



LUND UNIVERSITY  
School of Economics and Management

## Master in Economic Development and Growth

# MICROFINANCE IMPACT EVALUATION: THEORY VERSUS PRAXIS

Filippo Capurro  
[fi7550ca-s@student.lu.se](mailto:fi7550ca-s@student.lu.se)

*Abstract:* By conducting a comparative analysis of microfinance impact evaluation (IE) from the researchers and practitioners perspective, the paper aims to uncover the existing gap between theory and reality. To that purpose an academic literature review on IE, based on Randomized Controlled Trials (RCTs), is contrasted with the practitioners' social performance management approach (SPM). The two approaches differ not only in the results, but also methodologically, thus confirming the existence of such a gap. In particular the SPM framework appears to be more flexible, pragmatic and epistemologically grounded, while RCTs seem based on a more rigorous and statistical analysis, but both emerge as replicable and falsifiable. Besides the cooperation between the two environments is likely to bridge, or at least reduce, the gap because their advantages and disadvantages are more complementary than initially expected. On one hand practitioners can add rigorousness and quantitative analysis to the SPM approach and on the other hand academia can benefit from a higher degree of realism, pragmatism and philosophical solidity. Finally their merger, however with SPM as a guiding frame, appears also capable of strengthening both the scientific and pragmatic validity of microfinance IEs.

*Key words:* Impact Evaluation, Epistemology, Microfinance, Randomized Controlled Trials, Social Performance Management

EKHM52  
Master thesis, Second Year (15 credits ECTS)  
June 2017  
Supervisor: prof. dr. Benny Carlson  
Examiner: prof. dr. Cornelis Jaco Zuijderduijn  
Word Count: 18356



LUND UNIVERSITY

School of Economics and Management

## Contents

MICROFINANCE IMPACT EVALUATION: THEORY VERSUS PRAXIS.....	1
List of Abbreviations.....	4
List of Figures.....	5
List of Tables.....	5
I. Introduction .....	6
II. Literature Review.....	8
2.1. An historical, epistemological and methodological overview of the economic discipline.....	8
2.2 A microfinance overview.....	13
2.3 Randomized Controlled Trials .....	19
III. Comparative Analysis .....	25
3.1 Microfinance impact evaluation in academia .....	25
3.2 Microfinance impact evaluation by practitioners .....	39
IV. Discussion .....	49
V. Conclusions .....	56
Reference.....	60
Bibliography.....	60
Interview.....	65
Courses .....	65
Website list .....	65
Appendix.....	67
PART A – Cross referencing.....	67
PART B – SPM Universal Standards .....	70

## List of Abbreviations

AS: Adverse Selection

ASCA: Accumulating Savings and Credit Association

CGAP: Consortium Group to Assist the Poorest

CSA: Child Saving Account

E: Earmarking

FIA: Formal Impact Assessment

FP: Financial Provider

GL: Group Loan

GP: Grace Period

IIA: Informal Impact Assessment

IE: Impact Evaluation

IFRS: International Financial Reporting Standards

IL: Individual Loan

IPV: Intimate Partner Violence

MFI: MicroFinance Institution

MH: Moral Hazard

ODA: Official Development Assistance

RCT: Randomized Controlled Trials

ROSCA: ROTating Saving and Credit Association

SA: Saving Account

SHG: Self-Help Group

SMART: Specific, Measurable, Achievable, Relevant and Time-bound

SO: SOcial tie

SPM: Social Performance Management

SPMUS: Social Performance Management Universal Standards

SPTF: Social Performance Task Force

ST: STorage

## **List of Figures**

Figure 3. 1: Universal Standards of SMP .....	41
Figure 3. 2: SPM implementation process .....	43

## **List of Tables**

Table 3. 1: Academic findings .....	31
Table 3. 2: Requirements by client type .....	44
Table 3. 3: Outputs and Outcomes.....	44
Table 3. 4: D1 in practice.....	45
Table 3. 5: SMART characteristics.....	45
Table 3. 6: D3 in practice.....	47
Table A. 1: Cross reference overview.....	67
Table B. 1: Dimension 2.....	70
Table B. 2: Dimension 4.....	70
Table B. 3: Dimension 5.....	71
Table B. 4: Dimension 6.....	72

## I. Introduction

After having studied Economics for five years and having lived 10 years in South America, two in Brazil and eight in Bolivia, participated to multiple volunteering initiatives in Chile, Argentina, Peru, Romania and Tanzania, and worked in an NGO as project developer assistant for Ecuador, Guatemala, Morocco and Tunisia, I noticed that there is a leap within these personal experiences, they simply do not match. Such incompatibility and even contradiction between theory and reality appears to me an interesting paradox because, actually, theory or science aims, in its self-enhancement process, to shed light, understand, describe and eventually forecast real phenomena. Consequently theory and reality should be very interconnected and closer to each other and not two distant and almost diverse domains (Capra 1992). As a matter of fact, the etymology of economics comes from the ancient Greek words οἶκος and νόμος which literally stands for ‘house’ and ‘norm’ respectively. However the broader meaning has a strong practical facet, it refers to the household administration and the efficient management of its financial resources in order to cope with family needs (Rocci 1998). Today’s academic approach to economy cannot be further away from that interpretation.

One of the main sources of such detachment is the advent, during the 20th century, of rigorous hypothetical-deductive methods, combined with quantitative and reductionist approaches, aiming to elevate Economics, a social science, to a natural science. Indeed the abuse of mathematics and statistics involved a reductionism that generated logically robust models and theories which, however, became more and more unrealistic (Lawson 2001; Weintraub 2002). Besides, over the years, a great deal of effort has been dedicated to improve the internal validity of such models, theories and quantitative tools, while relatively little attention has been devoted to test their external and epistemological validity (Deaton 2010; Lawson 2001).

By conducting a comparative analysis of microfinance impact evaluation (IE) from the researchers and practitioners perspective, this paper aims to uncover the existing gap between theory and reality and eventually identify possible solutions or guidelines to bridge it in the future. More specifically, a review of academic literature on microfinance

IE will be contrasted with the practitioners' social performance management (SPM) approach<sup>1</sup>.

Microfinance is an interesting field for two reasons: firstly because, in the last decades, it has emerged as a dynamic alternative for “third-world-countries” economic and social development and consequently it has become very popular (Pizzo et al. 2013). Secondly because microfinance is a complex phenomenon involving, directly or indirectly, many different aspects such as politics, economics, history, culture, geography and institutions that in turn are connected, within other things, to health, innovation, education, environmental sustainability, gender empowerment and poverty (Pizzo et al. 2013). In this respect microfinance is everything but a closed system, the intricate interactions of the above elements make it difficult to isolate causality (e.g. simultaneity) and include all facets (e.g. omitted variables). Since the current mainstream approach in academia for microfinance IE is the so called Randomized Controlled Trial (RCT); its internal, external and epistemological validity is tested against the validity of practitioners IE approach, the SPM (Armendáriz et al. 2010; Hulme 1997; SPTF 2016).

Although both the literature reviews on microfinance economic and social IE and the academic papers on RCT are abundant, as is the literature covering the epistemology and methodology of economics, to the best of my knowledge no study has combined these three aspects. In other words no study has addressed the epistemological and methodological issues of the economic discipline and the reality and theory clash they raise by comparing how researchers and practitioners pursue the same goal, in this case the IE of microfinance interventions.

The thesis is structured in the following way: the next section, the literature review, briefly introduces the historical, epistemology and methodological evolution of Economics, an overview on the microfinance sector and on the RCTs. The alternative academic and practitioners approaches to microfinance IE are presented in Section III (Comparative Analysis). In Section IV the differences and implications of the two IE alternatives are more thoroughly discussed, together with a possible solution. Finally Section V concludes the thesis.

---

<sup>1</sup> Despite the SPM is emerging as the best practice within the microfinance practitioners, it still has not been

## II. Literature Review

### *2.1. An historical, epistemological and methodological overview of the economic discipline*

Both the etymology of economics and history suggest that simple arithmetic and basic accounting methods, that is the concept of quantification, have been always related to the management of resources and exchange of goods. Similarly old is the use of statistics for demographic purposes which can be traced back to Ancient Egypt (Di Vittorio 2002). However it was only during the Mercantilist Era and the upsurge of national states that the use of mathematics and statistics became rifer: in that respect ‘Political Arithmetick’ by William Petty published in 1690 is emblematic (Scrapanti et al. 2004)<sup>2</sup>. A further quantum leap happened between the end of the 18<sup>th</sup> and the early 19<sup>th</sup> centuries when Classical Economic theory emerged with the first economic models<sup>3</sup>, quickly followed, before the onset of the 20<sup>th</sup> century, by the Utilitarianism and Marginalist revolution<sup>4</sup>, when calculus was introduced in economic analysis (Scrapanti et al. 2004). Since then the evolution of the economic theory has been accompanied by a parallel and systematic ‘mathematicization’ of the discipline with a climax during the 20<sup>th</sup> century when econometrics was developed and institutionalized by R. Frisch and I. Fischer<sup>5</sup> (Scrapanti et al. 2004). It is worth noting that along the 20<sup>th</sup> century a whole set of new quantitative models, theories and methods have been developed like input-output models, game theory, linear modeling and so forth (Scrapanti et al. 2004). Evidence of such process was brought to the fore by Stigler et al. (1995), who conducted a vast study on the main English-language economic journals and found that the use of verbal and geometrical techniques in academic papers dropped from 98%, in 1892-1893, to 5.3% in 1989-1990; and concurrently Econometrics (and/or Algebra) and Calculus (and more advanced techniques) passed from being almost non-existent in 1892-1893, to have a share, in 1989-1990, of 38.8% and 55.9% respectively. Similarly Sutter et al. (2007)

universally adopted (Sinha 2006; SPTF 2016).

<sup>2</sup> During the 18<sup>th</sup> century, in France the Physiocratic stream of thought emerged in juxtaposition to Mercantilism and its main representative was F. Quesnay (Scrapanti et al. 2004).

<sup>3</sup> The major contributions to the Classical theory were developed by A. Smith, R. Malthus, D. Ricardo and S. Mill (Scrapanti et al. 2004).

<sup>4</sup> Also known as Neoclassical theory of which W. Jevons, L. Walras, W. Pareto, A. Marshall and C. Menger were some of the main exponents (Scrapanti et al. 2004).

<sup>5</sup> Founders in 1930 of the Econometric society (Scrapanti et al. 2004).

analyzed the 2003-2004 publications of the ten leading English-language economic journals, finding that only 1.5% of the papers were strongly ‘math-free’<sup>6</sup>, of which 40% were published in the *Economic Journal* alone.

As mentioned before the ‘mathematicization’ process aimed to elevate the economic discipline, a social science, to a natural science<sup>7</sup>. Indeed it was common belief that with the introduction of quantitative and rigorous methodology in economics, the field would have been able to replicate the scientific success and credibility that natural sciences were obtaining during the same period. Such upsurge of quantitative methods in economics was coupled with a shift from induction to deduction and a push towards a reductionist standpoint. The hypothetico-deductive model that took over with its seemingly irrefutable logic, was not however completely flawless. In fact while the ‘Swan example’<sup>8</sup> unveils all the dangers hidden in the inductive reasoning (Caldwell 1991), the deductive syllogism can also lead to unrealistic conclusions because logic is, in the end, only an instrument and it does not say anything about the actual truth of its content (Blaug 1980). The ontological proofs of the existence of God, developed by many western philosophers, are in that respect illustrative<sup>9</sup>. The syllogism is logically valid but most people, even religious believers, will not agree in such demonstration. The main source of inconsistency is the major hypothesis (or premise), because it is assumed to be true and unquestionable (Abbagnano et al. 2002a & 2002b), but as stressed by Mokyr (2005, p. 209): *‘Truth is at best consensus between experts’*. However the reductionist approach was necessary to isolate and prove causality, because the elements studied and their causal relationships were easily identifiable only in closed and simple systems; systems that were created *ad hoc* by researchers with hypotheses and assumptions which in turn fostered a progressive detachment from reality (Hulme 1997; Lawson 2001; Verschuren 2001).

<sup>6</sup> According to the authors, a strongly ‘math-free’ paper has no numbered equations, tables of regression results, mathematical or regression analysis and it does not fall under the experimental economics label nor it is part of a special issue or symposium (Sutter et al. 2007).

<sup>7</sup> The natural sciences are biology, chemistry and physics.

<sup>8</sup> *In my life I have seen only white swans and I have never seen swans of different colours, thus all swans are white*. The example unveils the dangers of basing scientific knowledge and research on the inductive reasoning, because experience can lead to erroneous conclusions (Caldwell 1991).

<sup>9</sup> The Anselm of Canterbury’s prove posits: ‘*God is absolute, the biggest thing we can think of. Existence is thus a quality that cannot be excluded from the absolute, because it is impossible to think something bigger than the absolute. Therefore God exists*

Indeed already at the beginning of 20<sup>th</sup> century V. Volterra and then G. C. Evans doubted that utility, within the *homo oeconomicus* domain, could be measured and aggregated to social level as the utilitarianists posited. For both scholars the hypotheses of mathematical models should be rooted in “real phenomena”, in reality and not be the outcome of abstract thought (Weintraub 2002). Similarly to Volterra and Evans, K. Popper suggested that science should be grounded on adduction, a mental operation by which individuals form a conjecture (i.e. hypothesis) about relationships based on real world experience. Initially such conjecture is not yet science and it may never be, but through a formal process of discovery and research based on deduction and hypothesis testing, the conjecture might become a provisional scientific law and ultimately scientific knowledge (Blaug 1980, Caldwell 1991). Likewise according to North (1968) the selection of relevant research objects, relationships and hypotheses should be soundly rooted not only in empirical evidence, but also in common sense<sup>10</sup>.

There has been also other opponents of the classical and neoclassical theory, assumptions and methods. E.g. the German School back in the 19<sup>th</sup> century challenged both the belief that an economic discipline based on deduction could achieve ‘universal laws’ and the meaningfulness of the human rationality assumption. The German School instead preferred to develop theories in terms of regularities in which both history and geography played an important role (Scrapanti et al. 2004). Just before the turn of the 20<sup>th</sup> century T. Veblen, considered the forerunner of the Institutional Economics, in his essay *Why Is Economics not an Evolutionary Science?*, defined the economic theory of his time as teleological, taxonomic and tautological. He also viewed Economics as an evolutionary science that should be devoted to the study of the process of change rather than static equilibrium, a change that, according to Veblen, must be both cumulative and path-dependent (Hodgson 2004; Scrapanti et al. 2004).

Another fervent opponent was K. Marx who formulated three chief arguments against the mainstream economic theory: firstly he pointed out the lack of theory on the origin of profits, which according to Marx was in the labour exploitation. Secondly mainstream economic theory was unable to acknowledge the historical aspect of capitalism, and finally it focused exclusively on exchange instead of production. Marx also

---

<sup>10</sup> Common sense prevents the use of premises as the ones in ‘God’s existence’ demonstrations.

viewed human nature as being both malleable and autopoietic<sup>11</sup>, which distanced him from the *homo oeconomicus* assumption (Screpanti et al. 2004).

