268</td>
<td>-0.004743</td>
<td></td>
</tr>
<tr>
<td>Household Size</td>
<td>Q</td>
<td>-0.016481 ***</td>
<td>-0.005000</td>
<td>-0.001460</td>
<td></td>
</tr>
<tr>
<td>Business Time (2017)</td>
<td>Q</td>
<td>0.002365</td>
<td>-0.000742</td>
<td>-0.001830</td>
<td></td>
</tr>
<tr>
<td>Working Hours/Week (2017)</td>
<td>Q</td>
<td>0.001761 **</td>
<td>0.002213 ***</td>
<td>0.002754 **</td>
<td></td>
</tr>
<tr>
<td>PPI_2015</td>
<td>Q</td>
<td>-0.006205 ***</td>
<td>-0.005910 ***</td>
<td>-0.007338 ***</td>
<td></td>
</tr>
<tr>
<td>Loan</td>
<td>D</td>
<td>0.046056</td>
<td>0.054611</td>
<td>0.198767 ***</td>
<td></td>
</tr>
</tbody>
</table>

Notes:  
Q = quantitative variable, D = dummy variable (1,0)  
*** significance level of 1%, ** significance level of 5%, * significance level of 10%
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of contents</td>
<td>2</td>
</tr>
<tr>
<td>Summary</td>
<td>3</td>
</tr>
<tr>
<td>Introduction</td>
<td>5</td>
</tr>
<tr>
<td>Part I – The clients and their businesses</td>
<td>6</td>
</tr>
<tr>
<td>Part II – The changes after the first loan cycle</td>
<td>9</td>
</tr>
<tr>
<td>The business outcomes</td>
<td>9</td>
</tr>
<tr>
<td>The household outcomes</td>
<td>11</td>
</tr>
<tr>
<td>Part III – The comparison with non-clients</td>
<td>14</td>
</tr>
<tr>
<td>Conclusions and Recommendations</td>
<td>17</td>
</tr>
<tr>
<td>References</td>
<td>18</td>
</tr>
<tr>
<td>Annexes</td>
<td>19</td>
</tr>
<tr>
<td>List of variables</td>
<td>19</td>
</tr>
<tr>
<td>Non-parametric tests</td>
<td>20</td>
</tr>
<tr>
<td>Regression results</td>
<td>24</td>
</tr>
</tbody>
</table>
Summary

This report is part of LENDWITHCARE (LWC) assessment project and focuses on the evaluation of LWC partner in Zimbabwe, THRIVE Microfinance. The report was prepared by the University of Portsmouth (UoP), partner in the project, after a second wave of a household survey to a sample of THRIVE clients who have been supported by the LWC crowdfunding platform.

The study sample includes 341 new THRIVE clients and 157 non-clients, first interviewed in 2016 (April to June) by a team of interviewers recruited from a local university. The second wave of interviews took place approximately one year later (June to August), when some of the clients were starting to repay their third loan. 245 clients and 110 non-clients were available to be interviewed.

The report offers an initial snapshot of how the lives of LWC supported entrepreneurs have changed since they have become THRIVE clients. While reading the report, it is important to take into account the exceptional economic and political circumstances in the country during this period. On May 2016, the Reserve Bank of Zimbabwe announced the launch of an internal currency to cope with the cash crisis faced by the country. The introduction of bond notes, although reducing the immediate effects of the lack of cash in the daily economic operations, did not solve the problem and created a double standard for the currency, with bond dollars rapidly depreciating to the US dollar, and instigating fears of a new hyperinflationary period. The situation has brought also political consequences culminating in the military intervention on November 2017.

From THRIVE Microfinance and its clients’ perspective, the deterioration of the economic situation and the cash crisis impacted businesses and daily lives. It became more difficult to pay, get paid or undertake any type of monetary transaction, forcing the institution and its clients into the ‘digital sphere’. As consequence, loan disbursements as well as repayment of loan instalments are now almost totally done by mobile payments.

The survey results show that THRIVE/LWC clients - who are all female - are mainly married (72%) and educated (91% attended secondary education or above). They live mostly in family houses (44%), in households with an average of 4.8 members. The average age is 42 years old and they have relevant business experience (8 years). 59% work alone in the business and they report working an average of 69h/week. Their first loan was, on average, $334 and was destined mostly to fund working capital. By the time of the second interview, 29% of the clients were on their third loan cycle, 35% had finished repaying a second loan, and 37% had completed repaying the first loan.

Looking at the overall results, some indicators seem to reflect a deterioration in living standards during the year. There was an increase in the number of clients reporting an external shock during the previous 12 months (from 29% to 44%), as well as health problems in the household (from 10% to 16%). Consumption declined, with a lower number of clients buying household fixed assets or making house improvements, and equally, Income and expenditure indicators, including average personal and household incomes and average household expenses, decreased during the period.

Mixed results were also obtained from the analysis of the variation of THRIVE-PAT (poverty assessment tool). Half of the clients (50%) have seen their household total score increase between 2016 and 2017, implying an improvement in the economic circumstances of their households, but there was a relevant percentage (42%) who have experienced a deterioration of the score.

Despite the difficulties, 81% of the clients reported higher or similar sales compared with the previous year (before receiving THRIVE loan), and 64% considered their quality of life to be overall better. Looking at the same survey questions for the comparison group, only 47% of the sample non-clients declared business revenues to be growing or stable and 29% considered their life to be better at the time of the second survey.
Moreover, there was a significant change on the savings practices among clients, even in a short period of time. 80% reported saving regularly, mainly using informal savings groups (ROSCAs). The development of a savings culture is one of THRIVE’s objectives and it is an important component of the training provided by the institution, which was highly valued by clients (99% considered it to be useful or very useful). Besides the training, there were two other institutional factors much appreciated by the clients: flexible repayment conditions and customer service. 77% of the clients stated that they would be willing to maintain their relationship with THRIVE and apply for future loans. Finally, the survey shows that clients in general had a positive attitude and demonstrated resilience in face of adversity.

**Recommendations:**

The two surveys implemented offer valuable data on the changes observed in the lives of THRIVE clients during the period of analysis. Household surveys, *per se*, do not provide definite explanations for the changes, but they allow for the identification of correlations between different factors and their results prompt some suggestions regarding the programme and the continuation of the evaluation project:

- The aggravated economic and political situation in Zimbabwe between the two surveys, and the uncertainty associated with this context, make it more relevant the continuation of the evaluation project and the possibility of repetition of the survey after a more stable period.
- Training was highly appreciated by the clients, but some of them complained about the number of sessions, implying opportunity costs associated with the time spent on the sessions and longer loan approval processes. Considering also the costs of training for the institution, the suggestion will be to reduce the number of sessions. Complementary actions may be used to reinforce the main messages to the clients, including the use of short phone messages.
- The report gives some detail on the existence of differences between the sample clients. These can be further explored to identify client segments with distinct needs in terms of training and prepare specific sessions. Among the potential training topics is digital literacy aiming to help clients adapt to the mobile money economy emergent from the cash crisis.
- The comparison between clients and non-clients included in the report alerts for the existence of a segment of poorer female entrepreneurs who are not being fully served by THRIVE. If the social mission of the institution includes reaching the poorest segments of women entrepreneurs, then it will be necessary to better understand the reasons for this exclusion and identify the best ways to respond to the needs of these potential clients, which may require changes in the main loan product or the design of new financial or non-financial products and services.
Introduction

LENDWITHCARE assessment project started in 2014 with the main objective of assessing the experience of the borrowers supported by the crowdfunding platform. By the end of 2014, Akhuwat in Pakistan had been chosen as the first field partner to participate in the evaluation, and the University of Portsmouth became also a partner in the project. In 2016, THRIVE Microfinance in Zimbabwe joined the project.

THRIVE Microfinance is a private financial company that started operations in 2012 offering business loans to low income entrepreneurs. The institution is among the few in the country with an explicit social mission, with its microcredit programme following a microfinance plus approach with a gender perspective. The MFI targets only female entrepreneurs, whose access to the loans is dependent on the formation of a group with 3 to 10 members, the completion of a training programme and compulsory savings.

THRIVE has been LWC partner since August 2015. By 31st October 2017, 553 (group) loans reaching 1,820 entrepreneurs have been funded, corresponding to a lending amount over $850k. The institution has five branches – Willowvale (Harare), Chitungwiza, Whitecliff (Harare), Juru and Marondera. The last three opened during 2017 and, for that reason, are not included in the study.

During the months of April to June 2016, a sample of 341 THRIVE new clients and 157 non-clients were interviewed by a team of independent interviewers recruited from a group of Bulawayo University finalists. The interviewers received training from the UoP/LWC team to apply the questionnaire, which was prepared by UoP with the support of LWC and THRIVE. The baseline survey allowed for the characterisation of the respondents and established the basis for comparative analysis in 2017, after the implementation of a second wave of the survey.

The second questionnaire was to be applied to the same clients and non-clients. The period between interviews was little over one year, although by this time some of the clients had already been granted a third loan due to the institution’s short loan cycle. The second survey included some questions identical to the first questionnaire, new queries about the changes in a selection of outcomes (at business and household level), and additional open questions aiming to capture the perception of the interviewees regarding the observed changes. In both surveys, data were collected to compute THRIVE-PAT, an internally developed poverty assessment tool.

Longitudinal studies present several challenges, one of the most significant being attrition. In the context of the aggravated economic situation in the country, the expectation was of increased difficulties to locate and interview some of the participants, especially those who have experienced negative evolutions of their businesses. The final dataset included valid data for the two periods for 245 clients and 110 non-clients. The geographical distribution of the survey sample is shown in Table 1.

<table>
<thead>
<tr>
<th>Branch</th>
<th>Clients 2015</th>
<th>Clients 2017</th>
<th>Non-clients 2015</th>
<th>Non-clients 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Willowvale</td>
<td>180</td>
<td>133</td>
<td>117</td>
<td>89</td>
</tr>
<tr>
<td>Chitungwiza</td>
<td>161</td>
<td>112</td>
<td>40</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>341</td>
<td>245</td>
<td>157</td>
<td>110</td>
</tr>
</tbody>
</table>

1 All figures and tables included in the report are based on survey data.
2 THRIVE-PAT is a poverty scorecard developed internally by the institution in the absence of externally ratified poverty scorecards for Zimbabwe.
3 Interviews took place between June and August 2017. 3 respondents included in the client baseline sample did not receive their loans as their group disintegrated after signing the contract. These entrepreneurs were re-classified as non-clients in the data analysis.
Part I - The clients and their businesses

All THRIVE clients are women. The average age of the clients in the sample is 42 years. Figure 1 illustrates the distribution of the clients by age groups and the concentration on the segments between 35 and 45 years. The focus is, therefore, on experienced entrepreneurs, which reflects also on the average business time (8 years for those with active businesses in 2017).

![Figure 1 - Clients Age Distribution 2017](image)

A common characteristic to most clients is their education level. 84% of the clients has completed 9 or 11 years of schooling (O-Level); a reduced number has progressed to A-Level and higher studies (7%), and similarly, a relatively small part of the clients has just attended primary school (9%). Data from the Census 2012 shows high literacy rates in Zimbabwe generally (98% in Harare), as well as a high percentage of students completing secondary or above education levels (68% for Harare).

Other common feature to most clients is their marital status – 72% are married. There is, however, a relevant number of widows (15%), white divorced/separated and single comprise 8% and 5% of the sample, respectively. The sample households have, on average, 4.8 members; a figure which is slightly higher than the data for the Harare province (3.9) in the 2012 Census. On average, the number of active workers in the sample households is 1.9. In 73 cases (80%), the client is the only source of income for the household.

House tenure is diversified in the sample. 44% of the clients live in houses owned by family members, 26% are home owners, and 27% live in rented property. The interesting result, most likely linked to the economic context, is the decrease observed in both the number of clients renting and living in their own houses (29% and 31% in 2016, respectively) in favour of living with family.

Decision making regarding how the household income is spent and how household activities are organised did not change significantly in the (short) period: 37% of the clients make decisions related to the household income on their own; 66% decide on household activities independently. Coincidentally, 37% is also the proportion of clients who hold leadership positions in community organizations. These organizations are in most cases linked to religious groups.

* The question regarding house tenure distinguishes between clients with property rights (either with or without a formal property title) from those clients living in houses owned by their husbands but where they have no ownership right. Those latter cases (60) were considered as living with family in both surveys.
The majority of THRIVE clients run activities in the trade sector (62%). As Figure 2 illustrates, among them, 17% are cross border traders, buying and selling different products (mainly clothes, footwear and food) in the neighbour countries (South Africa, Zambia, Mozambique and Botswana). As could be anticipated, the currency crisis and changes in the requirements at the South African border led to a decrease of the number of entrepreneurs in this type of activity, which represented 22% of the sample in 2016. Between the two surveys, 29% of the clients have changed their primary activity. No significant differences were identified between the two branches regarding the type of activities conducted by the clients.

![Figure 2 - Business Activities in 2017](image)

59% of the clients stated working alone in the business in 2017. Those who have employees report providing paid jobs to a total of 60 persons and occupation to 64 additional unpaid employees. In the 2017 survey, date was also collected on the working time of the entrepreneurs. For the 190 THRIVE clients who replied to the question, the average working load was 60 hours/week. Only 29 entrepreneurs declared working less than 40 hours/week, which suggests that part-time dedication to the business is clearly not predominant, despite an expectation that most women accumulate business and household tasks.

Given the short duration of THRIVE loans (typically 6 months), by the time of the second survey the clients who consecutively renewed their loans were already in their third loan cycle. This was the case for 71 clients, while 85 had completed their second loan, and 89 had finished repaying the first loan and did not receive further loans by THRIVE (Figure 3). On each loan cycle, the average loan has increased slightly from $334 on the first loan to $435 and $540 on the second and third loans, respectively. As is often the case with simple averages, this increase disguises different client funding strategies. The loan amounts range from $200 to $1,000, with a part of the clients opting to apply for similar or even lower amounts in the follow-up loans.

![Figure 3 - Client Distribution for Loan Cycle](image)
It is interesting to note that, independently of the number of loans received, 77% of the clients expressed their desire to apply for future loans at THRIVE. They present two main motivations. One is related to the business financial needs, which was mentioned by those clients who seek to continue funding working capital (30%), as well as by those who plan to expand their business or venture into new businesses (12%), albeit adverse economic conditions. Curiously, only three clients explicitly stated they were waiting for improvements in the economic situation to apply for further loans. The second main group of reasons, expressed by 45% of the clients, is associated with the institution performance and includes flexible repayment schedules, lower interest rates, training and good customer service.

As in the first wave of the survey, there was a low number of clients declaring other active loans (5%). Only 4 clients had successfully applied for loans in other MFI’s or Banks, and none reported resorting to moneylenders. These results seem to corroborate Global Findex data for 2014, which shows that informal lending (5%) and formal financial institutions (4%) are a resource used by a minority, family and friends were the main borrowing source.

Of significance is that almost one third of the sample (31%) has admitted to facing repayment problems at some stage. In a few cases (7%), illness of the client or a close relative led to less time dedicated to the business or increased medical expenses, but in most cases the reasons presented for defaulting were linked (directly or indirectly) to the economic situation and the cash crisis:

- Not enough sales or difficulties receiving timely from clients;
- Money or stocks stolen, or stocks confiscated by customs agents;
- Need to repay for other members of the group;
- Misappropriation of the instalment money by the treasurers of the group, or others in charge of making the repayment in the name of the client.

The difficult context seems also to be reflected in the survey on the higher proportion of clients that declared suffering an external shock during the previous year (44% of the client sample compared with 29% in 2016), and the prevalence of health problems in the households (28% compared with 20%). Death and illness of relatives are the main cause of external shocks (78%).

On a positive note, the training provided by THRIVE is undoubtedly one of its distinctive features and, in the two waves of the survey, it is possible to understand that it is widely appreciated by the clients. At baseline, clients had just finished the compulsory initial training, with 99% and 91% evaluating positively the contents and format of the training, respectively. The ones not satisfied with the format complained mostly about the long duration of the training, extending the loan approval time.

In the second questionnaire, clients were asked if they had received further training from the institution and how they globally evaluated the training. 66% had received further training, and 99.5% considered it useful or very useful. The appreciation of the training provided by THRIVE is also noticeable in the qualitative data collected through the survey open questions, with several clients linking their decisions or performance to the lessons learnt during the training sessions.

---

2 Data from the 2014 Global Findex survey [http://data.worldbank.org/financi-allowance/country/zimbabwe]. Adults (15 years) total borrowing during the previous year, independently of mode and channel, was 0.4%, a figure considerably higher compared with the average for Sub-Saharan Africa (5.5%) and Low-Income Countries (5.5%).
Part II – The Changes after 1 year

From a theoretical perspective, links can be established between microfinance and a wide range of socio-economic and well-being indicators. However, considering the logistical challenges associated with the project, only a few of the potential outcomes were selected to ensure manageable short questionnaires. The selection was made considering relevant impact studies in the sector, and the comments and suggestions from the field partner. The questionnaires included both business and household level outcomes.

In the interpretation of the results, it is important to keep in mind that a longitudinal survey of this nature does not give definite explanations for the identified changes but provides critical data on the lives of microcredit clients and the existence of relations between the factors studied. In this sense, it can become the foundation of a more in-depth analysis, and an important incentive to establish continuous evaluation processes in the MFIs involved in the project.

Business outcomes

Microcredit clients run small informal businesses, often with no organized accounting, making it more challenging to collect quantitative data on the businesses. The decision was to include in the questionnaire a qualitative appreciation of sales growth for the previous 12 months in both surveys, which enables the comparison of the sales/revenues level at the time of the two interviews.

The results are overall positive, with 51% of those with active businesses at the time of the second survey declaring growing sales, and further 39% reporting similar sales compared with the previous year. There are, notwithstanding, a group of clients for whom the business year was not positive, resulting in decreasing sales and business closures as illustrated in Figure 4.

![Figure 4 - Comparison Sales Revenues in 2017 vs Baseline (No. Clients)](image)

<table>
<thead>
<tr>
<th>Category</th>
<th>No. Clients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Growing</td>
<td>117</td>
</tr>
<tr>
<td>Similar</td>
<td>65</td>
</tr>
<tr>
<td>Decreasing</td>
<td>42</td>
</tr>
<tr>
<td>Closed Business</td>
<td>17</td>
</tr>
</tbody>
</table>

The majority of businesses in the sample (53.3%) managed to maintain the same level of employment (in many cases, the client’s own job), and 29.5% hired additional workers during the period. This was not enough, however, to compensate for those clients closing their businesses (7%) or reducing staff (19%), which resulted in a small net decrease of the total sample employment (~8 workers), and a slight increase in the proportion of clients working alone (from 54% in 2016 to 59%). It is interesting to note that, while most cases of business closure refer to situations of business failure, there are two clients whose sales were stable, but decided to shut down their businesses because they managed to secure formal jobs.

---

5 For further development see (Devarakonda et al. 2013), O’Hall (2015) and Banerjee et al. (2006).
6 Contributing also for the difficulties measuring business outcomes is the fact that it is common for the clients to change activity, or accumulate multiple activities, to adapt to the needs of the markets and maintain the household income level. In the characterization of the businesses, in both questionnaires was considered the main activity as declared by the clients.
The questionnaires included an open question on the constraints for the businesses to grow. ‘Missing capital’ was the most frequent constraint identified in both years, but answers were, in general, more detailed in the second survey. Figure 5 shows the different needs identified by the clients in the 2017 survey, for their businesses to survive and be able to grow.

![Figure 5 - Business Needs in 2017](image)

There were naturally more clients whose businesses were growing or stable reporting that they did not have any constraints and their businesses were performing well. In general, all clients provided detailed answers, with the economic conditions being mentioned by a more significant number of clients (compared with baseline), independently of their sales performance.

In Figure 5, the category ‘other’ includes those with no business (17), with four of them stating they need capital to re-start their business or venture into a different activity. It includes also three clients who refer to the need of changing their business strategy, increasing the time dedicated to the business or changing the rules to provide credit to their clients in order to reduce the number of debtors.