The hypothetico-deductive approach success, despite these critiques, derives from the fact it suits the hypothesis testing rationale very well. In other words the hypothetico-deductive model is consistent with Popper's concepts of falsifiability and replication. The former implies a theory or model in which it is possible to specify its core hypotheses and either disprove or confirm them through testing. The latter presumes that the conditions for the experiment can be specified and reproduced, possibly indefinitely, so that the hypothesis can be validated and verified with consistency (Blaugh 1980; Caldwell 1991). Specifically replication is facilitated if the methodology follows a rigorous and standardized pre-structured research design and a quantitative-deductive rationale (Verschuren 2001). In econometric analysis e.g. hypotheses testing assumes the form of the null versus the alternative hypotheses. The research is set up in such a way that it is only possible to reject or not reject the null hypotheses based on the data used. It is common practice to speak in terms of 'ability to reject or not the null hypothesis' and to avoid statements like 'the null can be accepted'. To be able to reject, a rejection rule is needed which depends on the alternative hypothesis and the chosen significance level. Such rule gives the probability of rejecting the null hypothesis when it is true<sup>12</sup>. When the data allows to reject the null then the alternative can be accepted whereas when the researcher fails to reject the null, the null itself cannot be considered true (Hill et al. 2008; Wooldridge 2013). The hypothesis testing design might reflect the falsifiability, replicability and rigor standards but it does not really help in the understanding of economic phenomena and reality. Using the "swan example", if it is assumed that "the swans are only white" (null hypothesis) and in the sample all swans are white, it is not possible to reject the null, but it is neither possible to validate the statement that 'all swans are white'. In that respect human knowledge about swans is not improved and the

---

<sup>11</sup> Human behaviour is shaped by history and the social context, therefore human nature should be considered endogenously in economic models and theory. Human nature is autopoietic in the sense that it is influenced by both the behaviour itself and the environment stimuli (Screpanti et al. 2004).

<sup>12</sup> Type I error is committed when the null hypothesis is rejected despite it was true, whereas the Type II error is committed when a false null hypothesis is not rejected. Note that the probability of the latter error type cannot be calculated (Wooldridge 2013).

'provisional theory' label<sup>13</sup> more than a sound justification appears to be a rhetoric loophole (Latouche 2010, McCloskey 1983).

The quantitative hypothetico-deductive method and its reductionist implications have also further downsides. First of all it does not solve the problem for which it was originally introduced in the economic discipline, that is to prove causality and produce universal laws like in the natural sciences. The main reasons for this are: Economics does not deal with (intrinsically and extrinsically) closed systems<sup>14</sup> and the mainstream quantitative models are basically correlation analyses and not causal theories (Hulme 1997; Lawson 2001; Verschuren 2001). In order to obtain closed system strong assumptions are needed (e.g. the rationality or perfectly competitive markets assumption) which leads to an extreme oversimplification of reality to the point that the models are so unrealistic as to have almost no practical use (Lawson 2001; Mokyr 2005; Verschuren 2001). Therefore the forecasting ability of such models is heavily undermined as well, as Kay (1995 cited in Lawson, 2001) put it: "*The differences between forecasts are trivial relative to the differences between all forecasts and what happens*". The main concern is however that the consequences of such an approach are more dramatic and real than the ones affecting other sciences using these research methods, as boldly highlighted by Hicks (1941 cited in Scarpanti et al., 2005):

*A man who is a mathematician and no more than a mathematician [...] harms no-one. An economist who is no more than an economist is a danger for his fellow men. Economics is not a thing on its own; it is the study of an aspect of the life of man in society [...] The economist of tomorrow (and sometimes of today) will certainly be aware of what to ground his economic advice on; but if [...] his economic knowledge is detached from any background of social philosophy, he runs a real risk of becoming a cheat, capable of implementing enterprising stratagems to find his way out of difficulties, but incapable of staying in contact with those fundamental virtues on which a sound society is founded. Modern economic science is subject to a real risk of*

---

<sup>13</sup> According to Popper all knowledge is provisional, conjectural and hypothetical. Scientific theories can only be 'provisionally' confirmed or 'condusively' refuted, but never categorically proven true. Thus scientists should refer to theories not as ultimately true, but just as highly corroborated by rigorous testing. In that respect theories may be provisionally retained as the best available until they are falsified, if it ever happens, or substituted by a better theory (Blaugh 1980, Caldwell 1991).

<sup>14</sup> Intrinsically closed means that *ceteris paribus* the objects studied will always behave in the same way, i.e. a combination of element **a** and **b** will always produce **c**. Extrinsically closed means that all the factors affecting **c** (in this case **a** and **b**) are entirely internalized in the system of analysis (Lawson 2001).

*Machiavellianism: the treatment of social problems as mere technical issues and not as an aspect of the general quest for the Good Life.*

The next two actual examples show the danger of excessively relying on the universality and ability to predict of reductionist hypothetico-deductive models and their mechanical approach (Verschuren 2001). One is the Washington Consensus, a set of policy prescriptions aiming to promote economic development in underdeveloped countries, in order to ease their convergence towards developed nations, which miserably failed (Rodrik 2006). Another is the deregulation dogma that, assuming among other things actors' rational behavior and the ability of markets to achieve efficiency alone, led to the 2008 financial crisis (Presbitero 2009). Such oversimplification proclivity seems very often more a misinterpretation of Occam's razor<sup>15</sup>, than a rigorous and scientific methodology (Blaugh 1980, Caldwell 1991).

A common researchers' reaction to these inadequacies is to either improve the model (e.g. by including more variables) or to vivisect further the phenomena because it is yet insufficiently atomistic (insufficiently simple and closed) for a proper understanding. These are obviously not solutions to the problem, as should be clear by now, since the problems are not strictly connected to the quantitative instruments of the hypothetico-deductive approach, but to the hypothetico-deductive model itself, because of its intrinsic epistemological and methodological framework (Lawson 2001).

In the following subsection, the microfinance sector and the context in which it operates will be presented and by doing so it will emerge more clearly that it is an open and complex system and consequently that the tunnel vision (reductionist) intrinsic in the hypothetico-deductive approach is unable to capture this complexity.

## 2.2 A microfinance overview

The historical origins of microfinance can be traced back to the medieval institutions called Mount of Piety in Italy during the 15<sup>th</sup> century. The Mounts of Piety

---

<sup>15</sup> It states that other things equal simpler hypotheses, not theories or explanations, are better, because they are more understandable and easily testable. It doesn't dismiss complexity, but it suggests that in case two theories well explain the same phenomena, the simpler one should be chosen (Abbagnano et al. 2002a & 2002b).

were managed by the Catholic Church and aimed to give small cash loans at low interest rates to the ones in need: it was an early form of institutionalized aid and charity<sup>16</sup>. To access the money, the person was required to deposit a valuable item (pledge), both in economic and emotional terms. The amount disbursed was proportional, but usually inferior, to the item actual economic value, while the attachment to the item guaranteed the borrower's commitment to pay back the debt. The Mounts of Piety were not however economically sustainable since they depended on the charity and donations of well-off individuals (Becchetti 2008; Pizzo et al. 2013).

The modern concept of microfinance was theorized and implemented by M. Yunus during the 70s, when some regions of Bangladesh were hit by a pervasive flooding and subsequently suffered from famine. On that occasion Yunus started lending small amounts of money to women in Jobra, a village close to the Chittagong University Campus, who managed to pay back these informal loans. The success of both such experiments and the ones that followed lead to the foundation of Grameen Bank in 1976 by the Bangladeshi economist (Armendáriz et al. 2010; Becchetti 2008; Pizzo et al. 2013). It was the first bank to address directly and exclusively the financial constraint of the poor and which strategy was based on trust and commitment rather than mainstream solvency requirements such as collateral, steady employment, verifiable credit history and others (Armendáriz et al. 2010; Bauchet et al. 2011; Pizzo et al. 2013).

The pioneer Grameen Bank experience showed that making profits and pursuing social goals (e.g. credit access to the unbanked, gender empowerment, education promotion, etc.) were not exclusive, but actually compatible objectives. Such 'win-win' conditions therefore attracted the attention not only of scientists and actors interested in development like governments and international agencies, but also of financial players interested in profit maximization. Consequently in the following decades, the 80s and 90s, microfinance established itself as a new bottom-up alternative, compared to the widespread top-down multilateral aid transfers and programs for poverty alleviation in particular in developing countries (ISPI Course 2014; Pizzo et al. 2013). Concurrently,

---

<sup>16</sup> Interest rates were formally forbidden, but it was common practice to charge them for values lower than 20% which was considered the usury rate. Interest rates, for the sole purpose of operational, were legitimated in 1513 by Pope Leo X to improve the long term viability of the Mounts (Archivio di Stato di Bologna 2016, Pizzo et al. 2013).

since microfinance was a relatively new and small market offering room and profit opportunities, financial institutions and investors, whose priorities were detached from social and economic development goals, joined the sector. The increasing flow of resources and notoriety of the microfinance industry also eased a process of ‘commercialization’ that has created some ideological turmoil, because it has been identified with a ‘mission drift’<sup>17</sup> and brought to spotlight by many experts. The win-win proposition was in part responsible for this shift because MFIs (MicroFinance Institutions) initially were subsidized through public funds, but afterwards they were progressively required to become both socially and economically viable. This pressure has modified their internal (mission, services, practices, etc.) and external (depth and breadth of outreach) structure, thus distancing MFIs from their original characteristics: small loans, serving the poorest of the poor, etc. (Armendáriz et al. 2010; Ausburg et al. 2013; Bauchet et al. 2011; Copestake 2007; Mersland et al. 2010; Murdoch 2000; Sinha 2006).

Nowadays microfinance involves a set of financial services like microloans, savings and insurance schemes, mainly addressing the financial constraints of households and small businesses that are unqualified to access the formal credit market. Such financial tools allow these vulnerable economic actors to cope with consumption needs, exogenous shocks, investment projects, etc., thus promoting their financial inclusion (Armendáriz et al. 2010; Bauchet et al. 2011; Pizzo et al. 2013; Van Rooyen et al. 2012). Each of these services materializes itself in a panoply of different financial arrangements depending on the context, the mission of the MFI, the type of borrowers, etc. For example loan contracts may be bounded to the purchase of inputs for either the farming or the business activity, which in turn may be tied to environmental and social sustainability requirements such as the purchase of organic fertilizers or local products. In other cases financial services may be devoted primarily to women for their empowerment or have

---

<sup>17</sup> There is not an unique definition of the phenomena, but it usually involves the trade-off between financial and social performance, the fact that in order to achieve financial viability the original mission, serving the poor at the bottom of the pyramid, may be given up by changing the clients target (better-off clients) and depending on the situation either expanding or reducing the pool of clients (Ausburg et al. 2013; Copestake 2007).

education requirements, that is the formal proof that beneficiaries' children are enrolled and regularly attend school (Bauchet et al. 2011; Becchetti 2008; Pizzo et al. 2013)<sup>18</sup>.

A similarly strong heterogeneity characterizes practitioners, indeed MFIs range from banks with an international network and wide outreach (e.g. the Mexican Banco Compartamos or the Bangladeshi Grameen Bank) to self-managed informal rural cooperatives with a limited pool of beneficiaries (Armendáriz et al. 2010; Becchetti 2008). Some MFIs are for-profit, others are non-profit, others aim just to break-even and reinvest the extra profits in new development initiatives; some are still subsidized by public funds others run exclusively on private equity. MFIs differ also in their mission, their internal and external structure, their legal form and the delivered services. These elements depend on the institutions' decisions, the wills of investors and donors and last but not least the context (Augsburg et al. 2013; Pizzo et al. 2013). One striking example is how MFIs and their services are shaped, in Islamic countries, by religious principles which, among other things forbid to charge interest (Pizzo et al. 2013). Another example is the financial regulations that, especially in western countries, do not allow MFIs to collect savings, a constraint that does not exist in other regions of the world (Armendáriz et al. 2010).

The main reason why some individuals are financially constrained, and consequently microfinance services needed, is the market failure, in this case the credit market failure. Information asymmetry and uncertainty are intrinsic credit market characteristics that lead everywhere to agency problems like adverse selection (AS) and moral hazard (MH) (Artoni 2005; Storey et al. 2010). AS implies that MFIs are not able to discriminate between risky and safe borrowers, because they have limited information, while MH can be either *ex ante* or *ex post*. The former refers to the situation in which the client has already received the loan, but can manipulate the success of the investment by changing unobservable aspects of his behavior (e.g. his effort, taking more risks, etc.); the latter refers to the situation in which the borrower, after the loan has been disbursed and the investment is made, hide important information or give false feedbacks, in jargon 'take the money and run away' (Armendáriz et al. 2010, Artoni 2005). In many rural and

---

<sup>18</sup> The term Microfinance plus refers to the provision of complementary services to clients alongside financial services. Such plus services usually involve business development support, health and literacy training, etc.

urban contexts in Africa, Latin America and Asia market failures are worse than the basic Pareto formulation<sup>19</sup>. Indeed, compared to developed countries, financial players in developing settings have to cope not only with market inefficiencies, but also with a weak or a complete lack of juridical system, financial and non-financial infrastructure, personal identification system, both financial and non-financial literacy and technology, collateral and economic, political and institutional stability. All these aspects exacerbate uncertainty at client and system level, which in turn aggravates MH and AS, thus also raise the MFI's operational costs. Extra costs on the one hand may imply higher interest rates, and higher interest rates on the other hand may undermine the feasibility of MFI's social commitment (Armendáriz et al. 2010; Pizzo et a. 2013).

Although often MFIs' reach and resources are limited and concurrently serve the riskiest customers<sup>20</sup>, they manage to cope with this high risk (higher costs) and to promote access to finance and to overcome the difficulties faced by traditional financial institutions by exploiting pre-existent informal solutions<sup>21</sup> (Armendáriz et al. 2010; Mauri et al. 2000; Pizzo et a. 2013). In fact consumption, insurance and investment needs have always been present in such environments and poor and isolated communities have developed informal, but effective solutions to limit e.g. MH and AS (Armendáriz et al. 2010; Collins et al. 2009). In that respect microfinance services, products and initiatives often combine these informal solutions that have an efficient legacy with tools developed by financial experts (Mauri et al. 2000; Pizzo et al. 2013).

Examples of informal solutions are the ROSCAs (ROtating and Saving Credit Associations) which involve a group of individuals who agree to regularly contribute money to a common “pot” that is allocated to one member each period. ROSCAs encompass meetings where dues are collected and proceeds allocated, moreover past recipients are excluded from getting the pot again until every member has had a turn with it. The degree of freedom in terms of spending the pot varies accordingly to what has

(Krumm 2010).

<sup>19</sup> For Pareto the main causes of market failure are public goods, information asymmetries, externalities, time inconsistent preferences and non-competitive markets (Artoni 2005; Scrpanti 2004).

<sup>20</sup> Limited reach and resources prevent MFIs to achieve economies of scale. Concurrently targeting few individuals with no collateral, landless or socially excluded increase the fixed cost per transaction.

<sup>21</sup> Collins et al. (2009) in presenting the financial diaries of the poor uncover the rife network of informal financial transactions and arrangements that characterize the life of such individuals. This credit system is based on trust and the possibility of being socially ostracized by the community in case of default (Mauri 2000).

been established in the first meeting. Compared to the ROSCAs, the ASCAs<sup>22</sup> (Accumulating Savings and Credit Associations) main additional feature is that they offer its members the possibility to save and withdraw money at any time. Other informal alternatives is borrowing from the local moneylenders, local traders, relatives, friends and neighbors (Armendáriz et al. 2010; Mauri et al. 2000).

Examples of professional solutions are dynamic incentives and frequent repayment instalments that allow MFIs to disburse progressively the loan and concurrently screen the investment development as well as borrower's behavior. The most famous solutions to reduce risk is probably group lending which in its simplest version consists in arrangements by individuals without collateral who get together and form groups to obtain loans; the loans are individual contracts, but all group members are responsible for the insolvency of another member<sup>23</sup> (Armendáriz et al. 2010; Odell 2010).

Conversely a combination of both informal and formal strategies is linkages to local markets which imply MFI's collaboration with local (informal) moneylenders who have a long-standing expertise and knowledge of the context (Floro et al. 1991). Such cooperation may require the moneylender to be involved in the distribution of loans, monitoring of borrowers behavior, collection of dues and information (Armendáriz et al. 2010). It is quite straightforward to see that peer monitoring, joint liability and moneylenders know-how can reduce risk and uncertainty, by limiting AS and MH, which subsequently allows MFIs to design looser contract terms (e.g. no collateral requirements, lower interest rates, ect.), thus promoting financial inclusion (Armendáriz et al. 2010; Pizzo et al. 2013).