The changes at business level were also included in the questionnaire through a second open question, aiming to understand the perception of the entrepreneurs on the changes and how they valued them. Table 2 shows the most common answers, which are not mutually exclusive and reflect mixed feelings regarding the business experience during the period.

<table>
<thead>
<tr>
<th>Change in Business</th>
<th>No Respondents</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diversification or Change of Activities</td>
<td>60</td>
<td>24%</td>
</tr>
<tr>
<td>Increase in Sales and/or Profits</td>
<td>99</td>
<td>40%</td>
</tr>
<tr>
<td>Increase in Number of Clients</td>
<td>19</td>
<td>8%</td>
</tr>
<tr>
<td>Increase in Business Assets</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>Increase in Number of Employees</td>
<td>6</td>
<td>2%</td>
</tr>
<tr>
<td>Working More Hours</td>
<td>10</td>
<td>4%</td>
</tr>
<tr>
<td>Difficulties Receiving from Clients</td>
<td>12</td>
<td>5%</td>
</tr>
<tr>
<td>External Shocks</td>
<td>4</td>
<td>2%</td>
</tr>
<tr>
<td>Business Closure</td>
<td>17</td>
<td>7%</td>
</tr>
<tr>
<td>Decrease in Sales and/or Profits</td>
<td>53</td>
<td>22%</td>
</tr>
<tr>
<td>Decrease in Number of Clients</td>
<td>9</td>
<td>4%</td>
</tr>
<tr>
<td>Working Less Hours</td>
<td>6</td>
<td>2%</td>
</tr>
<tr>
<td>Business Stable</td>
<td>27</td>
<td>11%</td>
</tr>
</tbody>
</table>
The data analysis shows that, in many cases, the increase of sales and profits was associated with the diversification to new activities, complementing or replacing the main business at the time of the first loan. The cash crisis has refrained consumption and made it more difficult for the entrepreneurs to get paid, especially for those who were not able to install SWPE machines. The coexistence of two or more activities appears to be one of the most adopted strategies to cope with these challenges, often associated with creative strategies such as trading clothes and blankets for maize in the rural areas, and then sell the maize in the urban areas. For those clients with growing businesses (or combination of businesses), a common expression used is “she now restocks more (frequently, volume, diversity of products)”. 

**Household outcomes**

At the household level, data was collected on personal income from business, total personal income (including income from other sources such as pensions or remittances), income from other household members and household total income, as well as food, education, health and total household expenses. These are indicators prone to error, especially in informal contexts, so particular attention was put on formulating the questions to minimise errors and missing values.

Looking at the main indicators in Table 3, it is easy to realise the negative effects of the cash crisis, aggravating an already difficult economic environment. Clients' income and consumption average levels have decreased in the period 2016-2017.

<table>
<thead>
<tr>
<th>Table 3 – Main Income and Expense Indicators</th>
<th>2016</th>
<th>2017</th>
<th>Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Business Income</td>
<td>$343.94</td>
<td>$298.47</td>
<td>-13%</td>
</tr>
<tr>
<td>Personal Total Income</td>
<td>$424.72</td>
<td>$373.16</td>
<td>-12%</td>
</tr>
<tr>
<td>Other HH Members Income</td>
<td>$318.91</td>
<td>$261.22</td>
<td>-18%</td>
</tr>
<tr>
<td>Household Total Income</td>
<td>$799.10</td>
<td>$635.50</td>
<td>-20%</td>
</tr>
<tr>
<td>Total Consumption Poverty Line (5 persons)</td>
<td>$501.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Poverty Line (5 persons)</td>
<td>$164.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Food Expenses</td>
<td>$110.89</td>
<td>$109.89</td>
<td>-1%</td>
</tr>
<tr>
<td>Education Expenses</td>
<td>$144.62</td>
<td>$60.09</td>
<td>-58%</td>
</tr>
<tr>
<td>Health Expenses</td>
<td>$19.94</td>
<td>$27.11</td>
<td>36%</td>
</tr>
<tr>
<td>Other HH Expenses</td>
<td>$88.23</td>
<td>$124.18</td>
<td>49%</td>
</tr>
<tr>
<td>Household Total Expenses</td>
<td>$356.27</td>
<td>$321.22</td>
<td>-10%</td>
</tr>
</tbody>
</table>

(1) Poverty Lines in June 2017 for the Harare province published by ZIMSTAT (http://www.zimstat.co.zw/prices-statistics-zimbabwe)

The effects of the crisis seem to be especially strong on education expenses (-58%), but it is important to highlight that this decrease does not translate into less households sending their children to school. In fact, out of the 245 clients, 204 report expenses with education, a higher number than in 2016 (197), which implies that the reduction is linked to the average amount spent by family. Contrarily, health and other expenses have increased on average.

The apprehension felt by the clients due to the economic environment reflected as well in the slight decrease of the share of clients willing to make house improvements (23% compared to 33% at baseline) or acquire household fixed assets (51% compared to 55% in 2016). It was probably also a factor contributing to the

---

1 The values are all considered at current values since inflation was not significant during the study period. The annual inflation rate in June 2017 was 0.3% (http://www.zimstat.co.zw/prices-statistics-zimbabwe). Speculation about a new inflationary period started after the second interviews were completed.
noteworthy change in savings. There was a strong reduction of the clients reporting never saving, from 57 to 22 clients (approximately half of them corresponding to clients who have closed their businesses). Even more relevant is the increase in the percentage of clients declaring to save regularly (88% compared to 73% before applying for the loan). For those saving regularly or occasionally, ROSCAAs are the main savings channel (81% in 2017), followed by saving at home (19.5%).

At the household level, data was collected to calculate a poverty scorecard. Poverty scorecards externally developed, such as the PPI® (Poverty Probability Index) or PAT® (Poverty Assessment Tool), were not available for Zimbabwe at the time of the interviews. Given this reality, THRIVE managers have decided to develop their own poverty assessment tool, reproducing the methodology adopted by Mark Schreiner’s team when developing the PPI. THRIVE-PAT lacks, however, the extensive testing required in refining a complex statistical process such as the construction of a poverty scorecard, and for this reason the external validity of the tool must be considered as limited.

This fact does not undermine its usefulness at internal level to compare clients and assess changes in their situation over time. In addition, THRIVE-PAT has an advantage compared with PPIs developed for other countries, as it uses local benchmarks (Zimstat data for the Harare province), and not national data, when defining the scores for each question. In this way, the accuracy of the analysis is enhanced with respect to the geographical context, which is significant in a country with strong regional inequalities.

THRIVE-PAT includes 8 multiple-choice questions about household characteristics and assets ownership. The answers to these questions have associated scores, and their sum gives the total score for the household. It ranges from 1 to 49 and, the higher the score, the lower is the probability of the household being considered poor. In the context of this study, THRIVE-PAT was used mainly to track household changes over time, complementing other indicators. No conclusions were drawn exclusively from the scores regarding the absolute poverty level of the respondents.

The average THRIVE-PAT for the sample clients increased from 23.5 at baseline to 24.2 in 2017, with the small increase in the score indicating that the probability of the clients being considered poor became slightly lower. The aggregate picture conceals, however, a diversity of situations that are best identified when looking at the variations of the individual scores. For half of the sample (50%), THRIVE-PAT has increased between 2016 and 2017, implying an improvement in their economic circumstances, but there is also 8% of the sample for whom there was no change, and 42% who have experienced a deterioration of the score. An important follow-up on this report will be the analysis of which factors (questions in THRIVE-PAT) contributed the most for the variations in both directions, as well as the amplitude of variation.

The existence of differences among clients regarding THRIVE-PAT scores, and respective variation, was further explored by using statistical tests, namely Mann-Whitney tests, to compare client sub-groups (according to some defined characteristic). The main results of these non-parametric tests include:

---

Further information on PPI® and PAT® available at [https://www.povertyindex.org/](https://www.povertyindex.org/) and [https://www.povertytools.org/](https://www.povertytools.org/), respectively.

The initial step in the analysis was the calculation of the main descriptive statistics as well as the computation of the Jarque-Bera test of normality for the variables of interest (list of variables in annex I). Given that for almost all tested variables, the hypothesis of normal distribution was rejected at a 5% significance level, the decision was to perform nonparametric tests, namely Mann-Whitney tests, less demanding in the assumptions regarding the population studied (statistically significant results are listed in annex II).
There were no significant differences between clients of the two branches for the average absolute scores, but there were significant differences when looking to their variation during the period. THRIVE-PAT score increased, on average, 15% for Willowvale clients and 6% for Chitungwiza clients.

Clients with education levels above O-Level achieved higher THRIVE-PAT scores than the other clients, in both waves of the survey. In the second survey, the score for those with ‘higher education’ was 26.6, on average, compared with 24.0 for the other clients. There were, however, no differences in the variation of the scores.

In both waves of the survey, the differences of THRIVE-PAT scores for those clients who are house owners, those who are renting and those who have made house improvements during the period between questionnaires were statistically significant.

<table>
<thead>
<tr>
<th></th>
<th>House Owners</th>
<th>Other Clients</th>
<th>Renting</th>
<th>Other Clients</th>
<th>Home improvements</th>
<th>Other Clients</th>
</tr>
</thead>
<tbody>
<tr>
<td>2016</td>
<td>27.6</td>
<td>22.3</td>
<td>18.2</td>
<td>25.5</td>
<td>25.6</td>
<td>22.8</td>
</tr>
<tr>
<td>2017</td>
<td>28.4</td>
<td>22.8</td>
<td>22.7</td>
<td>24.8</td>
<td>26.2</td>
<td>23.4</td>
</tr>
</tbody>
</table>

Parallel to the question on business changes, clients were also asked how the household had changed since the first interview. Table 5 presents the changes explicitly mentioned by the clients showing how diversified were the experiences and the focus of the clients.

<table>
<thead>
<tr>
<th>Table 5 – Client Perception on Household Changes</th>
<th>No Respondents</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased Consumption/Meeting All Needs</td>
<td>51</td>
<td>21%</td>
</tr>
<tr>
<td>Improvements in Food Consumption</td>
<td>59</td>
<td>24%</td>
</tr>
<tr>
<td>Improvements in Education</td>
<td>28</td>
<td>12%</td>
</tr>
<tr>
<td>Improvements in Housing Conditions</td>
<td>28</td>
<td>11%</td>
</tr>
<tr>
<td>Improvements in Clothing and HH Assets</td>
<td>48</td>
<td>19%</td>
</tr>
<tr>
<td>Satisfaction and Happiness</td>
<td>23</td>
<td>9%</td>
</tr>
<tr>
<td>Ability to Contribute to HH income and Help Family</td>
<td>9</td>
<td>4%</td>
</tr>
<tr>
<td>Better Self-Esteem</td>
<td>15</td>
<td>6%</td>
</tr>
<tr>
<td>Decrease in Food Consumption</td>
<td>19</td>
<td>8%</td>
</tr>
<tr>
<td>Decrease in Other Consumption</td>
<td>25</td>
<td>10%</td>
</tr>
<tr>
<td>Increased stress/less autonomy</td>
<td>11</td>
<td>4%</td>
</tr>
<tr>
<td>No changes</td>
<td>47</td>
<td>19%</td>
</tr>
</tbody>
</table>

As could be expected, food, education and housing, either improving or deteriorating, were the aspects most mentioned by the clients. In any of these areas, clients referred to changes in quantity (being able or not to fulfill these basic needs), and/or quality of the services. For instance, improvements in education include not only being able to pay school fees and keep the children at school (avoiding being “chased away”), but also moving the children to better schools. These, however, were not the only changes valued by the clients. It is interesting to note how clothing is mentioned by several clients as an important change, often associated with self-esteem and status.
Part III – Comparison with non-clients

One of the objectives of the project was to evaluate the ‘impact’ of the loans funded by LWC lenders. For this purpose, it was considered critically important to set up a comparison group of non-clients with characteristics similar to the client group. This was not an easy task from a logistical perspective. To achieve it, commercial areas in the neighbourhoods where clients lived and worked were used as basis to recruit participants. In 2016, 157 non-client female entrepreneurs were interviewed. The focus on commercial areas led, however, to the sample of non-clients to be different from the THRIVE clients in the sample in several variables, including those related to income and poverty levels. Nonetheless, the comparison results still provide relevant information to better understand the changes experienced by THRIVE clients.

To compare the two sub-samples of clients and non-clients interviewed at both waves of the survey, statistical tests (Mann-Whitney) were performed to identify statistically significant differences between the two groups regarding the variables included in the study (see Annexes I and II).

The results show that, on average, non-clients in the sample tended to be younger (36 years) compared with the clients (42 years). They lived in smaller households (4.3 members compared with 4.8 for the clients), and they more frequently lived in rented houses (65%) which sharply contrasts with the housing situation of most clients, with 70% living in family or their own houses and only 27% renting.

Other substantial difference between clients and non-clients refers to the average number of working hours per week. Non-clients seem to have a higher workload (66h compared with 60.2h for the clients). This finding, at least partially, derives from the prevalence of trade activities within the non-client sample (66% compared with 62% for the clients), and the fact that non-clients are more likely to work alone in their businesses (74% compared with 59% for the clients).

The differences between the two sub-groups are also significant regarding income, expenses and THRIVE-PAT scores. Table 6 summarizes the average results for the income, expenses and poverty scores.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Client</th>
<th>Average Non-client</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Business Income 2016</td>
<td>$349.94</td>
<td>$197.07</td>
</tr>
<tr>
<td>Personal Business Income 2017</td>
<td>$290.47</td>
<td>$144.50</td>
</tr>
<tr>
<td>Personal Total Income 2016</td>
<td>$490.15</td>
<td>$238.00</td>
</tr>
<tr>
<td>Personal Total Income 2017</td>
<td>$373.16</td>
<td>$156.83</td>
</tr>
<tr>
<td>Household Total Income 2016</td>
<td>$799.10</td>
<td>$407.47</td>
</tr>
<tr>
<td>Household Total Income 2017</td>
<td>$635.50</td>
<td>$429.29</td>
</tr>
<tr>
<td>Household Total Expenses 2016</td>
<td>$356.27</td>
<td>$238.24</td>
</tr>
<tr>
<td>Household Total Expenses 2017</td>
<td>$321.22</td>
<td>$213.43</td>
</tr>
<tr>
<td>THRIVE-PAT 2016</td>
<td>23.5</td>
<td>14.1</td>
</tr>
<tr>
<td>THRIVE-PAT 2017</td>
<td>24.2</td>
<td>20.9</td>
</tr>
</tbody>
</table>
The figures show that non-clients in the sample were, on average, poorer than clients in 2016, and they have continued to be in 2017. Both groups were affected by the economic context, with most of the indicators decreasing in the period.\textsuperscript{11} Interestingly, the reduction in the average personal business income was stronger for non-clients (-24\%) compared with -12\% for the clients), which suggests that the access to the loan(s) possibly acted as a ‘cushion’ for the effects of the economic crisis. Further analysis is required to confirm this finding. The contraction in consumption was similar for the two groups (-10\%).

Despite the economic context, THRIVE-PAT scores increased, on average, for both groups, which indicates a potential improvement in the economic condition of the households (i.e. they are less likely to be considered poor). However, non-clients seemed to have performed better in this indicator, with an increase of 78\% in the average score compared with 11\% for the clients. This statistically significant difference implies a reduction on the poverty gap between the two groups, but to confirm and better understand this result, further analysis of the scoreboard components and other eventual explanatory factors is necessary.

Departing from the statistical tests, and looking at other relevant findings of the surveys, there are two indicators that should be emphasized. There is a clear difference on savings practices, with the percentage of clients saving regularly (80\%) being much higher than in the first survey (59\%) and much higher than the same indicator for non-clients (59\%). Inversely, those declaring to have never saved during the previous 12 months are more significant among non-clients (27\%) compared with the clients (9\%).

Equally interesting is the comparison of answers to the question on how the respondents qualified the overall change in their lives (compared with the time of the first interview). Most clients considered their life to be better than before (64\%), with 24\% stating the situation was similar, and just 13\% of those interviewed declaring being in a worse moment of their lives. Non-clients, in turn, had a gloomier perspective about the evolution of their lives during the period, with 36\% considering to be in a worse position, 54\% reporting no changes and only 29\% declaring a positive change.

The analysis so far provides some interesting findings, but these are based on simple averages, thus not necessarily representative of all clients and non-clients. If, as expected, there is heterogeneity in the samples, these results will not tell the whole story. To start exploring this idea, the total sample was divided in 4 equal subgroups (quartiles) based on the variation (change) of the personal business income during the period. This variation was calculated dividing the personal business income reported in 2017 by the personal business income from 2016. The resultant new indicator (PI Bus 2017/PI Bus 2016) is larger than 1 for positive variations of the personal business income in the period, and between 0 and 1 for negative variations. Table 7 shows the quartiles 1 to 3 (the 4\textsuperscript{th} quartile corresponds to the maximum value of the indicator).

<table>
<thead>
<tr>
<th>Quartile</th>
<th>PI Bus 2017/PI Bus 2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (0.25)</td>
<td>0.44</td>
</tr>
<tr>
<td>2 (0.5)</td>
<td>0.80</td>
</tr>
<tr>
<td>3 (0.75)</td>
<td>1.50</td>
</tr>
</tbody>
</table>

\textsuperscript{11} Unlike the comparison of indicators for the two groups at each of the surveys, the variation of income and expenses during the period for the two groups were not statistically significant, which seems to indicate that the adverse economic conditions influenced equally clients and non-clients.
Survey respondents (clients and non-clients) in the lower quartile are those for whom the personal business income in the 2017 survey represented 44% or less than the personal business income reported in the 2016 survey. Inversely, those in the upper quartile have reported in the 2017 survey personal business incomes higher 150% or more relatively to the personal business income in 2016.

To understand which factors are statistically significant to explain the variations in the personal business income, a quartile regression model was implemented, using the variable ‘PI Bus 2017/PI Bus 2016’ as the explained variable. The potential explanatory variables included age, marital status, location (branch), educational level, house tenure, business time, working hours/week, working alone, THRIVE-PAT score and Personal Business Income in 2016 (baseline), external shocks and participation in the programme (access to THRIVE business loans).

The regression results can be consulted in annex III. They show that, at a level of confidence of 95%, there are different factors associated with the variation on the personal business income depending on the quartile. Here are highlighted the main results:

- There is a negative relationship between the initial personal business income (‘PI Bus 2016’ at baseline) and the variation of the personal business income during the period, common to all quartiles. Therefore, a respondent with a higher initial personal business income was more likely to have a lower variation of income compared with another respondent with a lower initial income.
- Being more experienced in the business (‘Business Time’) and having a lower poverty score at baseline (‘PAT_2016’) seems to have had a positive impact on the variation of the personal business income for those respondents which experience negative variations of personal business income (quartiles 1 and 2).
- Programme participation was found to be statistically significant for respondents in quartile 1. The access to a THRIVE loan seems to have functioned as a cushion for those clients more severely affected by the economic crisis, (those with personal business incomes lower than 44% of the income reported in the previous year), preventing the decrease of their income being greater compared with the non-clients. Being client or non-client did not have a significant impact on the variation of business income for the respondents in the other quartiles.

The quartile process was replicated using the variation of THRIVE-PAT score (PAT_2017/PAT_2016) as explained variable, and the same explanatory variables mentioned above.

<table>
<thead>
<tr>
<th>Quartile</th>
<th>PAT_2017/PAT_2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (0.25)</td>
<td>0.92</td>
</tr>
<tr>
<td>2 (0.5)</td>
<td>1.12</td>
</tr>
<tr>
<td>3 (0.75)</td>
<td>1.47</td>
</tr>
</tbody>
</table>

In this case, only quartile 1 corresponds exclusively to respondents who experienced negative variations of their poverty scores. The first conclusion from the regression analysis (see Annex IV) is that programme participation (access to THRIVE loan) is not statistically significant to explain the variation of the poverty scores during the period, for any of the sample quartiles.