In conclusion there is not 'one microfinance' and sector heterogeneity is high at all levels: institutional, service, actor and client. Contextual characteristics like culture, political and economic system matter and are often combined with a vast set of already existing informal financial transactions and arrangements that raise even more the openness and complexity of the system.

---

<sup>22</sup> Also known as Credit Cooperatives and Credit Unions.

<sup>23</sup> A contract format also known as joint liability.

### **2.3 Randomized Controlled Trials**

Up to now it has been broadly highlighted the main epistemological and methodological issues of the increasingly used quantitative hypothetico-deductive and reductionist approach in Economics. The focus now shifts to one of these methods, the RCTs, which is dominating microfinance IEs (Armendáriz et al. 2010; Bauchet et al. 2011).

IEs are important because they help to uncover the intervention causes, effects and their distribution (Armendáriz et al. 2010; Hulme 1997). Only recently RCTs have emerged as the golden standard in microfinance impact assessment, because they are believed to be successful in providing robust results (Armendáriz et al. 2010; Duvendack et al. 2011). RCTs are widely used in other fields, such as medicine<sup>24</sup>, and they have been progressively applied also in the evaluation of economic interventions. In the microfinance case the main forerunners were the members of the Poverty Action Lab at MIT, of which A. Banerjee and E. Duflo are the most prominent researchers (Bauchet et al. 2011; Deaton 2010).

Copestake (2012) includes RCTs within the Formal Impact Assessment<sup>25</sup> (FIA) methodologies which, according to the author, usually have a rigorous pre-structured research design and are based on a stimulus-response and quantitative-hypothetico deductive rationale. However RCTs entail also an experimental reductionism. Indeed the research subject is divided into units (i.e. participants in the sample) and variables, subsequently units and variables are isolated from the bigger context (i.e. experimental setting), so that the independent variables can be manipulated (i.e. supply of treatment) and their effect on the dependent variable identified (Hulme 2000; Verschuren 2001). More specifically the experimental design eases the computation of the intervention effects and eventually the unveiling of the causal mechanisms responsible for such changes (Deaton 2010; Duvendack et al. 2011). RCTs involve a baseline and follow-up survey or, more in general, at least two waves of data. During the first survey information about current economic and social conditions of participants is collected. The follow-up

---

<sup>24</sup> Both Mokyr (2005) and Deaton (2010) cast doubts on the meaningfulness of implementing methodologies coming from different fields and they define the justifications often brought up by researchers as mere rhetoric.

<sup>25</sup> Copestake (2012) also distinguishes between Informal Impact Assessment (IIA), in which the SPM approach is located, and broader Social Science Research approaches.

survey gathers the same information, ideally for the same individuals, but after the program has been implemented, so that by contrasting the *ex ante* and *ex post* situations the impact can be observed. The randomization can be of different types and it can be done at different levels<sup>26</sup>, but it implies that units are randomly assigned to either the treatment or control group. Thus the randomization can ensure that units are equivalent in terms of observable and unobservable characteristics and that the difference in performance is not contaminated by any particular individual characteristics like the unobservable entrepreneurial ability or the observable education and income level. Consequently, on average, any difference between treated and control group can be attributed entirely to the intervention or treatment (Armendáriz et al. 2010; Deaton 2010; Duvendack et al. 2011).

RCTs have become a golden standard in IE for many reasons: first of all they are rigorous, falsifiable and replicable, which, as previously seen, are necessary and desirable characteristics for scientific research methodologies (Blaugh 1980; Caldwell 1991). Secondly RCTs need less hypotheses compared to econometrical models and they appear to solve the selection bias problem. Finally because of its experimental design RCTs are thought to better perform in terms of proving causal links between the intervention and its effects (Armendáriz et al. 2010; Deaton 2010; Hulme 1997). An important caveat is that all the above elements hold conditional on the fact that it is possible to conduct an ideal RCT. Actually, the problem is that difficulties emerge and RCTs often have to be modified and adapted and by doing so the method loses its formal and scientific rigorously and it is likely to cross the blurred line between experimental and quasi-experimental (i.e. no randomization) or even the non-experimental label (i.e. impossibility to manipulate randomly the units with a treatment) (Copestake 2012; Deaton 2010).

RCTs have some drawbacks that undermine both their internal and external validity. The main internal issues are related to selection bias and program placement bias because an ideal randomization is almost impossible to conduct in reality. Indeed the

---

<sup>26</sup> Randomization can be conducted for e.g. at village level, thus all the members of one village are treated while the members of another village are not (level). Randomization at individual level can be done e.g. with the entire population, only the prospective clients or only the marginally credit worthy ones (type). The latter to avoid selection bias since there is evidence that households applying for microfinance differ significantly from the average individual: they tend to be richer, more educated, more entrepreneurial etc (Armendáriz et al. 2010; Bauchet et al. 2011).

control and treated groups may differ systematically in observable or unobservable characteristics (e.g. education level or entrepreneurial ability) and ultimately they may not be representative of the entire population, similarly a location may have fixed or variable characteristics (context favorable or unfavorable conditions) that either foster or hamper the intervention impact. Consequently the likelihood that the impact of intervention is either overestimated or underestimated is high (Armendáriz et al. 2010; Bauchet et al. 2011; Deaton 2010; Duvendack et al. 2011; Hulme 1997; Odell 2010). RCTs also suffer from the ‘omitted variable’ issue since the ability to capture the impact heavily rests on the choice of indicators, especially when the intervention encompass both direct and indirect social and economic effects. An extreme example would be the case in which the microfinance impact is measured only by means of economic indicators such as household income, investments, business profits, etc. being thus unable to compute its social impact. Furthermore the occurrence of both simultaneity and confounders challenge the ability to prove causality. In fact when it comes to the development field, the system under analysis is not a closed and simple one and isolating the causal mechanisms becomes complicated. For example, access to microfinance is believed to raise income, but income levels influence household demand of financial services, similarly successful clients may benefit from alternative and informal sources of credit so that attribution becomes controversial (Armendáriz et al. 2010; Collins et al. 2009; Deaton 2010; Hulme 1997; Mauri et al. 2000). In terms of confounders the main concerns are related to the presence of contamination and spillovers. Contamination can emerge when, in later stages of the experiment, access to financial services becomes available also in the control group area. Spillovers, instead, emerge when a neat distinction between treated and control members is not possible anymore: for instance when the treated, who received a loan, share the extra income with relatives or friends that are in the control group (Armendáriz et al. 2010). Another source of bias is attrition: if drop-outs from the experiment are not random, then they can either overstate or underestimate results. Indeed if the ones that drop-out are the successful participants, because they are now independent, the impact of the intervention will be underrated. Conversely if the ones that drop-out are the unsuccessful participants, the overall impact will be overrated (Armendáriz et al. 2010; Deaton 2010; Duvendack et al. 2011; Odell 2010). A further disadvantage is that RCTs measure the average difference in

'performance' between control and treated groups, but they do not supply any information about the median impact and the impact distribution. This is problematic from a policy perspective, because paradoxically the overall effect may be positive even if only few individuals are better-off after the intervention, if e.g. their aggregate gains are extremely high compared to aggregate marginal losses of the remaining participants. However a program producing these sort of outcomes will be generally withdrawn by policymakers and in that respect RCTs are not helpful in uncovering such situations (Armendáriz et al. 2010; Bauchet et al. 2011; Deaton 2010; Duvendack et al. 2011; Odell 2010; Van Rooyne et al. 2012).

RCTs external validity is weak as well because they are strictly connected, within other things, to the context, the financial service supplied and the MFI. This implies that a formally identical experiment, with the same financial tool, but conducted in another place or at larger scale would not lead to the same results. Similarly the RCT is likely to produce different results if, *ceteris paribus*, either the MFI or the financial service is changed<sup>27</sup>. Another point is that, because of the way RCTs are designed, they are more prone to answer if something (an intervention) works rather than explaining why it works, thus they do not really appear to improve our understanding of the phenomenon and its mechanisms, to answer the 'why question' (Armendáriz et al. 2010, Deaton 2010). Two additional and strictly related controversies, the ethical and the cost-efficiency components, add to the RCT results lack of generalizability and inadequacy to explain causality. The former stresses the problematic choice of who is going to be treated and who will not which, especially in poor and isolated contexts, raises all sorts of ethical issues<sup>28</sup>; the latter instead highlights that both a large sample and surveys, although necessary to obtain robust results, are very expensive. The problem is that RCT results without any further qualitative investigation are not conclusive and conducting additional research is also costly (Armendáriz et al. 2010; Duvendack et al. 2011; Hulme 1997; Odell 2010). The high costs affect also the duration of RCTs, generally around 2 years, and such short time frame in turn undermines the generalizability of their results (Khan et al. 2012). E.g. in terms of gender empowerment such a limited amount of time appears

---

<sup>27</sup> For example not all MFIs are equally effective, ethical and professional.

unable to gauge microfinance impact on gender, especially if it is considered that the gender revolution in western countries started around the 50s and it is still under way (Goldscheider et al. 2015; Khan et al. 2012; Kohler et al. 2002; Lesthaeghe 2010; Odell 2010).

The RCTs external validity in itself does not represent a problem if RCTs results are seen as a provisional and partial ‘theory’, however if they are generalized and rigidly used, without any consideration of geographical, cultural and political aspects, to formulate new policies and actual replications, they might become dangerous (Deaton 2010; Screpanti et al. 2005). At the same time there is not much that can be done to improve these limitations since they are rooted in the intrinsic methodological characteristics of RCTs and the microfinance sector. Conversely a lot of effort has been dedicated by researchers to improve the internal validity of RCTs. The improvements consisted in either the fine-tuning of statistical tools, the pursuit of ideal randomization, the development of robustness checks, etc. or practical solutions to selection and placement bias, simultaneity, etc. (Deaton 2010; Hulme 1997).

To overcome the narrowness of the average difference e.g., RCTs are often broadened with regressions that include a dummy variable for the treatment status and a vector of individual characteristics collected in the surveys. Thus it is possible to uncover how access to financial services interacts with these specific characteristics and the distribution of the intervention impact (Armendáriz et al. 2010; Deaton 2010). Regressions are also used to compute standard errors and significance in order to corroborate the reliability of the results. Randomization requires first of all that the probability law of control and treated group selection are the same, and secondly that such probability law is useful in determining how significant the differences in performance between the two groups are (Deaton 2010). The problem is that usually both the regressions and the probability laws require stricter assumption<sup>29</sup> that clash with the minimalist setting of RCTs and bring a new set of problems, which are however beyond

---

<sup>28</sup> Access to finance may constitute a life opportunity for some household and excluding them because of a proper randomization process becomes a pointless justification. Another example can be that social tensions may arise within the same village if some members benefit from access to finance whereas others are excluded.

<sup>29</sup> An example is that to compute the standard error with an OLS regression, the variance of treated and control group is assumed to be the same and concurrently it is assumed not to be affected by the experiment. Another

the scope of this work (Armendáriz et al. 2010; Deaton 2010). As brought to fore by McCloskey (1985), statistical significance is not the same as economic (or social) significance<sup>30</sup>. Thus statistical significant results are necessary but not sufficient for the understanding of the phenomena under analysis to the point that statistical significant results, if not accompanied by a discussion over their social and economic relevance, are meaningless (McCloskey 1985; Wooldridge 2013; Ziliak et al. 2004).

A practical solution that researches often adopt to avoid or, at least, reduce program placement and selection bias, simultaneity and contamination is to conduct the study in isolated areas where there are no other financial alternatives, where individuals have no incentives to divert resources simply because there are no competing profitable activities (Armendáriz et al. 2010; Bauchet et al. 2011; Deaton 2010; Hulme 1997; Odell 2010).

Independently if these solutions improve RCTs and are feasible, as it will emerge after the comparison between the researchers and practitioners way of conducting a microfinance IE, many of the issues that worry the academic world are, actually, often irrelevant. In fact context and clients selection is an integral part of MFIs strategy, because microfinance products are more efficient and successful both in locations where there is availability of complementary opportunities, such as markets, legal and physical infrastructure, etc. and when supplied to more resourceful, in economic and/or human capital terms, and less risky clients. Thus, overall RCTs validity appears more related to broader epistemological and methodological aspects.

---

example is the orthogonality assumption of the independent variables, which in an open system is quite unrealistic (Deaton 2010).

<sup>30</sup> Statistical significance is the probability that a certain relationship between two variables is caused by something else than random chance. Economic (and social) significance has to do more with the quantum and the sign of results and their possible implications in the real world (Wooldridge 2013).

### III. Comparative Analysis

#### ***3.1 Microfinance impact evaluation in academia***

As mentioned before, the academic microfinance IE section will be based on a review of the main publications on the topic: papers, articles and reports. Due to time, means and competences limitations, a selection has been necessary and the 15 publications overviewed were taken from: the literature of the ‘Finance and Development’ master course I attended at Groningen University<sup>31</sup> (3 papers); two systematic reviews, Van Rooyen et al. (2012) (4 papers) and Duvendack et al. (2011) (1 paper); the Bauchet et al. (2011) report (4 papers). The Banerjee et al. (2013) paper was included because it is considered, among scholars and practitioners, to be the most influential work (Armendáriz et al. 2010; Pizzo et al. 2013). The remaining 3 publications were selected because they are reviewed in both Banerjee et al. (2013) and Bauchet et al. (2011). Note that the papers selected often appear in each other’s bibliography if not in the review itself<sup>32</sup>. For a thorough cross-referencing overview see Appendix (Part A).

Initially the idea was to retrieve the 10-15 most cited papers (in descending order) on microfinance IE from two websites available at the Lund University Library: Scopus and Web of Science<sup>33</sup>. However three main problems emerged: first of all, different combinations of words returned different papers (thus a different order) both across and within online search engines<sup>34</sup>. Secondly some papers despite appearing in the list did not cover the microfinance field and some did not use RCTs. Both these issues prevented me from conducting a selection free of researcher bias (Verschuren 2001). Finally neither of the two search engines returned the Banerjee et al. (2013) at all. The impossibility to conduct an objective selection and to develop and implement a rigorous selection protocol persuaded me to opt for the ‘second best’ solution, that is to rely on the choice done by field and academic experts either rigorously, like in Van Rooyen et al. (2012) and

<sup>31</sup> The course code for the academic year 2015/2016 was EBM069A05 and further information can be found at <https://www.rug.nl/ocasys/gew/vak/show?code=EBM069A05>.

<sup>32</sup> Note that this does not necessarily increase the quality of selected papers since very often the authors, or at least one of them, are the same: in other words there is a lot of self-referencing.

<sup>33</sup> <https://www.scopus.com> and <http://login.webofknowledge.com>.

<sup>34</sup> Examples are: ‘impact assessment microfinance’, ‘impact evaluation microfinance’, ‘RCT microfinance’ and ‘Randomized Controlled Trials microfinance’. Note that the results also differed from the ones returned by the Google Scholar search engine (<https://scholar.google.com>).

Duvendack et al. (2011)<sup>35</sup>, or based on experience and knowledge like the master course coordinator B.W. Lensink, the report by Bauchet et al. (2011) and the Armendáriz et al. (2010) and Pizzo et al. (2013) handbooks.

Table 3.1 at the end of the section presents the main elements and results of the papers included in the review. Before turning to the table contents some brief guidelines on how to read it are needed. Following closely Bauchet et al. (2011), Duvendack et al. (2011) and Van Rooyen et al. (2012) for each study, the following features were highlighted: the location, the financial institution or MFI involved, the intervention rationale, the randomization and data gathering procedure. Concurrently for each paper the most striking economic and social impacts are presented. The economic subcategories are: household income, credit and savings<sup>36</sup> and business related effects, while the social ones are health, education and gender empowerment. Two further impact categories are reported: consumption and a more general one named “Other”. Consumption was separated from the economic impact because it includes also the consumption of non-durables that, despite of having a material dimension, are also related to a broader concept of well-being involving within other things security (e.g. food security) and health (e.g. nutrition levels). The label “Other” instead serves to encompass impacts that cannot be included in the other taxonomies.