---

12 Quantile regression is an econometric model which allows to overcome issues regarding the variable behaviour (including not normal distributions) and get further insight on the sample heterogeneity. For methodological reasons, all the quantitative variables were transformed into their logarithmic version for the analysis.
There are, however, five other variables which were found to be statistically significant at a confidence level of 95% for all quartiles - ‘Age’, ‘Own house’, ‘Working hours per week’, ‘Saving Regularly’ and ‘PAT_2016’. For the initial four indicators, the regression coefficients have a positive sign, signifying that being older, owning a house, working more hours per week and saving regularly likely had a positive impact on the variation of THRIVE-PAT scores. Conversely, the relation between initial THRIVE-PAT score and variation during the period is negative, which indicates that lower scores at baseline (the poorer respondents) were more likely to have higher variations of the scores, potentially reducing the initial poverty gap.

Conclusions and Recommendations

This report gives an insight into the changes in the lives of THRIVE clients after a period slightly over one year. Although THRIVE loan cycle is short, which meant that some of the clients in the sample were already repaying their third loan, several authors suggest that one year is insufficient to capture changes in poverty levels.11

Adding to the short period of analysis is the fact that the period was in many respects extraordinary in Zimbabwe. There were several negative consequences of the economic and cash crisis (reduced incomes, contraction of consumption, increase of external shocks and health problems), which presented important challenges for both clients and institution. Given that few financial institutions in Zimbabwe are willing and capable of working with the more vulnerable populations, demand will continue to exist, but risks are considerably higher in these uncertain conditions.

Nonetheless, clients, in general, have maintained a positive attitude, despite all the difficulties, and have looked for strategies to cope with the challenges raised by the cash crisis. Most clients, with active businesses by the time of the second interview, reported higher or similar sales in 2017 and 64% of all clients considered their life to be better than before applying for the business loan; interestingly, only 29% of the interviewed non-clients replied similarly to this question.

Other positive indications, identified in the survey, were the willingness of most clients to apply for further loans; the valorisation by the clients of the training provided as well as the loan conditions and customer service; and, finally, the increase of savings frequency, implying the further development of a savings culture among the clients.

The two surveys implemented provided valuable data that, in many cases, not providing definite explanations for the changes in the lives of THRIVE clients and how they compared with non-clients, allow for the presentation of a number of recommendations regarding the programme and the continuation of the evaluation project.

1. The extraordinary circumstances of the period make it more relevant to repeat the survey and identify (and attempt to explain) changes in a period expected to be more stable. The suggestion is that the repetition occurs after a period between one year to one year and six months, preferably. The definition of the timings should accommodate the need to ensure that local interviewers are available to start the implementation of the survey during the training period.

2. Although training was globally appreciated by the clients, several of them stated in the first survey that the duration of the initial training was excessive. Considering that the 6 plus 2 training sessions received by the sample clients contributed to a longer loan approval process and represented a relevant cost for the institution, the recommendation is to shorten the number of sessions, focusing on the crucial elements of the training. The main messages to the clients may then be reinforced by the loan officers in their visits and by phone messages, an instrument which has proven to be effective in different contexts.14

---

13 See Chen et al. (2010)
14 By the time of the second survey there were already changes in THRIVE training programme for the new clients (namely on its duration).
3. Linked to the shortening of the initial training is the idea of preparing additional training sessions on specific topics designed to meet the needs of particular segments among the clients. The report gives some detail on the differences between sub-groups of clients, and between clients and non-clients, which can be further explored for this purpose. Among the new subjects to be consider is digital literacy given that the use of mobile money is one of the permanent changes emerging from the cash crisis.

4. The differences between clients and non-clients in the baseline survey for a significant number of income and poverty related indicators suggest that THRIVE is not fully reaching poorer segments of female entrepreneurs, running mainly trade businesses and most often living in rented houses. The recommendation is to better understand why this exclusion occur – to what extent is this an auto-exclusion by the entrepreneurs who do not feel confident to take the risk associated with the loan or is it linked to the characteristics of the credit product offered by THRIVE or yet if the exclusion derives from a tendency of loan officers to pursue not-so-poor clients perceived as safer in terms of loan repayment. If the institution intends to work as well with these group of clients, it may be necessary to design new products, which not necessarily have to be business loans.

One final note should be on the importance of THRIVE participation in the project. The institution is working in the implementation of a continuous social performance management process, in which LWC assessment project is a valuable component. The evaluation project can potentially benefit all stakeholders involved – providing evidence to LWC on the performance of its partner as well as relevant management information to THRIVE managers on their clients; and ultimately, benefiting the clients with an improved programme that better suits their needs. Moreover, the lessons learnt throughout the process have the potential to be useful not only in the context of THRIVE but other LWC partners.

References


### Annex I

List of quantitative variables tested using Mann–Whitney Tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 2017</td>
<td>Age in 2017 (survey 2017)</td>
<td>No</td>
</tr>
<tr>
<td>Household Size 17</td>
<td>Household members in 2017 (survey 2017)</td>
<td>No</td>
</tr>
<tr>
<td>Workers 2017</td>
<td>Number of workers in the household (survey 2017)</td>
<td>No</td>
</tr>
<tr>
<td>Business Time 17</td>
<td>Years owning the business in 2017 (survey 2017)</td>
<td>No</td>
</tr>
<tr>
<td>Working Hours/Week 17</td>
<td>Average hours/week in 2017 from original data on 'working hours/day' and 'working days/week' (survey 2017)</td>
<td>No</td>
</tr>
<tr>
<td>First Loan Amount</td>
<td>First loan amount (survey 2017)</td>
<td>USD</td>
</tr>
<tr>
<td>Second Loan Amount</td>
<td>Second loan amount (survey 2017)</td>
<td>USD</td>
</tr>
<tr>
<td>Third Loan Amount</td>
<td>Third loan amount (survey 2017)</td>
<td>USD</td>
</tr>
<tr>
<td>PAT_2016</td>
<td>THRIVE-PAT score at baseline (survey 2016)</td>
<td>No</td>
</tr>
<tr>
<td>PAT_2017</td>
<td>THRIVE-PAT score in 2017 (survey 2017)</td>
<td>No</td>
</tr>
<tr>
<td>PI Business 16</td>
<td>Monthly personal income from business in 2016 (survey 2016)</td>
<td>USD</td>
</tr>
<tr>
<td>PI Business 17</td>
<td>Monthly personal income from business in 2017 (survey 2017)</td>
<td>USD</td>
</tr>
<tr>
<td>PI Total 16</td>
<td>Monthly personal income from other sources in 2016 (survey 2016)</td>
<td>USD</td>
</tr>
<tr>
<td>PI Total 17</td>
<td>Monthly personal income from other sources in 2017 (survey 2017)</td>
<td>USD</td>
</tr>
<tr>
<td>HHI Other Members 16</td>
<td>Monthly income of other household members in 2016 (survey 2016)</td>
<td>USD</td>
</tr>
<tr>
<td>HHI Other Members 17</td>
<td>Monthly income of other household members in 2017 (survey 2017)</td>
<td>USD</td>
</tr>
<tr>
<td>HHI Total 2016</td>
<td>Monthly household total income in 2016 (survey 2016)</td>
<td>USD</td>
</tr>
<tr>
<td>HHI Total 2017</td>
<td>Monthly household total income in 2017 (survey 2017)</td>
<td>USD</td>
</tr>
<tr>
<td>HH Total Expenses 16</td>
<td>Monthly household total expenses in 2016 (survey 2016)</td>
<td>USD</td>
</tr>
<tr>
<td>HH Total Expenses 17</td>
<td>Monthly household total expenses in 2017 (survey 2017)</td>
<td>USD</td>
</tr>
<tr>
<td>Weight Food/ HHE 17</td>
<td>Proportion of food expenditure in the household total expenses in 2017 (survey 2017)</td>
<td>No</td>
</tr>
<tr>
<td>Weight Education/ HHE 17</td>
<td>Proportion of education expenditure in the household total expenses in 2017 (survey 2017)</td>
<td>No</td>
</tr>
<tr>
<td>THRIVE-PAT Relative Variation</td>
<td>Percent difference between THRIVE-PAT scores in 2016 and 2017 (survey 2016 and survey 2017)</td>
<td>%</td>
</tr>
<tr>
<td>Total Employment Absolute Variation</td>
<td>Absolute difference between total employment in 2016 and 2017 (survey 2016 and survey 2017)</td>
<td>No</td>
</tr>
<tr>
<td>PI Business Relative Variation</td>
<td>Percent difference between personal income from business in 2016 and 2017 (survey 2016 and survey 2017)</td>
<td>%</td>
</tr>
<tr>
<td>PI Total Relative Variation</td>
<td>Percent difference between total personal income in 2016 and 2017 (survey 2016 and survey 2017)</td>
<td>%</td>
</tr>
<tr>
<td>Household Income Relative Variation</td>
<td>Percent difference between household income in 2016 and 2017 (survey 2016 and survey 2017)</td>
<td>%</td>
</tr>
<tr>
<td>Household Expenses Real Variation</td>
<td>Percent difference between household expenses in 2016 and 2017 (survey 2016 and survey 2017)</td>
<td>%</td>
</tr>
</tbody>
</table>

**Note:** All monetary indicators are presented in USD (United States Dollars) at current values. Inflation during the study period was not significant.
Annex II

Non-parametric Tests

The analysis was carried out for all the variables included in Annex I. Mann-Whitney tests were used to test the null hypothesis of the distribution of the selected variables being equal for defined subgroups. The subgroups are associated with clients’ characteristics and they were chosen considering the results from the baseline survey and qualitative information collected throughout the project. In order to apply the tests, dummy variables were constructed for:

- Location (Willowvale; Chitungwiza)
- Marital Status (married; widow)
- Educational level (primary education – grades 1 to 7; secondary education – O’ Level, form 1 to 4; higher education – A’ Level and university)
- Type of activity (trade; crossborder trade; production; services; agriculture and animal creation);
- Savings Frequency (regularly saving; never saving)
- House ownership (home owner – with or without property title; renting house)

The table below shows the results statistically significant at a significance level of 5% (***) or 1% (****), demonstrating the existence of differences between the subgroups for the studied variable. As guidance in reading the table, here is an example:

Clients developing trade activities were found to be different from clients with other type of businesses in terms of the first loan amount with “trade” clients accessing lower amounts, on average ($314.87 compared with $550.97).

<table>
<thead>
<tr>
<th>Location</th>
<th>Average Willowvale</th>
<th>Average Chitungwiza</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI Business Relative Variation</td>
<td>0.00</td>
<td>0.56</td>
<td>5,449 ***</td>
</tr>
<tr>
<td>PI Total Relative Variation</td>
<td>0.06</td>
<td>0.53</td>
<td>5,883 ***</td>
</tr>
<tr>
<td>THRIVE-PAT Relative Variation</td>
<td>0.15</td>
<td>0.06</td>
<td>6,092.5 **</td>
</tr>
<tr>
<td>First Loan Amount</td>
<td>$319.92</td>
<td>$351.35</td>
<td>5,880 ***</td>
</tr>
<tr>
<td>Working Hours/Week 2017</td>
<td>56.4h</td>
<td>64.2h</td>
<td>3,428.5 **</td>
</tr>
<tr>
<td>HHI Other Members 17</td>
<td>$358.74</td>
<td>$245.73</td>
<td>5,259.5 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Marital Status</th>
<th>Average Married</th>
<th>Average Other Status</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 2017</td>
<td>39.9</td>
<td>46.7</td>
<td>3,941 ***</td>
</tr>
<tr>
<td>Household Size 2017</td>
<td>5.1</td>
<td>5.9</td>
<td>3,661 ***</td>
</tr>
<tr>
<td>Workers 2017</td>
<td>2.1</td>
<td>1.6</td>
<td>3,782 ***</td>
</tr>
<tr>
<td>PI Total 2017</td>
<td>$368.64</td>
<td>$597.96</td>
<td>4,878 ***</td>
</tr>
<tr>
<td>HHI Other Members 2016</td>
<td>$396.83</td>
<td>$116.09</td>
<td>2,686.5 ***</td>
</tr>
<tr>
<td>HHI Other Members 2017</td>
<td>$342.86</td>
<td>$51.10</td>
<td>1,664.5 ***</td>
</tr>
<tr>
<td>HHI Total 2016</td>
<td>$873.62</td>
<td>$605.15</td>
<td>4,461.5 ***</td>
</tr>
<tr>
<td>HHI Total 2017</td>
<td>$707.94</td>
<td>$449.06</td>
<td>3,941 ***</td>
</tr>
<tr>
<td>HHI Total Expenses 2017</td>
<td>$225.94</td>
<td>$308.92</td>
<td>4,938.5 **</td>
</tr>
</tbody>
</table>
### LWC – THRIVE REPORT 2017

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Widow</th>
<th>Average Other Status</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>THRIVE-PAT Relative Variation</td>
<td>-0.06</td>
<td>0.14</td>
<td>2,478.5 ***</td>
</tr>
<tr>
<td>Age 2017</td>
<td>27.5</td>
<td>22.8</td>
<td>2,333.5 ***</td>
</tr>
<tr>
<td>Household size 2017</td>
<td>4.1</td>
<td>4.9</td>
<td>2,830 ***</td>
</tr>
<tr>
<td>Workers 2017</td>
<td>1.6</td>
<td>2.0</td>
<td>2,678 ***</td>
</tr>
<tr>
<td>Business Time 2017</td>
<td>13.8y</td>
<td>7.5y</td>
<td>2,249 **</td>
</tr>
<tr>
<td>PI Total 2016</td>
<td>$542.56</td>
<td>$404.43</td>
<td>2,933.5 **</td>
</tr>
<tr>
<td>PI Total 2017</td>
<td>$463.53</td>
<td>$357.60</td>
<td>2,644 ***</td>
</tr>
<tr>
<td>HHI Other Members 2016</td>
<td>$128.56</td>
<td>$351.70</td>
<td>1,963.5 ***</td>
</tr>
<tr>
<td>HHI Other Members 2017</td>
<td>$43.33</td>
<td>$299.11</td>
<td>1,478.5 ***</td>
</tr>
</tbody>
</table>

#### Educational Level

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Secondary Education</th>
<th>Average Other Education</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>THRIVE-PAT Relative Variation</td>
<td>0.13</td>
<td>0.01</td>
<td>3,128 **</td>
</tr>
<tr>
<td>Age 2017</td>
<td>40.3y</td>
<td>49.8y</td>
<td>2,389 ***</td>
</tr>
<tr>
<td>First Loan Amount</td>
<td>$328.40</td>
<td>$355.38</td>
<td>3,185 ***</td>
</tr>
<tr>
<td>HH Total Expenses 16</td>
<td>$343.33</td>
<td>$423.97</td>
<td>3,129.5 **</td>
</tr>
<tr>
<td>PAT_2016</td>
<td>23.0</td>
<td>25.7</td>
<td>3,045 **</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Primary Education</th>
<th>Average Other Education</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age 2017</td>
<td>53.6y</td>
<td>40.7y</td>
<td>1,067.5 ***</td>
</tr>
<tr>
<td>Business Time 17</td>
<td>14.1y</td>
<td>7.8y</td>
<td>1,502 **</td>
</tr>
<tr>
<td>HHI Other Members 17</td>
<td>$85.34</td>
<td>$377.86</td>
<td>1,508.5 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Higher Education</th>
<th>Average Other Education</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Loan Amount</td>
<td>$388.24</td>
<td>$390.27</td>
<td>1,253 **</td>
</tr>
<tr>
<td>PI Total 2016</td>
<td>$581.59</td>
<td>$413.03</td>
<td>1,220 **</td>
</tr>
<tr>
<td>PI Total 2017</td>
<td>$687.65</td>
<td>$490.71</td>
<td>1,330.5 ***</td>
</tr>
<tr>
<td>HHI Total 2016</td>
<td>$1,130.71</td>
<td>$774.38</td>
<td>1,150.5 ***</td>
</tr>
<tr>
<td>HHI Total 2017</td>
<td>$1,163.53</td>
<td>$595.78</td>
<td>1,130.5 ***</td>
</tr>
<tr>
<td>HH Total Expenses 2016</td>
<td>$475.94</td>
<td>$347.27</td>
<td>987.5 ***</td>
</tr>
<tr>
<td>HH Total Expenses 2017</td>
<td>$518.12</td>
<td>$306.54</td>
<td>1,080 ***</td>
</tr>
<tr>
<td>PAT_2016</td>
<td>27.8</td>
<td>23.1</td>
<td>1,160 ***</td>
</tr>
<tr>
<td>PAT_2017</td>
<td>26.6</td>
<td>24.0</td>
<td>1,276.5 **</td>
</tr>
</tbody>
</table>

#### Economic Activity

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Crossborder</th>
<th>Average Other Activity</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment Absolute Variation</td>
<td>0.3</td>
<td>-0.1</td>
<td>3,201.5 ***</td>
</tr>
<tr>
<td>Second Loan Amount [1]</td>
<td>$468.75</td>
<td>$426.61</td>
<td>3,346.5 **</td>
</tr>
<tr>
<td>PI Business 2016</td>
<td>$440.77</td>
<td>$323.83</td>
<td>3,275.5 **</td>
</tr>
<tr>
<td>PI Business 2017</td>
<td>$425.86</td>
<td>$271.35</td>
<td>3,154.5 ***</td>
</tr>
<tr>
<td>PI Total 2016</td>
<td>$506.67</td>
<td>$407.32</td>
<td>3,468 **</td>
</tr>
<tr>
<td>PI Total 2017</td>
<td>$492.60</td>
<td>$347.74</td>
<td>3,500 **</td>
</tr>
</tbody>
</table>

[1] Average for clients who have applied for a second loan

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Trade</th>
<th>Average Other Activity</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>PI Business Relative Variation</td>
<td>0.50</td>
<td>0.06</td>
<td>6,086 **</td>
</tr>
<tr>
<td>PI Total Relative Variation</td>
<td>0.47</td>
<td>0.11</td>
<td>6,259 **</td>
</tr>
</tbody>
</table>
# LWC – THRIVE REPORT 2017

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Agriculture/Animal</th>
<th>Average Other Activity</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>THRIVE-PA Relative Variation</td>
<td>0.17</td>
<td>0.05</td>
<td>6,078.5 **</td>
</tr>
<tr>
<td>Working Hours/Week 2017</td>
<td>64h</td>
<td>57.4h</td>
<td>3,186.5 ***</td>
</tr>
<tr>
<td>First Loan Amount</td>
<td>$314.97</td>
<td>$350.57</td>
<td>5,648.5 ***</td>
</tr>
<tr>
<td>PI Business 2016</td>
<td>$305.00</td>
<td>$376.20</td>
<td>5,888.5 ***</td>
</tr>
<tr>
<td>PI Total 2016</td>
<td>$352.10</td>
<td>$484.88</td>
<td>5,733.5 ***</td>
</tr>
<tr>
<td>HHI Total 2016</td>
<td>$857.70</td>
<td>$907.95</td>
<td>6,286 **</td>
</tr>
<tr>
<td>Weight Food/Total Expenses 2017</td>
<td>0.36</td>
<td>0.41</td>
<td>6,168 **</td>
</tr>
<tr>
<td>PAT 2016</td>
<td>21.9</td>
<td>24.7</td>
<td>5,517 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Agriculture/Animal</th>
<th>Average Other Activity</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>THRIVE-PAT Relative Variation</td>
<td>-0.09</td>
<td>0.13</td>
<td>1,682 ***</td>
</tr>
<tr>
<td>Age 2017</td>
<td>49y</td>
<td>41y</td>
<td>2,020.5 **</td>
</tr>
<tr>
<td>Working Hours/Week 2017</td>
<td>49.9h</td>
<td>61.5h</td>
<td>1,096 ***</td>
</tr>
<tr>
<td>First Loan Amount</td>
<td>$396.15</td>
<td>$326.95</td>
<td>1,908.5 ***</td>
</tr>
<tr>
<td>Second Loan Amount</td>
<td>$505.26</td>
<td>$425.55</td>
<td>2,151.5 **</td>
</tr>
<tr>
<td>PI Total 2016</td>
<td>$619.31</td>
<td>$401.56</td>
<td>1,894 **</td>
</tr>
<tr>
<td>PI Total 2017</td>
<td>$492.46</td>
<td>$302.56</td>
<td>2,068.5 **</td>
</tr>
<tr>
<td>HHI Total 2016</td>
<td>$1,268.19</td>
<td>$743.41</td>
<td>2,176.5 **</td>
</tr>
<tr>
<td>HHI Total 2017</td>
<td>$838.77</td>
<td>$611.14</td>
<td>2,117.5 **</td>
</tr>
<tr>
<td>HHI Total Expenses 2016</td>
<td>$481.35</td>
<td>$341.29</td>
<td>1,882 ***</td>
</tr>
<tr>
<td>HHI Total Expenses 2017</td>
<td>$388.87</td>
<td>$313.19</td>
<td>2,153.5 **</td>
</tr>
<tr>
<td>PAT 2016</td>
<td>39.1</td>
<td>22.8</td>
<td>1,365.5 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Services</th>
<th>Average Other Activity</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Time 2017</td>
<td>11.3y</td>
<td>7.6y</td>
<td>2,883 ***</td>
</tr>
<tr>
<td>HHI Other Members 2017</td>
<td>$817.26</td>
<td>$596.42</td>
<td>3,427 **</td>
</tr>
</tbody>
</table>