Regarding the Table 3.1 symbology, in the impact part an **X** indicates that such an aspect has not been assessed by the study. All the results are reported as either statistically significant or insignificant. The latter is highlighted with the label “Statistically Insignificant”. The former type is indicated by a (+) in case of positive and different from zero effect, (-) in case of negative and different from zero effect and (Null) when the effect is non-existent or very close to zero. Note that the above (+) and (-) symbols do not refer to the sign of the statistical outcomes, but constitute a personal interpretation of reviewed papers results accordingly to the table category they are assigned. E.g. in the Field et al. (2013) paper, the beneficiaries of a grace period appear to take more risks thus the default rate of their businesses rise. In Table 3.1 the “increase in

---

<sup>35</sup> For a more thorough insight on methodology see the corresponding section in the publications. Note that the rigorous selection procedures implemented in the reviews are not manageable by a single researcher with limited availability of time, knowledge and resources.

default rates” under the “Business” category is interpreted as a negative impact, thus (-). Many studies complement the average effects with a heterogeneity analysis but, due to space limitations, only the most striking results are reported<sup>37</sup>.

Some general features of the papers can be highlighted. Firstly all studies conduct an IE using RCTs despite both the impact assessed (economic and/or social) and the randomization procedure may differ. Secondly there is a high degree of heterogeneity at all levels. Location varies in terms of country and area (urban, rural or both); the financial institutions involved range from big national banks and MFIs to very small village banks and ROSCAs<sup>38</sup>; also participants are heterogeneous: both men and women, only women, only entrepreneurs or only students. Diversity emerges likewise from the type of intervention, indeed some papers focus on improving access to savings, others to credit; and some studies combine traditional microfinance tools with complementary services such as health or business training (microfinance plus), others assess exclusively the impact of different contractual terms. Data gathering also differ: there are papers based only on a baseline and endline survey and others that, by exploiting financial institutions databases and surveys, benefit from a more continuous flow of information. A final point is that overall the IEs were limited in time, in fact of the 15 publications two had a length below the year and 4 papers a length above the 2 years, but none above the 3 years.

Turning to the results, both economic and social impact are assessed by 11 out of 15 studies; of the remaining 4, one focuses only on social impact while the last 3 only on the economic effects. The big picture that emerges from the review is that microfinance does not appear to be such a miraculous tool in terms of poverty alleviation and social development as often has been depicted, but concurrently there is no evidence that it is harmful as skeptics often posit<sup>39</sup>. Furthermore it is worth noting that statistical and economic (and social) significance are not the same thing, and although the reported

<sup>36</sup> It involves effects on access to credit and savings, the size of loans and saving accounts, frequency or likelihood of getting a loan or making a deposit, etc.

<sup>37</sup> Deaton (2010) underlines the fact that the insignificant (statistically, economically and social) results had often led researches to dig into the data to see if there is at least a positive effect for certain subcategories of microfinance beneficiaries.

<sup>38</sup> Three studies: Augsburg et al. (2015), De Mel et al. (2009) and Fafchamps et al. (2014) do not use any microfinance intermediary in the experiment implementation.

<sup>39</sup> This is particularly true for the average results.

results are the statistically significant ones, the following paragraph discusses very sketchily their economic and social significance.

In terms of economic impact the direct effect of microfinance interventions on household income appears very weak; out of the 8 papers that quantify the income effect, 4 recorded a null effect, 1 a mixed effect, 2 a minor increase and only 1 an economic significant increase. The most positive impacts are recorded in the “Savings/Credit” category. The only negative effect is found by Augsburg et al. (2015) “decline in savings” which however may be due to the extra investment in the business whereas, of the remaining 8 publications assessing either a credit or savings effect or both, 5 papers found a strong economically significant improvement and the other 3 documented at least a positive one, though not strong. The microfinance effects on business are more mixed indeed: on the one hand the impact on profits appears to be null or at best weak, but on the other hand a more consistent positive impact is recorded for revenues, investments and labor supply. Besides individuals benefiting from microfinance seem also more likely to start a business. The main concern is that these results differ for different participants: the positive effects involve only already existing businesses, certain type of businesses, or the economically stronger businesses, so it becomes difficult to reach conclusive inferences.

When it comes to social impact again microfinance impact looks weaker than expected. Many outcomes are either statistically insignificant or null, if not both. Out of 7 papers evaluating the impact on gender empowerment only Angelucci et al. (2012) and Attanasio et al. (2011) document a weak positive effect, a slightly stronger decision making power within the household and an increased likelihood of having a female-owned business respectively. Microfinance appears to generate positive outcomes on health, but mainly when the interventions address health goals directly, like in Dupas et al. (2013b) and Pronyk et al. (2008). Moreover out of 6 papers evaluating educational effects of microfinance, only 2 report statistically significant results, one is positive “improved ability to pay fees” (Banerjee et al. 2013) whereas the other is negative since 16-19 year youngsters are more likely to drop out of school than to work in the family business (Augsburg et al. 2015). In terms of consumption the overall impact seems positive, in fact all the 9 publications, evaluating household expenditure, report at least

an increase in consumption in one of the following aspects: non-durables, durables, average per capita. Nevertheless the most consistent positive result (+) is a decline in temptation goods consumption.

Before turning to the practitioners IE approach, a few considerations on the validity of these results must be made. Different statistical and methodological choices may undermine both the statistical and economic (and social) significance of results. In fact a larger sample may affect statistical significance whereas an alternative indicator or a shorter experiment length may weaken the ability to capture both social impacts (e.g. gender empowerment) and long term economic outcomes, thus the economic and social significance. Concurrently the frequency of data gathering may increase the likelihood of disturbance, due to changes in unobservables occurring in the meanwhile. In that respect results of studies conducting only a baseline survey and an endline survey 2-3 years later are more likely to suffer from such bias than the ones based on a more continuous monitoring of participants. Moreover, being the microfinance field a multidimensional and complex system, there might be a multiplicity of hidden mechanisms affecting the results. One example is that a business start-up involves a series of extra costs that may in the short run lead to negative (or null) returns, a situation that can be turned around later on (Augsburg et al. 2015; Storey et al. 2010). Similarly entrepreneurs tend, once they benefit from a loan, to pool savings and the new capital to invest in the business, which may result in lower saving rates, decline in consumption or household income, etc. to support the investment effort (Banerjee et al. 2013; Fafchamps et al. 2014). These two cases unveil that negative or weak results do not necessarily imply that the microfinance intervention has been ineffective, on the contrary they uncover the need of a more holistic contextualization and interpretation of the IE outcomes.

A final reflection regards to what extent the results obtained in these papers can be generalized. The heterogeneity above underlined suggests there is not ‘one microfinance’ and the sum of the experiments outcomes cannot ultimately define the effectiveness of the sector as a whole. Furthermore, even within the same study results often differ from participants’ categories, e.g. an intervention may benefit in different degrees men and women, farmers and businessmen, etc. Consequently conclusions on the overall usefulness of microfinance appear at best bold, if not completely wrong. Despite the

important contribution of each paper, the inferences remain very specific. Indeed at best, it is possible to conclude that given specific criterions<sup>40</sup>, within a certain time frame and under a certain methodology and hypotheses a specific microfinance tool appears more or less effective than expected. Statements or research questions such as ‘Microfinance is effective/ineffective’ or ‘Does microfinance reduce poverty?’ should be therefore taken with a pinch of salt.

---

<sup>40</sup> For example the context, the type of beneficiaries, intervention and financial provider.



**Table 3. 1:** Academic findings

PAPER	LOCATION	ACTOR	LENGTH	INTERVENTION	RANDOMIZATION
Angelucci et al. (2012)	<b>Mexico</b> – rural & urban Sonora state	Banco Compartamos - largest micro lender in the country	+2 years	Compartamos expanded the delivery of credit services to areas that it did not serve previously. The intervention involved individual loans to women with joint liability. Informal and formal credit alternatives were available before, during and after the program, but Compartamos was the first active MFI in the area. The study evaluates the impact of improved access to credit on economic and social spheres of clients' life.	Areas never served before were randomly selected by the MFI. A door-to-door promotion combined with a baseline survey was supplied to all potential borrowers. The ones that required a service became the treated group whereas the control group were the prospective clients who however did not have immediate access to financial services (it took a while for the MFI to start operations in all the identified areas). A part from the follow-up surveys the endline survey was conducted around 26 months later.
Ashraf et al. (2009)	<b>Kenya</b> – rural districts of Githugu & Kirinyaga	Pride Africa – NGO implementing the DrumNet program	1 year	A service package (involving promotion of linkages with banks, retailers, etc; marketing & business training; and credit) was delivered to small horticultural farmers. The aim was to see if they would then be capable of entering export markets and increase their income. The study is an IE of the program on mainly the economic sphere of clients.	Self-Help Groups (SHGs) were randomly assigned to three groups: treated with credit & services, treated only with credit, and control. Note that all the farmers member of a SHG were in the same group type. The study is based on a baseline and follow-up survey conducted 1 year later.
Attanasio et al. (2011)	<b>Mongolia</b> - rural villages in 5 northern provinces	XacBank - big national bank	+1 year	The study aims to unveil the differences in social and economic performance of microfinance female clients receiving either an individual loan (IL), a group loan (GL) or no credit. Note that female clients were particularly poor and generally would not have been able to apply for a XacBank loan, besides access to financial services in such areas was limited.	Information sessions were organized in 40 villages and prospective clients knew they would randomly receive an IL a GL or no loan. When around 30 women signed up in each village the baseline survey was conducted and then the randomization took place at village level. Thus all women living in a village were assigned to one of the groups. The endline survey was conducted 18 months later and a follow-up survey 2 years later.
Augsburg et al. (2015)	<b>Bosnia and Herzegovina</b> - rural and urban areas	X - large local MFI	+1 year	The experiment involved the IE of extending microcredit to marginally rejected individuals by the MFI. The clients, relatively poor individuals, were offered an IL with individual liability. The loan should be used preferably for business purposes to help clients to exploit profitable opportunities. Loan conditions were: 22% interest rates, yearly average maturity, monthly repayments and no grace period (similar to market conditions).	Marginally rejected loan applicants were identified in each MFI branch. After completing the baseline survey and an interview, they were randomized to either the control or treatment group, the treated received a loan. The follow-up survey was conducted 14 months later; to motivate the controls to respond, since they do not receive any financial benefit, they were offered little cash and a SIM card.

PAPER	ECONOMIC IMPACT			SOCIAL IMPACT				<i>Consumption</i>	<i>Other</i>
	<i>Income</i>	<i>Savings/Credit</i>	<i>Business</i>	<i>Health</i>	<i>Education</i>	<i>Gender emp.</i>			
Angelucci et al. (2012)	(Null) average effect	(+) <i>strong</i> increase in formal borrowing (Null) effect on savings	(+) effect on revenues and input investment (Null) effect on average profits*  * positive for the 90 <sup>th</sup> percentile	(+) Statistically Insignificant Expenditure on health	(+) Statistically Insignificant Expenditure on education	(+) improved female power in household decisions (Null) effect on intra household violence	(Null) effect on durables, assets and groceries (+) decline in temptation goods	(+) less stress (+) more trust in the community	
Ashraf et al. (2009)	(Null) average effect (+) <i>strong</i> increase for first-time growers	(+) more likely to obtain loans from formal sources (+) more likely to have deposits	(+) more likely to grow an export crop* (+) more land dedicated to cash crops*  * more marked for first-time growers	X	X	X	(Null) effect for inputs (+) lower expenditure for marketing*  * only for prior growers	X	
Attanasio et al. (2011)	(Null) average effect for individual loans (IL) and group loans (GL)	(+) much <i>stronger</i> likelihood to obtain a credit for both IL and GL beneficiaries	(+) higher profits after long microfinance access* (+) more likely to start a business*  *only for GC	X	X	(+) more likely to have a female-owned business for IL and GL beneficiaries	(+) higher for durables, slightly higher for non-durables*  * mainly for poorer women	X	
Augsburg et al. (2015)	(+) increase in average income from self-employment (-) decline in average wage income	(+) <i>strong</i> improvement in access to credit (-) decline in savings	(Null) effect on average profits* (+) more investments (+) more labor supply (hours) of child and adults  * once the top 1% is trimmed out profits become positive	X	(Null) effect 6-15 years old attendance (-) effect 16-19 years old attendance	X	(-) decline of food, non-durables and durables (+) decline of temptation goods	(Null) effect on alternative measures of well-being	

PAPER	LOCATION	ACTOR	LENGTH	INTERVENTION	RANDOMIZATION
Banerjee et al. (2013)	<b>India - Hyderabad</b>	Spandana – big MFI	+1 year	The MFI opened branches in areas where no formal financial alternatives were available before. Women in some areas were offered immediate access to credit whereas for others the access became available 1 year later. Access to credit was delivered as a group loan with joint liability. The study thus evaluates the impact of access to credit on social and economic spheres of female clients.	Spandana identified neighbourhoods where it never supplied before and where people were poor. After conducting a baseline survey in each neighbourhood prospective clients were paired up based on observable characteristics similarity. Finally they were randomly assigned to either the control or treatment group, so that all women within an area received the same assignment. A first endline survey was conducted after 18 months and another one after 3 years.
Crépon et al. (2011)	<b>Morocco – rural villages</b>	Al Amana - largest MFI in the country	2 years	New MFI branches opened in areas that previously had no access to microcredit and by doing so the study aims to evaluate the impact of improved access to credit on a multiplicity of social and economic aspects of clients' life.	After waves of baseline surveys, villages with similar observable characteristics were paired up. Then in treatment villages, randomly selected within each pair, Al Amana started immediately to supply microfinance services, while in the control villages access became available only 2 years later when the endline survey was conducted.
De Mel et al. (2009)	<b>Sri Lanka - Kalutara, Galle &amp; Matara districts</b>	X – grants directly supplied to beneficiaries	+2 years	Selected male and female owned microenterprises received a financial shock consisting in either a cash or equipment grant. The firms were active in the retailing, manufacturing and service sector. The aim was to evaluate the impact of different exogenous positive shocks on business performance.	After the baseline survey, microenterprises matching specific requirements were selected from each district & pooled together. Then they were randomly assigned to the in-kind or cash treatment. 11 waves of follow-up surveys were conducted the last one taking place almost more than 3 years later (subsamples).
Dupas et al. (2013a)	<b>Kenya - rural town of Bumala (Busia district)</b>	Local village bank	6 months	Since saving constraint is believed to be rife within poor entrepreneurs. By providing small business owners with access to savings accounts (SA), the experiment unveils constraints that are stifling microenterprise development. The micro entrepreneurs involved were self-employed as vendors (mainly women) and bicycle taxi drivers (men).	After stratifying for gender & occupation, participants were randomized to the treatment or control group. The treatment consisted in access to noninterest-bearing bank accounts. There were 3 waves of data flow: first the baseline survey was conducted when the SA was opened (for treated), 3 months later the treated started compiling the logbook (2° wave) for the next 3 months (3° wave).

PAPER	ECONOMIC IMPACT			SOCIAL IMPACT				<i>Other</i>
	<i>Income</i>	<i>Savings/Credit</i>	<i>Business</i>	<i>Health</i>	<i>Education</i>	<i>Gender emp.</i>	<i>Consumption</i>	
Banerjee et al. (2015)	X	(+) <u>strong</u> improvement of access to credit	(+) more creation (+) more investments (+) more labor supply (hours)	(+) Statistically Insignificant Less ill children	(+) improved ability to pay fees	(+) Statistically Insignificant Power in decision making	(+) more durables & non-durables (+) less temptation goods	(Null) Statistically Insignificant Other welfare indicators
Crépon et al. (2011)	(Null) effect on net average income	(+) <u>strong</u> improvement in access to credit (+) bigger loan size (+) higher savings only for business owners, especially livestock owners	(+) expansion of existing businesses (Null) probability to start a business  Effect on business sales, revenues & labor (Null) non-agricultural (+) agricultural	X	(Null) Statistically Insignificant Expenditure on education	(Null) Statistically Insignificant Power in decision making, but the number of women in the sample is very low	(Null) effect on average per capita (+) more food and durables*  *only for the ones without business	X
De Mel et al. (2009)	(+) slightly higher average income*  * due to business profits	X	(+) slightly higher profits, assets and investments*  * only for male clients	(Null) Statistically Insignificant*  * Sri Lanka has a relative cheap public health system	(Null) Statistically Insignificant*  * Sri Lanka has a relative cheap public education system	(Null) Statistically Insignificant Bargaining power within the household	(+) more durables	X
Dupas et al. (2013a)	(+) higher average income	(+) higher saving frequency* (+) savings growth*  *mainly for women vendors	(+) business investment*  *mainly for women vendors	(+) improved ability to deal with health shocks due to savings	X	X	(+) higher consumption per capita	X

PAPER	LOCATION	ACTOR	LENGTH	INTERVENTION	RANDOMIZATION
Dupas et al. (2013b)	<b>Kenya</b> – rural district	Local ROSCAs	3 years	The intervention aims to study the impact on preventative health investment and health shock vulnerability by loosening savings constraint with the supply of 4 technologies involving different psychological mechanisms: storage (ST), earmarking (E) & social tie (SO). The save box (ST), the lock box (ST&E), the health pot (ST, E&SO) & the health saving account (ST&E). The experiment also uncovered the key saving barriers in the context.	District ROSCAs participated to the census and the baseline survey. Selected ROSCAs also participated to the meeting, received a health training and were randomized to one of the 5 groups, 1 control & 4 treatment groups for each technology. All ROSCAs' members either benefited from one of the technologies or none. 2 follow-up survey were conducted 6 and 12 months later, whereas the endline survey took place 3 years later.
Fafchamps et al. (2014)	<b>Ghana</b> - Accra & Tema	X – grants directly supplied to beneficiaries	1 year	The intervention aims to evaluate the impact of a financial shock on business development. In that respect male and female business owners in urban and industrial areas benefited from a cash or in-kind grant.	After the baseline survey was conducted some businesses were selected. Then they were randomly assigned to one of the three groups: treatment either with cash or in-kind grants and control. Businesses were monitored at quarterly intervals for a time span for a maximum of 12 months.
Field et al. (2013)	<b>India</b> -Kolkata	Village Financial Services -small local MFI	3 years	The experiment aims to evaluate the impact of different microfinance contractual terms on entrepreneurial performance among the urban poor, specifically repayment schemes. Low income women were given individual liability loans with either immediate repayment scheme or one with a two-months grace period.	Women that were second-time borrowers of the MFI were asked to form groups of 5 individuals. Then the groups were randomly assigned to one of the two loan taxonomies. The experiment involved a survey just after women joined the program, one after the loan cycle ended and a follow-up survey for long term effects 2 years later the program has ended.

PAPER	ECONOMIC IMPACT			SOCIAL IMPACT				<i>Other</i>
	<i>Income</i>	<i>Savings/Credit</i>	<i>Business</i>	<i>Health</i>	<i>Education</i>	<i>Gender emp.</i>	<i>Consumption</i>	
Dupas et al. (2013b)	X	(+) <i>stronger</i> saving activity	X	(+) investments on preventive health goods <b>for lock &amp; safe box</b> (+) reduced vulnerability to health shocks <b>for health SA</b> (+) higher likelihood to reach health goals <b>for health pot &amp; safe box</b>	X	X	(+) more health products (+) less temptation goods & transfers	(+) spillover effects: more people using technologies* (+) take-up and saving for all technologies remained high* * after 3 years
Fafchamps et al. (2014)	X	X	Business returns (+) higher for women with already wealthy businesses & in-kind grant (+) slightly higher for men with both grant types <b>(Null)</b> effect for women with poor businesses & cash grant	X	X	X	X	X
Field et al. (2013)	(+) effect on net average income for Grace Period (GP)	X	(+) more likely to start a business* (+) higher capital & profits* (-) more likely to conduct risky practices* (-) more likely to default* * only for GP	X	X	X	X	X

PAPER	LOCATION	ACTOR	LENGTH	INTERVENTION	RANDOMIZATION
Karlan & Valdivia (2011)	<b>Peru</b> - Lima & Ayacucho	FINCA Peru – quite big MFI	2 years	The intervention rationale lays on the belief that self-employed poor rarely had any formal business training, thus they might lack of business skills and just by corroborating their human capital it may be possible to improve their livelihood. To some female entrepreneurs, already clients of FINCA, it was offered in addition to their group loan a business training.	Pre-existing FINCA lending groups were randomly assigned to 3 groups: mandatory treatment, voluntary treatment and control. Treated groups received a (mandatory or optional) business training after the compulsory loan weekly meetings. Subsequently the baseline survey, 2 more surveys were conducted after 1 and 2 years respectively.
Karlan & Zinman (2011)	<b>Philippines</b> – Rizal, Cavite, & the National Capital Region provinces	First Macro Bank – small local bank	+1 years	The experiment aims to evaluate the impact of improved access to credit for marginal creditworthy clients on some social and economic spheres of their life. Marginal creditworthy clients in general are not the poorest and they have some assets, they usually run a business and have some basic education.	Individuals applying at bank branches in each province were eventually labelled as marginal applicants by a credit scoring algorithm. Both treated and control participants were selected from this pool of marginal applicants. The treatment consisted in access to credit which normally would not have been possible. The baseline survey was conducted when individuals applied for loans while the endline survey took place 22 months later.
Pronyk et al. (2008)	<b>South Africa</b> - Limpopo Province (rural area)	Microfinance for Aid and Gender Equity – IMAGE program	2 years	By combining group lending (microloans for business development) with gender and HIV training the study aims to evaluate the intervention effects on HIV risk behaviours and intimate partner violence (IPV).	8 villages were paired up according to their observable characteristics similarity, then one from each pair was randomly selected to receive the credit and the training (treatment). For each woman joining the intervention, a woman of similar age and poverty status was randomly selected from comparison villages into the control group. After the baseline survey the endline one was conducted 2 years later.
Ssewamala et al. (2010)	<b>Uganda</b> - Rakai District (rural area)	Centenary Rural Development Bank – SUUBI program	10 months	The intervention aimed to evaluate the impact of economic empowerment of male and female primary school orphans on their attitudes towards sexual risk-taking behaviours. The empowerment involved the opening of a Child Saving Account (CSA) and basic financial & business training. Note that both treated and control students received the same basic medical care and education.	Primary schools in the district where randomly assigned to the treatment or control group. All orphans (male and female) attending the same school received the same assignment, hence in case of treatment they opened a CSA and received the training. A baseline survey was administered before the opening of the CSA and one was conducted after 10 months the account was activated.

PAPER	ECONOMIC IMPACT			SOCIAL IMPACT			<i>Consumption</i>	<i>Other</i>
	<i>Income</i>	<i>Savings/Credit</i>	<i>Business</i>	<i>Health</i>	<i>Education</i>	<i>Gender emp.</i>		
Karlan & Valdivia (2011)	X	X	(Null) effect on sales & revenues (Null) effect on entrepreneur ability to cope with fluctuations (Null) effect on labor supply (+) improved business skills*  *only for the less educated	X	(+) Statistically Insignificant School enrolment	(Null) effect on women decision making power	X	For the MFI (+) higher client retention rates (+) lower default rates (Null) effect on loan & savings size  (+) Statistically insignificant less child labor
Karlan & Zinman (2011)	X	(+) increased access to formal and informal credit	(-) decline in the number of business and employees	X	X	(Null) women do not appear to benefit more than men from the treatment	X	(+) higher ability to cope with risk (-) slight decline in well-being (+) stronger social ties
Pronyk et al. (2008)	X	X	X	(+) improved HIV-related knowledge (+) probability of household sex/HIV dialogue (+) more uptake of voluntary counselling & tests for HIV (+) decline of unprotected sex	X	X	X	X
Ssewamala et al. (2010)	X	(+) increased savings attached to a goal (payment of secondary school education)	X	(Null) effect on attitudes to sexual risk-taking behaviours for girls (+) improved attitudes to sexual risk-taking behaviours for boy	X	X	X	X



# LUND UNIVERSITY

School of Economics and Management

## 3.2 Microfinance impact evaluation by practitioners<sup>41</sup>

The information on the practitioners IE approach, the Social Performance Management (SPM) framework, here presented is based on an interview conducted with the microfinance expert Giampietro Pizzo<sup>42</sup> who, in addition to his standpoint on the issue, kindly recommended me to use, within others, the Social Performance Task Force (SPTF) Platform<sup>43</sup> to retrieve relevant material. For a complete list of the online sources see the Reference section (websites list).

When it comes to IE results obtained by practitioners, there is the need to distinguish between the financial and social performance of the financial providers (FPs). The former is testified in the annual audited financial statement and complementary documents required by law. These reports contain important information on FP's economic viability and they are usually based on the International Financial Reporting Standards (IFRS). IFRS is a set of rigorous and standardized principles and quantitative indicators that allows cross boundary comparison of FP's economic performance (Wardle 2014). Conversely the social performance, which encompasses both FP's economic and social impact on its clients, is often based on quantitative, non-rigorous and non-standardized data, qualitative and anecdotal information gathered or formulated by practitioners. It also tends to be over optimistic about microfinance value as a poverty alleviator tool (Collins et al. 2009; Pizzo et al. 2013)<sup>44</sup>. A practice that has casted doubts

<sup>41</sup> The microfinance sector comprehends a multiplicity of different MFIs, the IE strategy here presented has been developed and is mainly used by MFIs that pursued economic and social viability independently, thus by not receiving public funds, or where the amount of public funds is limited. Besides these MFIs are also committed to promote both economic and social development.

<sup>42</sup> Pizzo is an economist and microfinance specialist with 28 years of international experience. His knowledge and experience cover several areas with a specific focus on: financial inclusion, rural finance and SMEs finance. He worked in the field as a consultant, project manager and regional coordinator in Latin America, Caribbean, Sub Saharan Africa and MENA. He spent five years (2003-2008) at the Italian Ministry of Foreign Affairs (MAE) as Team Leader of a IMFA/UE Program in the framework of a communitarian program (FESR funds) aimed to support the implementation of regional programs of internationalization, development and cohesion. Pizzo is one of the founding members of Microfinanza Srl ([www.microfinanza.com](http://www.microfinanza.com)) and holds the position of President of the company since 2007; he is a founding member of RITMI, the Italian Microfinance Network and he has also been the vice President of the European Microfinance Network ([www.european-microfinance.org](http://www.european-microfinance.org)). The interview took place the 09/02/2017 via Skype.

<sup>43</sup> [www.sptf.info](http://www.sptf.info)

<sup>44</sup> The Afghanistan (pp. 263-271), Madagascar (pp. 427-437) and Senegal (pp. 510-521) cases in Pizzo et al. (2013) are in that respect emblematic.

on the reliability of such results (Banerjee et al. 2013; Hulme 2000; Van Rooyen et al. 2012).

As a matter of fact, during the last decade, the microfinance sector has been under attack and often accused of mission drift, opaqueness, financial and social unsustainability and other evils (Armendáriz et al. 2010; Copestake 2007; Mersland et al. 2010; Murdoch 2000). The concerns were corroborated by the collapse of the microfinance industry, during the 2010, in Andhra Pradesh, India (Pizzo et al. 2013). Such turmoil promoted debate, movement towards common standards, efforts in highlighting the best practices and a more general field reexamination and transformation. The objective was to bring back the sector to its roots and renew its international credibility by dedicating attention to FP's social performance and the SPM approach was the concrete response to that challenge (Campion et al. 2008; Pizzo et al. 2013; SPTF 2016; Wardle 2014).

SPM is a comprehensive approach that helps the FPs to define their profile and mission, pinpoint their social goals, assess their efficiency in pursuing them and eventually help institutions to improve their performance over time by changing their internal (governance, clients and employees protection, profit allocation, etc.) and/or external (services, products, delivery channels, etc.) structure. Providers' social performance, instead, reflects its effectiveness in achieving its stated social goals and creating value for clients. Usually a provider, that implements strong SPM practices, is more likely to achieve a robust social performance. The main SPM innovation consists in putting customers at the center of provider's strategies and operations. This is achieved both by delivering financial services in a way that is transparent, fair, safe, likely to generate benefits for clients (client protection) and by developing client-centric products and services (Koning et al. 2014; SPTF 2016; Wardle 2014).

In 2010 the SPTF invited all of its members and representatives of other organizations<sup>45</sup> to participate in the SPM Universal Standards (SPMUS) development. The first version of the SPMUS was published in 2012 and successive feedbacks, innovations, experiences, etc. have led to two revisions (Wardle 2014).

The SPMUS encompass six dimensions, as shown in Figure 3.1. Dimension 1 (D1), 'Define and monitor goals', demands FPs to clearly and formally outline its mission, its social goals, its targets and the indicators used to monitor the progress as well as the strategy to pursue such objectives. It also requires that the provider both collects and releases accurate client data accordingly to the stated goals. D2, 'Ensure board, management and employee commitment to social goals', consists in a series of practices providing on the one hand orientation and training on MFI mission, and goals strategies to staff at all levels and on the other hand monitoring that employees serve FP's mission.

**Figure 3. 1:** Universal Standards of SMP



*Source: SPTF (2016, p. 8)*

D3, 'Design products, services and delivery channels that meet clients' needs and preferences', involve two main aspects. Firstly the adoption of a client centric approach when it comes to design new or change existing products, services, and delivery frameworks. Secondly ensuring that, despite favoring client needs and benefits, the products, services and delivery channels are still aligned with the FP's mission. D4, "Treat Clients Responsibly", focuses on client protection, thus particular attention is addressed to the issue of over-indebtedness and, more generally, the transparent, fair and respectful

---

<sup>45</sup> Smart Campaign, the Imp-Act Consortium, MFTransparency, MicroSave, CERISE, the Progress out of

treatment of customers coupled with client data security and efficient and effective mechanisms for complaints resolution. D5, ‘Treat employees responsibly’, is dedicated to staff protection. It covers aspects such as: providing fair salaries and safe working conditions, securing worker’s rights, monitoring of employees satisfaction and turnover, supplying proper training, transparent communication of responsibilities and performance evaluation accordingly with the position held. Finally D6, ‘Balance financial and social performance’, highlights the importance of improvements in internal efficiency to achieve financial sustainability without compromising FP’s social goals and mission (Koning et al. 2014; SPTF 2016; Wardle 2014).

The implementation of the SPM approach involves a process that encompasses five main stages, as sketched in Figure 3.2. First of all there is the need to learn about SPM and to spread it within the institution, the information campaign should involve preferably all staff members at all levels. The assessment phase begins with data gathering on FP’s internal and external operations so that areas where the institution has the potential to improve can be identified. Subsequently, to achieve the targeted improvements, a strategic plan has to be tailored and this can be done with the support of a social audit tool, the SPI4<sup>46</sup>. After plan implementation the results should be reported, internally and externally, to uncover how successful where the interventions, if further efforts instead are needed and eventually highlight best-practices for the industry (SPTF 2016; Wardle 2014).

For the sake of comparison between the academic and practitioners approach D1 and D3 will be presented in a more exhaustive and practical way, but a schematic overview of D2, D4, D5 and D6 is presented in the Appendix (Part B)<sup>47</sup>. Note that any attempt of SPM implementation is grounded on two important prerequisites: the context and FP’s profile. The former involves basic information (e.g. economic and political conditions, quality of infrastructure, distance to economic centers, rural or urban area, etc.) about the zone of intervention, whereas the latter consists in basic information (e.g.