## Savings Frequency

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Regularly Saving</th>
<th>Average Not Regularly Saving</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment Absolute Variation</td>
<td>0.1</td>
<td>-0.4</td>
<td>3,608.5 ***</td>
</tr>
<tr>
<td>THRIVE-PAT Relative Variation</td>
<td>0.13</td>
<td>-0.01</td>
<td>3,554 ***</td>
</tr>
<tr>
<td>Workers 2017</td>
<td>2</td>
<td>1.7</td>
<td>3,564 ***</td>
</tr>
<tr>
<td>Third Loan Amount</td>
<td>$543.38</td>
<td>$466.67</td>
<td>3,305.5 ***</td>
</tr>
<tr>
<td>PI Business 2017</td>
<td>$325.14</td>
<td>$188.88</td>
<td>2,907.5 ***</td>
</tr>
<tr>
<td>PI Total 2017</td>
<td>$401.25</td>
<td>$257.90</td>
<td>3,146.5 ***</td>
</tr>
<tr>
<td>HHI Total 2017</td>
<td>$655.73</td>
<td>$553.51</td>
<td>3,470.5 ***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Never Saving</th>
<th>Average Saving</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employment Absolute Variation</td>
<td>-0.8</td>
<td>0.0</td>
<td>1,273.5 ***</td>
</tr>
<tr>
<td>HHI Total Relative Variation</td>
<td>-0.02</td>
<td>0.32</td>
<td>1,793 **</td>
</tr>
<tr>
<td>THRIVE-PAT Relative Variation</td>
<td>-0.10</td>
<td>0.13</td>
<td>1,378.5 ***</td>
</tr>
<tr>
<td>Workers 2017</td>
<td>1.5</td>
<td>2</td>
<td>1,448 ***</td>
</tr>
<tr>
<td>Working Hours/Week 2017</td>
<td>76.9h</td>
<td>59.4h</td>
<td>453.5 **</td>
</tr>
<tr>
<td>Third Loan Amount</td>
<td>$600.00</td>
<td>$589.29</td>
<td>1,806.5 **</td>
</tr>
<tr>
<td>PI Business 2016</td>
<td>$341.68</td>
<td>$354.03</td>
<td>1,753.5 **</td>
</tr>
<tr>
<td>PI Business 2017</td>
<td>$121.56</td>
<td>$315.64</td>
<td>1,112.5 ***</td>
</tr>
<tr>
<td>PI Total 2017</td>
<td>$176.82</td>
<td>$392.53</td>
<td>1,166.5 ***</td>
</tr>
<tr>
<td>HHI Other Members 2017</td>
<td>$117.73</td>
<td>$275.50</td>
<td>1,777 **</td>
</tr>
</tbody>
</table>
### LWC – THRIVE REPORT 2017

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average House Owner</th>
<th>Average Non-house owner</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHI Total 2017</td>
<td>$294.55</td>
<td>$369.44</td>
<td>869 ***</td>
</tr>
<tr>
<td>HHI Total Expenses 2017</td>
<td>$370.61</td>
<td>$316.34</td>
<td>1,727 5 **</td>
</tr>
<tr>
<td>PAT 2017</td>
<td>21.8</td>
<td>24.5</td>
<td>1,654 **</td>
</tr>
</tbody>
</table>

#### House Ownership

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Rented House</th>
<th>Average Non-rented house</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>THRIVE-PAT Relative Variation</td>
<td>0.35</td>
<td>0.02</td>
<td>3,046 5 ***</td>
</tr>
<tr>
<td>Age 2017</td>
<td>37y</td>
<td>44y</td>
<td>3,056 ***</td>
</tr>
<tr>
<td>Household Size 2017</td>
<td>4.4</td>
<td>5.0</td>
<td>4,852 **</td>
</tr>
<tr>
<td>First Loan Amount</td>
<td>$808.72</td>
<td>$644.77</td>
<td>4,406 5 ***</td>
</tr>
<tr>
<td>PI Total 2017</td>
<td>$287.12</td>
<td>$405.58</td>
<td>4,838 **</td>
</tr>
<tr>
<td>HHI Total 2017</td>
<td>$483.52</td>
<td>$693.26</td>
<td>4,715 **</td>
</tr>
<tr>
<td>Weight Food/Total Expenses17</td>
<td>0.42</td>
<td>0.38</td>
<td>3,791 5 ***</td>
</tr>
<tr>
<td>Weight Education/Total Expenses17</td>
<td>0.18</td>
<td>0.16</td>
<td>4,534 5 ***</td>
</tr>
<tr>
<td>PAT 2016</td>
<td>18.2</td>
<td>25.5</td>
<td>2,129 ***</td>
</tr>
<tr>
<td>PAT 2017</td>
<td>22.7</td>
<td>24.8</td>
<td>4,406 **</td>
</tr>
</tbody>
</table>

#### Other Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Clients</th>
<th>Average Non-Clients</th>
<th>U statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>41.8y</td>
<td>36.1y</td>
<td>5.363152 ***</td>
</tr>
<tr>
<td>Household Size</td>
<td>4.8</td>
<td>4.3</td>
<td>2.448088 ***</td>
</tr>
<tr>
<td>Working Hours/Week</td>
<td>60.2h</td>
<td>68.8h</td>
<td>3.204275 ***</td>
</tr>
<tr>
<td>Personal Business Income 2016</td>
<td>$343.54</td>
<td>$197.07</td>
<td>5.739680 ***</td>
</tr>
<tr>
<td>Personal Business Income 2017</td>
<td>$298.47</td>
<td>$144.90</td>
<td>6.058736 ***</td>
</tr>
<tr>
<td>Personal Total Income 2016</td>
<td>$460.19</td>
<td>$213.00</td>
<td>6.842681 ***</td>
</tr>
<tr>
<td>Personal Total Income 2017</td>
<td>$573.16</td>
<td>$156.83</td>
<td>8.451286 ***</td>
</tr>
<tr>
<td>Household Other HH Members</td>
<td>$318.01</td>
<td>$187.29</td>
<td>3.798804 ***</td>
</tr>
<tr>
<td>Income 2016</td>
<td>$799.10</td>
<td>$407.47</td>
<td>6.985092 ***</td>
</tr>
<tr>
<td>Household Total Income 2016</td>
<td>$625.50</td>
<td>$429.29</td>
<td>6.281868 ***</td>
</tr>
<tr>
<td>Household Total Income 2017</td>
<td>$365.27</td>
<td>$238.24</td>
<td>3.519882 ***</td>
</tr>
<tr>
<td>Household Total Expenses 2016</td>
<td>$321.22</td>
<td>$213.43</td>
<td>5.266460 ***</td>
</tr>
<tr>
<td>Thrive-PAT 2016</td>
<td>23.5</td>
<td>14.1</td>
<td>10.290615 ***</td>
</tr>
<tr>
<td>Thrive-PAT 2017</td>
<td>24.2</td>
<td>20.9</td>
<td>5.945824 ***</td>
</tr>
<tr>
<td>Thrive-PAT Relative Variation</td>
<td>0.11</td>
<td>0.78</td>
<td>8.574786 ***</td>
</tr>
</tbody>
</table>
### Annex III – Quantile Regression 1

Dependent Variable: PI Business 2017 / PI Business 2016 (Ln)
Method: Quantile Regression
Sample: 955
Included observations: 272