---

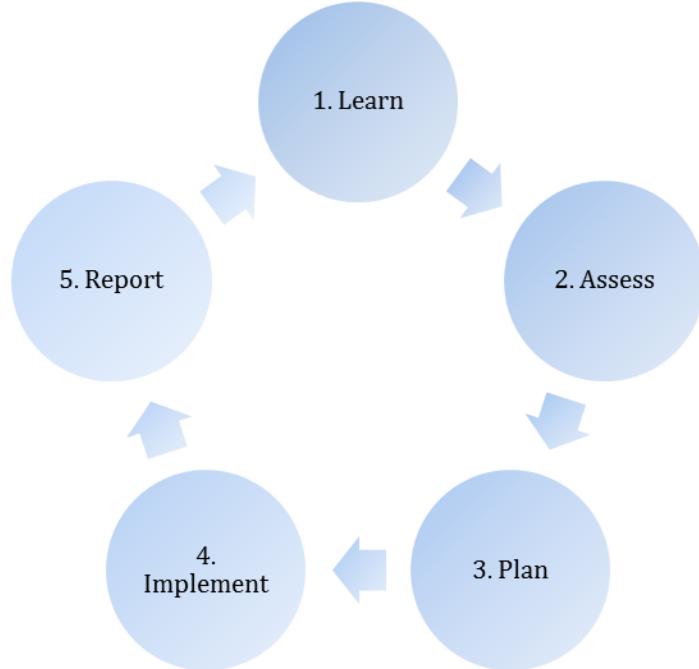
Poverty Index, MIX, and the International Labor Organization (see the reference section for the websites).

<sup>46</sup> SPI4 (Social Performance Indicators 4) is a tool developed by CERISE. It addresses all the Universal Standards Essential Practices, and it allows to benchmark the institution on each practice ([www.cerise-spi4.org](http://www.cerise-spi4.org)).

<sup>47</sup> For a more comprehensive explanation of SPMUS refer to Wardle (2014).

legal form, current budget, profit allocation, etc.) about the institution itself. Both these profiles represent an initial compass on how to direct the SPM development (Sinha 2006).

**Figure 3. 2:** SPM implementation process



*Source: SPTF (2016, p. 2)*

D1 (Define and monitor goals) has four main standards, the first implies the formulation of a good mission statement. A good mission statement is short and clear, it summarizes FP's social goals and contains information about three key elements: who is the target population, how the target population will be served, what is the expected impact. The mission statement should describe concrete, measurable, and plausible impacts and be very specific. An example of good mission statement is the following:

Provide competitive financial products that empower smallholder farmers and rural enterprises to create sustainable agri-businesses and improve their livelihoods. Furthermore female customers will receive a health training which is expected to improve household health and sanitation conditions.

*Source: Wardle (2014, p. 50)*

Another important standard is to define target clients as precisely and consistent with the context as possible. Note that client taxonomies are periodically updated and

reviewed because of either the success (failure) of the intervention or other exogenous reasons. Table 3.2 continues the above example<sup>48</sup>.

**Table 3. 2:** Requirements by client type

CLIENT	REQUIREMENTS
Smallholder farmer	> Owns less than an X amount of assets > Owns a land of maximum X hectares > etc
Female costumer	> Any women actively using/purchasing financial product or service

*Source: Wardle (2014, p. 51)*

For each costumers' category it is necessary to indicate also the specific strategies and products that are offered to them, elements that are covered in the two other standards of D1: the definition of social goals and targets. Beforehand it is, however, necessary to distinguish between outputs, whom the FP will serve, and outcomes, how the FP will serve them: both are part of the social goals (Table 3.3). Moreover, each goal must be attached to a target. The following tables give respectively practical examples (Table 3.4) and an overview of the SMART properties a target should have: Specific, Measurable, Achievable, Relevant and Time-Bound (Table 3.5).

**Table 3. 3:** Outputs and Outcomes

OUTPUTS	OUTCOMES
Provide crop insurance to smallholder rural farmers	Reduce vulnerability of rural farmers to shocks
Provide health training to female clients	Clients will improve health and sanitation practices

*Source: Wardle (2014, p. 52)*

The final D1 standards refers to the data which is a central element. Indeed the FP should know which information is needed and how it can be collected, since data mining is both time consuming and costly. The FP should consider also the following aspects: what resources are available for data collection, which data exists already and is eventually available, the frequency of data flows, what client sample the data should be based on, how burdensome the data gathering will be for clients and staff. A second important step is the definition of data protocols: which employees will collect and analyze the information, where the data will be stored, which employees will verify the accuracy and reliability of information, how the data will be reported and finally to and for whom it will be reported. Two further caveats are worth noting. Firstly, if the FP

---

<sup>48</sup> Note that a small farmer in a certain context may differ from another farmer with the same amount of land

employees are responsible for data gathering, then professional training should be supplied (alternatively an external expert can be used). Secondly, given the resources and expertise available, the data should be disaggregated as much as possible to obtain the highest level of segmentation (SPTF 2016; Wardle 2014).

**Table 3. 4:** D1 in practice

SOCIAL GOALS (Outputs & Outcomes)	TARGETS
Provide crop insurance to smallholder rural farmers	> X% of all clients this year will be farmers owning assets less than US\$ X & owning less than X hectares of land. > X% of all new clients in this group will purchase FP's crop insurance product.
Reduce the vulnerability of rural farmers to shocks	> Rural farmer reporting financial distress from unexpected crop loss will decrease by X% in t0 & by another X% in t1
Provide health training to female clients	> X female clients will X sessions of health training during a period t
Female clients will improve health and sanitation practices	> Verbal tests administered t months after completion of the training will show that at least X% of trainees are able to describe X health and sanitation procedures correctly (e.g. malaria prevention). > Surveys administered t months after the program will show a decrease in the number of work days missed due to illness, compared to baseline survey results

Source: Wardle (2014, p. 53)

**Table 3. 5:** SMART characteristics

CHARACTERISTIC	QUESTIONS
<i>Specific</i>	> Which type of clients and employees are involved in the target > The target is clear and transparent
<i>Measurable</i>	> Will the FP be able to assess changes? > Are quantitative measures used? If yes which ones & how they are used > Are qualitative measures used? If yes which ones & how they will be measured > Can the FP realistically collect the data required?
<i>Achievable</i>	> Can the FP meet the targets with the budget? > Which changes are needed to achieve the targets? (e.g. strategies, products, etc.)
<i>Relevant</i>	> Is the target consistent with the mission? > Is the target consistent with one or more social goals?
<i>Time-Bound</i>	> What is the reasonable time frame within which this target is achievable?

Source: Wardle (2014, p. 54)

---

and revenues in another location.

D3 (Design products, services, delivery models, and channels that meet clients' needs and preferences) is based on two main standards: the first is that the institution should understand the needs and preferences of different types of clients. This standard in turn involves three essential practices: data gathering on both client (e.g. their behavior, their day-to-day and life-cycle financial needs, their economic and social opportunities or barriers, common risks and emergencies they are exposed to, their satisfaction towards current services and products, suggestions for product and service improvements, etc.) and market (e.g. number of competitors, economic and political conditions, available technology and infrastructure, etc.). Although there are multiple ways of retrieving data, it should be ensured the possibility to segment the data, cover a representative sample of clients and non-clients and usually have neutral parties analyzing the information. Another essential practice is monitoring client satisfaction through multiple channels of investigation, such as interviews, field observations, surveys, discussion and meetings (level of trust in the institution, actual user of the products and services, satisfaction with the timeliness, etc.). The third practice requires awareness of client exit (attrition for academics) and the understanding of its reasons. Indeed, clients' exit may be due to dissatisfaction, graduation (i.e. clients successfully improve their living conditions and do not belong anymore to the target group) or other reasons<sup>49</sup>.

The second standard of D3 suggests that: the institution's products, services, delivery models, and channels are designed to benefit clients, in line with the institution's social goals. It involves five main practices: first of all it demands for a credit-centric approach when it comes to designing products and services. Therefore all the main aspects (size, delivery channels, fees, delivery modes, contract terms, etc.) are considered from the client perspective. This practice is strongly related to the following three ones: understanding of client needs and preferences to reduce their barriers to financial inclusion; delivery of products and services that allow clients to reduce their risk and cope with emergencies; creation of benefits for clients by enabling them to invest in economic opportunities and address anticipated household needs. Indeed a more thorough knowledge of the client is likely to help the FP to tailor efficient and effective financial tools that reduce clients' constraints and risks, cope with clients' emergencies

---

<sup>49</sup> In order to avoid biases in the data, the dormant clients should be included as well. They are clients who, in

and needs, and ultimately foster their ability to seek economic and social opportunities. A final essential practice, although the FP should promote new and existing products, is that the sale strategies should always respect the client's right to refuse (SPTF 2016; Wardle 2014). Table 3.6 gives a practical example on D3.

**Table 3. 6:** D3 in practice

PRODUCT OR DELIVERY CHANNEL	DATA MINING ON PRODUCT OR DELIVERY CHANNEL	ANALYSIS OF THE DATA TO UNDERSTAND CLIENT NEEDS	DELIVERY CHANNEL DECISION BASED ON CLIENT NEEDS
<p><b>Delivery channel options:</b></p> <ul style="list-style-type: none"> <li>&gt; Branchless banking</li> <li>&gt; Traditional banking (branch)</li> </ul>	<p><i>Segmented by client characteristics</i></p> <ul style="list-style-type: none"> <li>&gt; Client distance from branch offices, by region</li> <li>&gt; Mobile phone ownership, by client income or poverty level</li> <li>&gt; Delivery channel preferences (client survey)</li> </ul>	<p><i>Segmented client data show:</i></p> <ul style="list-style-type: none"> <li>&gt; Clients in the eastern &amp; western provinces live, on average, 2km and 10km away respectively from their local branch office.</li> <li>&gt; 85% of clients own mobile phones (both provinces). Of these, 70% are “poor.”</li> </ul> <p><i>Additional market data show:</i></p> <ul style="list-style-type: none"> <li>&gt; Clients prefer the mobile banking (both provinces)</li> <li>&gt; Western clients spend 3 times more on transportation to the closest branch compared to eastern clients</li> </ul>	<p>The FP decides to pilot mobile banking, as a majority of its clients have expressed a preference for it.</p> <p>It begins with some branches located in the west, as these clients live further from their local branch &amp; spend more money on transportation</p>
<p><b>Product:</b></p> <ul style="list-style-type: none"> <li>&gt; Agricultural insurance</li> </ul>	<ul style="list-style-type: none"> <li>&gt; Demand by region</li> <li>&gt; Asset value by business type (e.g. farmer or livestock owner)</li> <li>&gt; Preferences on product terms by business type (e.g. interview)</li> <li>&gt; Responses on household needs by business type (e.g. survey)</li> </ul>	<p><i>Segmented client data show:</i></p> <ul style="list-style-type: none"> <li>&gt; Clients living in the north suffer from natural disasters at 3 times the rate of the southern clients</li> <li>&gt; Farmers' assets are worth twice as much as livestock owners' assets.</li> </ul> <p><i>Additional market data show:</i></p> <ul style="list-style-type: none"> <li>&gt; Farmers want to insure both their crops &amp; farming inputs.</li> <li>&gt; Livestock owners already have access to government insurance for their assets.</li> </ul>	<p>The FP adjusts its current agricultural insurance product to target <i>farmer</i> clients rather than <i>livestock</i>.</p> <p>The basic product covers crops, but farmers are given the option to purchase supplemental insurance for their farming inputs.</p> <p>Farmers in the north are offered additional coverage specifically for natural disasters.</p>
<b>Product:</b>	> Deposit size and	<i>Segmented client data show:</i>	The FP pilots 2 savings products:

spite of being formally clients, have not demanded or required any service and/or product for a while.

> Voluntary savings	frequency by client age > Product term preferences by business type & age (e.g. interview)  > Responses on household needs by business type & age (e.g. survey)	> Younger clients (21-35 years) make larger, less frequent deposits  <i>Additional market data show:</i> > Income flows are more variable for younger clients  > Younger clients save mainly for housing & weddings	wedding & home purchase savings  The accounts offer higher interest rates but require larger minimum deposits that must be made at least twice a year.  Clients may withdraw savings after a minimum of one year.
---------------------	--	--	---

Source: Wardle (2014, p. 94)

In conclusion SPM is a flexible and adjustable framework, based on specificity and transparency that can be used by heterogeneous FPs. After outlining its organizational profile and the regional profile where it is operating, the FP is required to state its mission, target clients, social goals, SMART targets and data protocols (D1). Furthermore the FP should ensure that its internal (D2, D4 and D5) and external structures (D3) are aligned to its mission. Alignment to the mission demands also a balance between FP's financial and social performance, and in the relationship between investors and FP (D6); such balance is only possible through a transparent communication of intentions. In other words the implementation of SPM allows the FP to create an *a priori* frame within which conducting an IE is feasible. The IE will not be perfect because of measurement errors and because some effects (either positive or negative) will fall outside such a frame. However, since FPs usually have a long life, the SPM approach becomes a continuous and repeated process of data mining, evaluation, planning, adjustments implementation, data gathering, evaluation, planning and so on (Figure 3.2). Therefore the FP, after each cycle, is able to fine-tune targets consistently to its stated social goals and eventually make adjustments in its internal and external structures<sup>50</sup>. In other words it can expand the *a priori* frame to capture previously excluded positive effects and implement changes to reduce the negative effects (Campion et al. 2008; Sinha 2006; SPTF 2016; Wardle 2014).

---

<sup>50</sup> A self-updating mechanism that reflects the critical rationalism in Caldwell (1991) preventing the strategy to become perpetual even if inefficient or obsolete.

## IV. Discussion

Now that both the academic and practitioners microfinance IE have been introduced, a closer comparison and the implications of these two different approaches are discussed in this section. Moreover an attempt to bridge the gap between science and reality is outlined, attempt based on the merger of these two apparently contrasting methodologies.

The academic approach, i.e. the rigorous quantitative hypothetico-deductive model combined with a reductionist perspective (RCTs), aims primarily to prove causality, in this specific case whether microfinance is able to reduce poverty and more generally promote social and economic development. As already highlighted due to the complexity and heterogeneity that characterizes the microfinance field, the ability of RCTs to prove causal relationship between intervention and impact (i.e. attribution) is undermined. Furthermore since RCTs have an experimental design, they can at best determine if something (an intervention) is effective rather than explaining why it is effective. YES (effective) and NO (ineffective or harmful) answers do not enhance the understanding of the phenomena and its mechanisms, consequently they also do not supply any guidelines to reduce the intervention detrimental effects and amplify the beneficial ones. The intrinsic reductionism required in experimental designs may help in dealing with microfinance complexity and heterogeneity, but undercuts the external validity of its results. In fact, even putting aside these issues, the RCTs inferences remain very specific. At best it is possible to conclude that e.g. given a specific context, specific beneficiaries, a specific intervention, a specific financial provider, within a certain time frame and under a certain methodology and hypotheses, a specific microfinance tool appears more or less effective than expected. Secondly and more striking, the fetishism of statistical and methodological rigorousness often demands an experimental setting which is far from reality and any common sense. The following metaphor discussed with Pizzo during the interview is in that respect informative:

*"If tomato seeds are planted in the desert, without any irrigation, fertilizer or any other alternative support and care, we would not be surprised if they do not produce any fruit or if they do not blossom at all"*

Rephrasing the metaphor, it should not come as a surprise if microfinance products (tomato seeds) are not successful when implemented in adverse and isolated contexts (desert, heat, lack of water) where there are no complementary activities or opportunities, (forward and backward linkages to markets) or where there is not a proper financial or technological infrastructure (irrigation system, fertilizers, etc.) that can foster its effectiveness (plant propensity to grow and produce fruit). The example is, in turn, strictly related to the excessive and tunnel-vision attention researches dedicate to the internal validity. Indeed solutions often imply technical improvements addressing randomization, control of selection bias, simultaneity, etc., which however lack thorough microfinance understanding and common sense. Implementing the experiments in isolated and adverse areas where there are no credit alternatives, where the closest market and complementary businesses are many kilometres away and there are no paved roads and reliable means of transport, where people never even had access to credit before may enhance the statistical robustness, but it is not realistic (Pizzo interview 09/02/17). Indeed FPs usually conduct ex ante feasibility studies and, after identifying their target clients and feasible social goals, design the supposed beneficial client-specific tools (see in D1 and D3 in section 3.2). In that respect, the overall weak microfinance impact, found by many of the papers reviewed in section 3.1, should not be surprising once the reader is aware of their 'poor' epistemological foundations. This is reflected in e.g. the variable selection. In academic publications the variables selection is based on previous research and theories and no insight on the selection process and the rationale behind it is presented. For practitioners the choice of indicators is an extremely delicate task, indeed indicators must be both consistent with the stated social goals and have the SMART (Specific, Measurable, Achievable, Relevant and Time-Bound) properties. Thus expecting a microfinance effect on gender empowerment after one or two years is neither achievable nor well time-bounded from a practitioners' standpoint, simply because it is not

realistic<sup>51</sup>. Also the interpretation of results is more thoroughly contextualized by practitioners than by researchers, which is very important. For example, in rural patriarchal communities with no or very low literacy rates, where women previously had no decision power whatsoever and after the intervention they have systematically a say on intra-household decision, we can speak of a massive impact. This is so independently of the absolute number of women that achieved such degree of emancipation, independently of the fact that, if compared to western women, that have a broad range of freedoms and rights, the result is not impressive anymore (ISPI Course 2014)<sup>52</sup>. Similarly, selection bias, a big issue in academic research, is for practitioners an integral stage of their operational strategies, as seen in section 3.2. A further example that unveils the gap between reality and science is the attrition problem. For researchers attrition is connected to two main issues: one regarding the sample size and thus its ability to return statistical significant results, the other if the drop-out is random within participants, random in terms of their individual observable characteristics. Conversely practitioners call attrition “client exit”, which represents also a concern, but a concern that has to be monitored, understood and eventually solved in subsequent cycles. These are the types of inconsistencies present in the scientific publications, deriving from a poor understanding of how actual microfinance works. The above points uncover another problem with academic research: the studies are one-time evaluations, fact that undermines the ability to capture the long run effects and follow the economic and social development of both the communities and the clients. At the same time, the one-time feature, although it does not directly compromise RCT’s rigor, replicability and falsifiability, limits the room for improvement and cumulative understanding of both the methodology used and the phenomenon analysed. Indeed, on the one hand, despite all the limitations underlined, the role of RCTs in microfinance IE does not appear downsized. This contrasts with the self-updating mechanism so warmly advocated by Caldwell (1991), which is supposed to prevent the methodologies to become a ‘golden standard’ even if inefficient, inadequate or obsolete. On the other hand the lack of a cumulative understanding has not only

---

<sup>51</sup> The lack of impact is not necessarily connected to microfinance intervention itself, but the selection of the indicators, the feasibility of obtaining a cultural change in such limited amount of time.

<sup>52</sup> In fact in the long term empowered women are likely to be followed by other women in the same or nearby communities, generating a spillover effect (ISPI Course 2014).

undermined the practical use of academic research, but in the end also the reliability and relevance of its contribution to scientific knowledge.

Turning to the SPM, Copestake (2007) places the approach among the Informal Impact Assessment (IIA) methodologies. IIA strategies are looser and based on interpretation of data and analogy (Copestake 2007; Verschuren 2001). The data involved is usually both qualitative and quantitative, although the latter is not analysed through rigorous statistical procedures, and comes from a multiplicity of sources (e.g. direct observations, information systems, staff experiences, client complaints, surveys, etc.). According to Copestake (2007), a study, in order to be considered an IE, needs to have a systematized framework of interpretation and to directly tackle the problem of attribution, elements that are both present in the SPM approach. However due to its intrinsic features, flexibility and specificity, the SPM also lacks of sound external validity. Indeed the SPMUS (SPM Universal Standards) are strongly bounded, among other things, to the institution's profile, mission, target clients, social goals and the context where it operates. It is also true that the SPM *raison d'être* has no scientific root. In that respect two aspects are worth noting. Firstly FPs are interested in IE, ultimately in causality, mainly because they need to prove that their activity is impactful and beneficial in order to attract prospective clients and investors and to retain the current ones. Secondly attribution and understanding (i.e. answering the why question) are more directed towards supplying the FP with *ad hoc* insights and solutions on how to enhance its positive impact and limit the negative one through adjustments in its internal and external structure rather than coming up with scientific knowledge. However, although such solutions are institution-specific, if shared in conferences, platforms and other networks, they can foster capacity building, in particular if they are carefully tailored and adapted, and eventually become best practices. Consequently, the SPM framework may not be scientifically rigorous, but has at least practical use. The former aspect is also strictly related to the SPM internal validity, which is based on institutions ethics and professionalism. Since the SPM approach lacks of a rigorous and impartial methodological design an obvious question arises: how reliable are the reports and results published by these institutions? First of all FPs are encouraged to hire, or to be assisted by, external professionals in the most delicate aspects of SPM, e.g. when conducting a survey and

during the data interpretation process. However and more importantly it exists a system of external monitoring for FPs. The network includes national agencies designated to the regulation and supervision of financial markets, a set of formal requirements like the annual audited financial statement, licences and other documents, and rating institutions. Currently MicroRate (US), MicroFinanza Rating (Italy), Planet Rating (France) and MCRIL (India) are the most relevant rating institutions and another important source of information is the MIX, a platform founded in 2002 by the CGAP which provides sectoral information mainly on FPs specialized in supplying low-income individuals around the world (Pizzo et al. 2013)<sup>53</sup>. Nevertheless, independently on how objective and strict is such external rating and supervision, it is not completely free of flaws as the 2008 crisis and the microfinance sector breakdown in India have shown. Institutions ethics and professionalism still matter a great deal, although both are aspects that are ultimately not possible to control (Pizzo et al. 2013; Presbitero 2009)<sup>54</sup>. Apart from these undeniable internal and external limitations, the SPM approach appears to have a more solid epistemological base<sup>55</sup>. Because of the way the SPM is structured, it requires an *a priori* strong understanding and awareness of the microfinance field, the institution, the context and the clients which is achievable through feasibility studies and/or long standing expertise. Subsequently such knowledge allows to ‘forecast’ impacts, plan the intervention thoroughly and monitor its development. As presented in section 3.2, the choice of social goals and targets (i.e. variables or indicators for academics) within the SPM framework, is much more philosophically grounded compared to the academic approach, or at least is intended to be so. Indeed the rationale behind the choice procedure reflects to a certain extent the complexity and holism of thought of the ‘capability approach’ theorized by A. Sen during the 80s, that is a commitment to

<sup>53</sup> These rating institutions are specialized in the microfinance sector, thus they have developed alternative assessment standards that suit more the microfinance industry, known as Social Rating. For a more thorough understanding of social rating see Clark et al. (2013) while for a brief overview on rating see Pizzo et al. (2013, p 711-718).

<sup>54</sup> In the case of microfinance, three main reasons emerge as controversial: first of all the data on which these external evaluations are based on is from the assessed FPs. Secondly the experts working in the rating institutions have previously often worked in one or more of the FPs they are currently evaluating. Finally FP can conduct illegal practices to blur important and sensible information (Pizzo et al. 2013).

<sup>55</sup> Note that it is not completely correct to use epistemology when talking about the practitioners approach. In fact epistemology is the branch of science that addresses the conditions under which it is possible to obtain scientific knowledge or in a narrower sense is the ‘philosophy of science’, thus the discipline that embraces the grounding hypothesis and methods of the scientific disciplines (Latouche 2010). The SPM approach is dearly not

embrace a broader view and set of elements in the assessment (Robeyns 2004; Robeyns 2006; Sen 1994)<sup>56</sup>. Consequently a microfinance intervention that creates access to financial services, thus an opportunity, it is already successful, from a ‘capability approach’ standpoint, independently on how many will decide take it. Furthermore since FP’s activity is a process, the learning and understanding is path dependent and cumulative, but not necessarily always incremental. Therefore, within this frame, the SPM allows, after each intervention or cycle, to eventually enhance FP’s activity, thanks to the new insights, by implementing consistent adjustments. The evaluation as a process not only augment the ability to capture long term effects, but by being based on a more continuous flow of data and information, it is also closer to the actual social and development progression. Because of its solid philosophical roots, flexibility, specificity and pragmatism SPM eventually allows e.g. an FPs to support clients along all their journey from being well below the poverty line and without access to any type of credit, to have enough assets and credit history to be able to access the formal credit market, where collateral and other forms of guarantee are required. Such development requires clients’ needs and preferences to be frequently monitored, current products and services to be evaluated and eventually tailored to cope with new challenges, especially when clients ‘upgrade’ (SPTF 2016). Finally the introduction of SPMUS has improved the worldwide replicability and rigorousness of the approach and its intrinsic updating mechanism, indeed the SPM process feature guarantees falsifiability in a broader sense (i.e. through continuous improvements)<sup>57</sup>. Hence the SPM framework appears to satisfy somehow the scientific requirements outlined by Popper.

None of the two approaches is exempt from limitations and concurrently both have some key strengths. Actually by their pros and cons, they appear more complementary

scientific, however its ‘philosophical’ foundation appears stronger than the academic one, which defines itself as scientific.

<sup>56</sup> The ‘capability approach’ was introduced by Sen as an alternative metric for well-being evaluation. Given the available resources (e.g. food) a set of capabilities (e.g. eating or starving) is determined for each person, but only some of them will be transformed in functionings (e.g. being nourished or malnourished) and generate an utility for the individual. The functionings are defined as the states of ‘being and doing’ for an individual, whereas the capabilities are the set of functionings that a person can effectively achieve. Other important elements are human agency and freedom, the former involves a choice, indeed starving and fasting are intrinsically different, whereas the latter is the ultimate requirement to both make a choice and the concrete possibility to transform resources and capabilities in functionings (Robeyns 2004; Robeyns 2006; Sen 1994).

<sup>57</sup> Falsifiability refers primarily to the hypothesis testing involved in scientific research, as introduced in section 2.1. However the SPM approach allows after each cycle to review and eventually change, before a new cycle starts, the indicators, social goals, products and services, target clients and even the mission.

than competitive alternatives, thus by promoting dialogue, cooperation and partnerships between academia and practitioners it may be possible to trigger an interesting exchange of views, expertise, knowledge, skills and practices. The SPM framework has the advantage of answering more successfully the ‘why-question’, having a more solid philosophical background and being both more pragmatic and falsifiable. Although the publication of the SPMUS has increased SPM degree of replicability and rigorousness, they still remain the weakest points of the approach. The quantitative hypothetico-deductive model has the main advantage of being easily replicable, falsifiable and more rigorous and objective, while its ability to cope with ‘why-question’, its epistemological background and its pragmatism are limited. Having said that, since the SPM approach is the result of a thorough and broader understanding of the microfinance field, it should be adopted as IE guideline by the academic environment (Firebaugh 2008). The collaboration however can be beneficial also for practitioners, in fact researchers can improve the reports by adding quantitative analysis and rigorousness. For example, in the stage of data mining researchers are likely to be more aware of all the survey biases and related issues, thus they can help in implementing procedures that lead to higher quality of data (Hulme 1997; Rossi et al. 2013). Similarly by using the data, the clients (sample), the products and services (intervention) and targets (variables or indicators), pinpointed after a SPM assessment and thus more philosophically grounded, the researchers can conduct RCTs or econometric analysis to verify if FPs’ results are also statistically corroborated. The participation of the academia in the SPM process is likely to strengthen further the reliability, objectivity and transparency of FP’s internal assessment. Concurrently academic research will benefit from a pragmatic epistemological base capable of reducing its detachment from reality and therefore capable of re-evaluating its overall usefulness and contribution to scientific knowledge.

## V. Conclusions

Before presenting the conclusions, it is worth highlighting the main limitations of this thesis. Firstly the selection of microfinance IE publications could have been based on a more structured and rigorous protocol, which however would have required an unfeasible effort, because of the limited availability of time, resource and knowledge. Secondly it would have been more interesting to conduct multiple interviews with experts and practitioners. Unfortunately to contact them was time consuming and in almost all the cases unsuccessful. The final weaknesses regard the limited number of sources from where the material on SPM was retrieved and the fact that SPM framework has not yet been universally adopted by the microfinance sector. Despite these limitations, the elements discussed in this dissertation appear sufficient, at least for the microfinance case, to unveil the existing gap between reality and science, the striking differences between the academics' and practitioners' IE approaches and to allow the formulation of some non-normative inferences about the real implications of such a gap.

Within the microfinance case, a gap between reality and science appears to exist and to be quite relevant. In fact the practitioner's and academic's approach differ not only in the results, more optimistic the former whereas more pessimistic the latter, but also methodologically (Hulme 2000). The SPM framework, in spite of being more flexible, very specific, based on both qualitative and quantitative data and less rigorous in the interpretation, seems to satisfy the scientific requirements suggested by Popper. Indeed on the one hand the efforts to standardize the methodology (i.e. SPMUS) have progressively made SPM a more replicable and rigorous approach and on the other hand, because of its process feature, the framework is also intrinsically falsifiable (self-updating). Moreover SPM provides FPs with practical solutions and an increasingly deeper microfinance understanding (i.e. of causality, but not in the strict scientific sense). Conversely RCTs, despite following a rigorous and quantitative hypothetico-deductive model, are limited by the reductionism trap and distance themselves from reality and, by doing so, also from science. In other words they fail to prove the causality that matters. Indeed strong hypothesis, experiment design adjustments and unrealistic experiment settings undermine both the scientific validity and pragmatism of RCTs results. Besides

RCTs are replicable and falsifiable: not the methodology in itself, but rather the research hypothesis which can be easily challenged. Indeed although RCTs flaws are well acknowledged, they have remained the microfinance IE ‘golden standard’ for many years. More strikingly between the two approaches there is a big gap in terms of microfinance understanding and consequently also in their philosophical solidity. Some choices and strategies adopted by researchers on time frame, variables, targets, etc. clash with FPs’ policies, because actually they are neither part of microfinance at all nor realistically achievable, measurable, etc. This raises concerns on the validity of RCTs results, even if they would be the outcomes of a flawless methodology.

The attention dedicated to such a gap is justified because it matters and it has significant real implications, as it will become clearer. Before continuing it is worth quoting once again Hicks (1941 cited in Scropanti et al., 2005):

*[...] Economics is not a thing on its own; it is the study of an aspect of the life of man in society  
 [...] The economist of tomorrow (and sometimes of today) will certainly be aware of what to ground his economic advice on; but if [...] his economic knowledge is detached from any background of social philosophy, he runs a real risk of becoming a cheat, capable of implementing enterprising stratagems to find his way out of difficulties, but incapable of staying in contact with those fundamental virtues on which a sound society is founded. Modern economic science is subject to a real risk of Machiavellianism: the treatment of social problems as mere technical issues and not as an aspect of the general quest for the Good Life.*

IEs can influence governments, investors and public opinion, in particular publications that are labelled as scientific, because of the credibility they have. Specifically, when it comes to the Official Development Assistance (ODA) to developing countries, pessimistic and sceptical IEs are translated into a reduction of funds, capital and investment in such a sector. This has a detrimental effect not only on the sector itself (e.g. loss of jobs), but also negative consequences for current and prospective beneficiaries who would not be able to access credit or other financial services, in the case of microfinance. Thus conducting a proper and sound IE is as important as obtaining reliable results so that funds are allocated to the most efficient institutions and/or interventions (Hulme 2000; ISPI Course 2014; Pizzo interview 09/02/2017).