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Type</th>
<th>Quartile 1 (0.25)</th>
<th>Quartile 2 (0.50)</th>
<th>Quartile 3 (0.75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age_LN</td>
<td>Q</td>
<td>0.1138</td>
<td>0.3858 **</td>
<td>0.7216 ***</td>
</tr>
<tr>
<td>Loan (Client)</td>
<td>D</td>
<td>0.5686 ***</td>
<td>0.3276 *</td>
<td>0.1124</td>
</tr>
<tr>
<td>Willowvale</td>
<td>D</td>
<td>0.0183</td>
<td>0.0135</td>
<td>0.0049</td>
</tr>
<tr>
<td>Married</td>
<td>D</td>
<td>-0.0712</td>
<td>0.0164</td>
<td>0.0062</td>
</tr>
<tr>
<td>Widow</td>
<td>D</td>
<td>-0.1892</td>
<td>-0.1905</td>
<td>-0.18567</td>
</tr>
<tr>
<td>Primary School</td>
<td>D</td>
<td>0.0360</td>
<td>-0.0245</td>
<td>-0.4442</td>
</tr>
<tr>
<td>Secondary School</td>
<td>D</td>
<td>0.0611</td>
<td>0.0902</td>
<td>-0.2455</td>
</tr>
<tr>
<td>Rented House</td>
<td>D</td>
<td>0.0594</td>
<td>0.0255</td>
<td>0.1186</td>
</tr>
<tr>
<td>Own House</td>
<td>D</td>
<td>0.0387</td>
<td>0.0652</td>
<td>0.2079</td>
</tr>
<tr>
<td>Business Time_LN</td>
<td>Q</td>
<td>0.1134 ***</td>
<td>0.1546 ***</td>
<td>0.0507</td>
</tr>
<tr>
<td>Working Hours/Week_LN</td>
<td>Q</td>
<td>0.1555</td>
<td>0.2073 *</td>
<td>0.1985 *</td>
</tr>
<tr>
<td>Work Alone</td>
<td>D</td>
<td>-0.0464</td>
<td>-0.1666</td>
<td>-0.1700</td>
</tr>
<tr>
<td>Shocks</td>
<td>D</td>
<td>-0.0485</td>
<td>-0.0865</td>
<td>-0.2424 **</td>
</tr>
<tr>
<td>Saving Regularly</td>
<td>D</td>
<td>0.2096</td>
<td>0.1998</td>
<td>0.2179</td>
</tr>
<tr>
<td>PAT_2010_LN</td>
<td>Q</td>
<td>0.4382 ***</td>
<td>0.3363 **</td>
<td>0.3193 *</td>
</tr>
<tr>
<td>PI Business 2016_LN</td>
<td>Q</td>
<td>-0.6917 ***</td>
<td>-0.7299 ***</td>
<td>-0.7556 ***</td>
</tr>
</tbody>
</table>

Notes: Q – quantitive variable, D – dummy variable (1,0)
*** significance level of 1%, ** significance level of 5%, * significance level of 10%
Annex IV – Quantile Regression 2

Dependent Variable: PAT_2017/ PAT_2016 (ln)
Method: Quantile Regression
Sample: 555
Included observations: 273

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Type</th>
<th>Quantile 1 (0.25)</th>
<th>Quantile 2 (0.50)</th>
<th>Quantile 3 (0.75)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age_LN</td>
<td>Q</td>
<td>0.2012 ***</td>
<td>0.3043 ***</td>
<td>0.4326 ***</td>
</tr>
<tr>
<td>Loan (Client)</td>
<td>D</td>
<td>-0.0014</td>
<td>0.0122</td>
<td>0.0411</td>
</tr>
<tr>
<td>Widow</td>
<td>D</td>
<td>0.0812</td>
<td>0.0820 *</td>
<td>0.0344</td>
</tr>
<tr>
<td>Married</td>
<td>D</td>
<td>0.0912</td>
<td>0.0700</td>
<td>0.0660</td>
</tr>
<tr>
<td>Widow</td>
<td>D</td>
<td>-0.0043</td>
<td>-0.0281</td>
<td>-0.05104</td>
</tr>
<tr>
<td>Primary School</td>
<td>D</td>
<td>0.0485</td>
<td>0.1630 **</td>
<td>0.0649</td>
</tr>
<tr>
<td>Secondary School</td>
<td>D</td>
<td>0.1539**</td>
<td>0.2371 ***</td>
<td>0.1382</td>
</tr>
<tr>
<td>Rented House</td>
<td>D</td>
<td>0.1411 **</td>
<td>0.1283 ***</td>
<td>0.0297</td>
</tr>
<tr>
<td>Own House</td>
<td>D</td>
<td>0.1688 ***</td>
<td>0.2083 ***</td>
<td>0.2058 ***</td>
</tr>
<tr>
<td>Business Time_LN</td>
<td>Q</td>
<td>-0.0191</td>
<td>-0.2191</td>
<td>-0.0405 *</td>
</tr>
<tr>
<td>Working</td>
<td>Q</td>
<td>0.1457 ***</td>
<td>0.1162 ***</td>
<td>0.1306 ***</td>
</tr>
<tr>
<td>Working Hours/Week_LN</td>
<td>D</td>
<td>0.0539</td>
<td>0.0275</td>
<td>-0.0023</td>
</tr>
<tr>
<td>Work Alone</td>
<td>D</td>
<td>0.0166</td>
<td>-0.0277</td>
<td>-0.0470</td>
</tr>
<tr>
<td>Saving Regularly</td>
<td>D</td>
<td>0.1209 **</td>
<td>0.0809</td>
<td>0.1202 ***</td>
</tr>
<tr>
<td>PAT_2016_LN</td>
<td>Q</td>
<td>-0.6578 ***</td>
<td>-0.6561 ***</td>
<td>-0.7302 ***</td>
</tr>
<tr>
<td>Pi Business 2016_LN</td>
<td>Q</td>
<td>0.0367</td>
<td>0.0167</td>
<td>0.0271</td>
</tr>
</tbody>
</table>

Notes: Q – quantitative variable, D – dummy variable [1,0]
*** significance level of 1%, ** significance level of 5%, * significance level of 10%
26 March 2018

Joana Silva Afonso  
PhD Student  
Portsmouth Business School

Dear Joana

<table>
<thead>
<tr>
<th>Study Title:</th>
<th>Evaluating outcomes in microfinance: challenges and lessons learnt from Pakistan and Zimbabwe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethics Committee reference:</td>
<td>E380 [Substantial Amendment]</td>
</tr>
</tbody>
</table>

Thank you for submitting your documents for ethical review. The Ethics Committee was content to grant a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, revised in the light of any conditions set, subject to the general conditions set out in the attached document, and with the following stipulation:

The favourable opinion of the EC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including University of Portsmouth, prior to the start of the study.

**Summary of any ethical considerations:**

A favourable ethical opinion to the amendment has been granted on condition that you lodge with FEC the amended questionnaire that Lendwithcare will be using.
Documents reviewed

The documents reviewed by Peter Scott, taking chair’s action in approving further amendments.

<table>
<thead>
<tr>
<th>Document</th>
<th>Version</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notice of substantial Amendment</td>
<td>V3</td>
<td>19/03/18</td>
</tr>
<tr>
<td>Application form</td>
<td>V3</td>
<td>19/03/18</td>
</tr>
</tbody>
</table>

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements set out by the University of Portsmouth.

After ethical review

Reporting and other requirements

The attached document acts as a reminder that research should be conducted with integrity and gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Notification of serious breaches of the protocol
- Progress reports
- Notifying the end of the study

Feedback

You are invited to give your view of the service that you have received from the Faculty Ethics Committee. If you wish to make your views known please contact the administrator, Christopher Martin.

Please quote this number on all correspondence: E380

Yours sincerely and wishing you every success in your research

Chair

Enclosures: “After ethical review – guidance for researchers”
26 March 2018

Joana da Silva Afonso  
PhD Student  
Faculty of Business and Law

Dear Joana

<table>
<thead>
<tr>
<th>Study Title:</th>
<th>Evaluating outcomes in microfinance: challenges and lessons learnt from Pakistan and Zimbabwe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethics Committee reference:</td>
<td>E334 [Substantial Amendment]</td>
</tr>
</tbody>
</table>

Thank you for submitting your substantial amendment document for ethical review. The Ethics Committee was content to grant a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, revised in the light of any conditions set, subject to the general conditions set out in the attached document, and with the following stipulation:

The favourable opinion of the EC does not grant permission or approval to undertake the research. Management permission or approval must be obtained from any host organisation, including University of Portsmouth, prior to the start of the study.

Summary of any ethical considerations:

- Documents reviewed

The documents reviewed by Peter Scott, taking chair’s action in approving further amendments
Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements set out by the University of Portsmouth.

After ethical review

Reporting and other requirements

The attached document acts as a reminder that research should be conducted with integrity and gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Notification of serious breaches of the protocol
- Progress reports
- Notifying the end of the study

Feedback

You are invited to give your view of the service that you have received from the Faculty Ethics Committee. If you wish to make your views known please contact the administrator, Sharman Rogers.

Please quote this number on all correspondence: E334

Yours sincerely and wishing you every success in your research

Chair

Email:

Enclosures: “After ethical review – guidance for researchers”

Copy to: Joe Cox
FORM UPR16
Research Ethics Review Checklist

Please include this completed form as an appendix to your thesis (see the Research Degrees Operational Handbook for more information).

Postgraduate Research Student (PGRS) Information

<table>
<thead>
<tr>
<th>Student ID:</th>
<th>768067</th>
</tr>
</thead>
</table>

PGRS Name: Joana Da Silva Afonso

Department: Economics/Finance

Start Date: February 2015

First Supervisor: Joe Cox

Study Mode and Route:

- Full-time
- PhD
- Professional Doctorate

Title of Thesis: Evaluating Outcomes in Microfinance: Challenges and Lessons from Pakistan and Zimbabwe

Thesis Word Count: 72,397

If you are unsure about any of the following, please contact the local representative on your Faculty Ethics Committee for advice. Please note that it is your responsibility to follow the University’s Ethics Policy and any relevant University, academic or professional guidelines in the conduct of your study.

Although the Ethics Committee may have given your study a favourable opinion, the final responsibility for the ethical conduct of this work lies with the researcher(s).

UKRI/C Finished Research Checklist:
(If you would like to know more about the checklist, please see your Faculty or Departmental Ethics Committee reps or see the online version of the full checklist at: http://www.ukri.org/what-we-do/ethics-practice-for-research)

- a) Have all of your research and findings been reported accurately, honestly and within a reasonable time frame? YES ☒ NO ☐
- b) Have all contributions to knowledge been acknowledged? YES ☒ NO ☐
- c) Have you complied with all agreements relating to intellectual property, publication and authorship? YES ☒ NO ☐
- d) Has your research data been retained in a secure and accessible form and will it remain so for the required duration? YES ☒ NO ☐
- e) Does your research comply with all legal, ethical, and contractual requirements? YES ☒ NO ☐

Candidate Statement:

I have considered the ethical dimensions of the above named research project, and have successfully obtained the necessary ethical approval(s).

Ethical review number(s) from Faculty Ethics Committee (or from NRES/SCREC):

E334 / E386

If you have not submitted your work for ethical review, and/or you have answered 'No' to one or more of questions a) to e), please explain below why this is so.

Signed (PGRS): [Signature]

Date: 30.07.2018

UPR16 – April 2018