Fund allocation decisions have a real impact because MFIs and other organizations implementing development projects, although having a limited scope and reach, do usually generate a positive impact on people's life. The fact that, taken singularly or altogether, these initiatives do not eradicate global poverty, does not mean they do not have an impact and that they do not e.g. take many of their beneficiaries out of poverty. As an outsider, it is difficult to perceive the impact of even a simple programme that allows a few children to progress in their education or helps in the purchasing of some mosquito nets. However the former basically gives a future to younger generations<sup>58</sup>, while the latter actually 'saves life' (ISPI Course 2014). In these situations, low absolute headcount numbers should be interpreted with care, because the results refer to human lives and not to income, interest rates, GDP, etc. Similarly probabilities should be contextualized, a higher or lower likelihood of opening a business should be treated differently from the higher or lower probability of conducting HIV risky behaviours (like unprotected sex), which is directly connected to human life.

Recently seven new potentially habitable planets have been discovered. Although such a finding may have an important impact on human life in the next centuries, if in the following decades the discovery is proven wrong, no immediate consequences will affect our daily lives (Strickland 23/02/2017, CNN). Conversely, if financial resources stop being allocated to development institutions it can have dramatic consequences in the short run. Microfinance plus interventions, like the ones reviewed in Pronyk et al. (2008) and Ssewamala et al. (2010), that combine saving and credit services with HIV<sup>59</sup> related training, do save lives today independently of the statistical significance of results. Likewise in Dupas et al. (2013b), where one of the saving technologies was attached to a health goal: the purchase of a mosquito net to prevent malaria<sup>60</sup>. Awareness and attention

<sup>58</sup> If one member of the family manages to graduate from high school, move to the city, enrol to university and finally obtain a high paid job, it is a success. Such a family member represents a sort of family insurance by means of his salary and concurrently a networking resource, indeed he can host and support other relatives when they want to move to the city, he can pay the tuition fees for education and so forth, generating thus a sort of spillover effect (Armendáriz et al. 2010; ISPI Course 2014; Massey et al. 1993).

<sup>59</sup> Since the HIV epidemic started in the mid-80s more than 70 million people have been infected with the virus, of which around 35 million have died (WHO, [www.who.int](http://www.who.int); AVERT, [www.avert.org](http://www.avert.org)).

<sup>60</sup> Malaria is an infectious disease and one of the leading causes of death, especially in the developing countries. Although preventable and treatable it causes over than 438,000 deaths every year, mainly in sub-Saharan Africa.

should be involved when interpreting the results, their social and economic significance. In the planet case, bold scientific statements may not have an immediate and real effect, but when it comes to human lives inferences should be more careful. Saving even a single human life is important, despite in technical terms it may be labelled as a ‘null effect’.

The following example is more related to reviewed papers dealing mainly with the economic impact of microfinance interventions. Such interventions aim to improve access to credit or other financial services, and they indirectly promote the economic emancipation of beneficiaries. Many of the immigrants attempting nowadays to cross the Mediterranean Sea are defined as economic immigrants, especially the ones coming from Northern Africa. They are individuals seeking better economic and life opportunities who however put at risk their lives and safety in pursuing that objective. Development projects in countries of origin aim to create a fertile economic and social environment and by doing so, among other things, they also reduce the incentives to migrate under such hazardous conditions. When funds are moved away from ODA, the ability to deal with this migration crisis is undermined, in particular the ability to deal with its structural causes (Rizzi 12/05/17, *Il Fatto Quotidiano*). The consequences of such events are there for all to see, real and dramatic. These examples unveil the importance of conducting sound IEs. Only by doing so, it will be possible to spot the most effective interventions and simultaneously prevent funds to fly away from development initiatives because of reports resulting from both poor methodological and epistemological evaluations.

Fortunately the practitioners and academic approaches are more complementary than initially expected. Collaboration between the two environments can benefit practitioners SPM by improving its rigorousness and quantitative analysis (e.g. statistical analysis) involved. Concurrently it can benefit academia as well with a higher degree of realism, pragmatism and philosophical solidity. The combination of the two approaches, with however SPM as a guideline, seems capable of strengthening both the scientific and practical validity of microfinance IEs, and ultimately bridging, or at least reducing, the gap between science and reality.

---

The most vulnerable individuals are pregnant women and children under five years of age (Malaria Consortium, [www.malariaconsortium.org](http://www.malariaconsortium.org)).

## Reference

### Bibliography

- Abbagnano N., Fornero G. (2002a). *Itinerari di filosofia. Vol. 1: Dalle origini ad Aristotele-Dall'ellenismo alla scolastica.* Paravia.
- Abbagnano N., Fornero G. (2002b). *Itinerari di filosofia. Vol. 2: Dall'umanesimo all'empirismo-Dall'illuminismo a Hegel.* Paravia
- Angelucci, M., Karlan, D., & Zinman, J. (2012). *Win some lose some? Evidence from a randomized microcredit program placement experiment by Compartamos Banco.* J-PAL working paper.
- Armendáriz B., Morduch, J. (2010). *The economics of microfinance.* MIT press.
- Artoni, R. (2005). *Elementi di scienza delle finanze.* Il Mulino.
- Ashraf, N., Giné, X., & Karlan, D. (2009). *Finding missing markets (and a disturbing epilogue): Evidence from an export crop adoption and marketing intervention in Kenya.* American Journal of Agricultural Economics, 91(4), pp. 973-990.
- Attanasio, O., Augsburg, B., De Haas, R., Fitzsimons, E., & Harmgart, H. (2011). *Group lending or individual lending? Evidence from a randomised field experiment in Mongolia.* WZB Discussion Paper (No. SP II 2014-303).
- Augsburg, B., & Fouillet, C. (2013). *Profit empowerment: The microfinance institution's mission drift.* Studies of impact and performance: the credibility of microcredit, 199.
- Augsburg, B., De Haas, R., Harmgart, H., & Meghir, C. (2015). *The impacts of microcredit: Evidence from Bosnia and Herzegovina.* American Economic Journal: Applied Economics, 7(1), pp. 183-203.
- Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2013). *The miracle of microfinance? Evidence from a randomized evaluation.* American Economic Journal: Applied Economics, 7(1), pp. 22-53.
- Bauchet, J., Marshall, C., Starita, L., Thomas, J., & Yalouris, A. (2011). *Latest findings from randomized evaluations of microfinance.* In Access to finance forum (Vol. 2). Washington, DC: CGAP.
- Blaug M. (1980), 'From the received view to the views of Popper', in The methodology of economics or how economists explain, Cambridge: Cambridge University Press, pp. 1-29.
- Becchetti, L. (2008). *Il microcredito: una nuova frontiera per l'economia.* Il mulino.

- Caldwell, B. J. (1991). *Clarifying popper*. Journal of Economic literature, 29(1), pp. 1-33.
- Campion, A., Linder, C., & Knotts, K. (2008). *Putting the 'social' into performance management*. Brighton, Institute of Development Studies at the University of Sussex.
- Capra, F. (1992). *Le Tao de la physique*. Sand.
- Clark H., Sinha F. (2013). *Social Rating Guide*. The Rating Initiative. Retrieved: [https://sptf.info/images/social%20rating%20guide\\_english\\_nov%202014.pdf](https://sptf.info/images/social%20rating%20guide_english_nov%202014.pdf) [Accessed 13 April 2017]
- Collins, D., Morduch, J., Rutherford, S., & Ruthven, O. (2009). *Portfolios of the poor: how the world's poor live on \$2 a day*. Princeton University Press.
- Copestake, J. (2007). *Mainstreaming microfinance: Social performance management or mission drift?*. World Development, 35(10), pp. 1721-1738.
- Copestake, J. (2012). *The Heterogeneity of Microfinance Impact and its Assessment*. In UMM Workshop Report.
- Crépon, B., Devoto, F., Duflo, E., & Parienté, W. (2011). *Impact of microcredit in rural areas of Morocco: Evidence from a Randomized Evaluation*. MIT Working paper.
- Deaton A. (2010). *Instruments, randomization, and learning about development*. Journal of Economic Literature, 48(2): pp. 424-55.
- De Mel, S., McKenzie, D., & Woodruff, C. (2009). *Are women more credit constrained? Experimental evidence on gender and microenterprise returns*. American Economic Journal: Applied Economics, 1(3), pp. 1-32.
- Di Vittorio, A. (2002). *Dall'espansione allo sviluppo: una storia economica d'Europa*. G. Giappichelli.
- Dupas, P., & Robinson, J. (2013a). *Savings constraints and microenterprise development: Evidence from a field experiment in Kenya*. American Economic Journal: Applied Economics, 5(1), pp. 163-192.
- Dupas, P., & Robinson, J. (2013b). *Why don't the poor save more? Evidence from health savings experiments*. The American Economic Review, 103(4), pp. 1138-1171.
- Duvendack, M., Palmer-Jones, R., Copestake, J. G., Hooper, L., Loke, Y., & Rao, N. (2011). *What is the evidence of the impact of microfinance on the well-being of poor people?*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

- Fafchamps, M., McKenzie, D., Quinn, S., & Woodruff, C. (2014). *Microenterprise growth and the flypaper effect: Evidence from a randomized experiment in Ghana*. Journal of development Economics, 106, pp. 211-226.
- Field, E., Pande, R., Papp, J., & Rigol, N. (2013). *Does the classic microfinance model discourage entrepreneurship among the poor? Experimental evidence from India*. The American Economic Review, 103(6), pp. 2196-2226.
- Firebaugh, G. (2008). *The First Rule, from Seven Rules for Social Research*. Introductory Chapters.
- Floro S. L., Pan A. Y. (1991). *Informal credit markets and the new institutional economics: the case of Philippine agriculture*. Westview press.
- Goldscheider, F., Bernhardt, E., & Lappegård, T. (2015). *The gender revolution: A framework for understanding changing family and demographic behavior*. Population and Development Review, 41(2), pp. 207-239.
- Hill, R. C., Griffiths, W. E., & Lim, G. C. (2008). *Principles of econometrics*. Fourth edition. Hoboken, NJ: Wiley.
- Hodgson, G. (2004). 'Veblen and Darwinism', International Review of Sociology, Vol. 14(3), pp. 343-361.
- Hulme, D. (1997). *Impact assessment methodologies for microfinance: A review*. AIMS, USAID.
- Hulme, D. (2000). *Impact assessment methodologies for microfinance: theory, experience and better practice*. World development, 28(1), pp. 79-98.
- Karlan, D., & Valdivia, M. (2011). *Teaching entrepreneurship: Impact of business training on microfinance clients and institutions*. Review of Economics and statistics, 93(2), pp. 510-527.
- Karlan, D., & Zinman, J. (2011). *Microcredit in theory and practice: Using randomized credit scoring for impact evaluation*. Science, 332(6035), pp. 1278-1284.
- Khan, R. E. A., & Noreen, S. (2012). *Microfinance and women empowerment: A case study of District Bahawalpur (Pakistan)*. African Journal of Business Management, 6(12), 4514.
- Kohler, H. P., Billari, F. C., & Ortega, J. A. (2002). *The emergence of lowest-low fertility in Europe during the 1990s*. Population and development review, 28(4), pp. 641-680.
- Koning, A., & Wardle, L. (2014). *Embedding Social Performance Management in Financial Service Delivery*. The World Bank (No. 20261).
- Krumm, D. (2010). *Microfinance Plus programmes*. In UMM Workshop Report (p. 7).
- Latouche, S. (2010). *L'invenzione dell'economia*. Bollati Boringhieri.

- Lawson T. (2001). *Two responses to the failings of modern economics: the Instrumentalist and the Realist*. Review of Population and Social Policy, 10: pp. 155–81.
- Lesthaeghe, R. (2010). *The unfolding story of the second demographic transition*. Population and development review, 36(2), pp. 211-251.
- Massey, D. S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., & Taylor, J. E. (1993). *Theories of international migration: A review and appraisal*. Population and development review, 19(3), pp. 431-466.
- Mauri A., Conti C. (2000). *Finanza informale, finanza etica e finanza internazionale nelle piccole e medie imprese*. Giuffrè Editore.
- McCloskey, D. N. (1983). *The rhetoric of economics*. Journal of Economic Literature, 21(2), pp. 481-517.
- McCloskey, D. N. (1985). *The loss function has been mislaid: The rhetoric of significance tests*. The American Economic Review, 75(2), pp. 201-205.
- Mersland, R., & Strøm, R. Ø. (2010). *Microfinance mission drift?*. World Development, 38(1), pp. 28-36.
- Mokyr J. (2005), "Is there a theory of economic history?", in K. Dopfer (ed.), *The Evolutionary Foundations of Economics*, Cambridge: Cambridge University Press, pp. 194-218.
- Morduch, J. (2000). *The microfinance schism*. World development, 28(4), pp. 617-629.
- North D.C. (1968). 'Economic History', in D.L. Sills (ed.), *The International Encyclopedia of the Social Sciences*, vol. 6(7), pp. 468-474, -: Macmillan and the Free Press.
- Odell, K. (2010). *Measuring the impact of microfinance*. Grameen Foundation, Washington, pp. 1-38.
- Pizzo, G., & Tagliavini, G. (2013). *Dizionario di microfinanza: le voci del microcredito*. Carocci.
- Presbitero, A. F. (2009). *La crisi 2007-?: Fatti, ragioni e possibili conseguenze*. Money and Finance Research Group.
- Pronyk, P. M., Kim, J. C., Abramsky, T., Phetla, G., Hargreaves, J. R., Morison, L. A., & Porter, J. D. (2008). *A combined microfinance and training intervention can reduce HIV risk behaviour in young female participants*. Aids, 22(13), pp. 1659-1665.
- Rizzi F. (12/05/17). *Europa e migrazioni nel Mediterraneo: guerre e crisi o nuova lotta di classe?*. Il Fatto Quotidiano. Retrieved: <http://www.ilfattoquotidiano.it/2015/05/12/europa-e-migrazioni-nel-mediterraneo-guerre-e-crisi-o-nuova-lotta-di-classe/1672834/> [Accessed 14 May 2017]

- Robeyns, I. (2003). *Sen's capability approach and gender inequality: selecting relevant capabilities.* Feminist economics, 9(2-3), pp. 61-92.
- Robeyns, I. (2006). *The capability approach in practice.* Journal of Political Philosophy, 14(3), pp. 351-376.
- Rocci, L. (1998). *Vocabolario greco-italiano.* Dante Alighieri.
- Rodrik, D. (2006). *Goodbye Washington consensus, hello Washington confusion? A review of the World Bank's economic growth in the 1990s: learning from a decade of reform.* Journal of Economic literature, 44(4), pp. 973-987.
- Rossi, P. H., Wright, J. D., & Anderson, A. B. (2013). *Handbook of survey research.* Academic Press.
- Screpanti, E., Zamagni, S. (2004). *Profilo di storia del pensiero economico: Dalle origini a Keynes.* Carocci.
- Screpanti, E., Zamagni, S. (2005). *An outline of the history of economic thought.* Oxford University Press on Demand.
- Sen, A. (1994) "Population and Reasoned Agency: Food, fertility, and economic development", in Kiessling- Lindahl, K. & Landberg, H. (eds.) Population, Economic Development and the Environment: The making of our common future. Oxford: OUP, 51-78.
- Sinha, F. (2006). *Social Rating and Social Performance Reporting in Microfinance.* Towards a Common Framework, the SEEP Network for the Argidius Foundation.
- SPTF – Social Performance Task Force (2016). *Universal Standards for Social Performance Management. Version 2.0.* Retreived: <https://sptf.info/images/USSPM-Manual-English-v2.0-Nov2016.pdf>
- Ssewamala, F. M., Ismayilova, L., McKay, M., Sperber, E., Bannon, W., & Alicea, S. (2010). *Gender and the effects of an economic empowerment program on attitudes toward sexual risk-taking among AIDS-orphaned adolescent youth in Uganda.* Journal of Adolescent Health, 46(4), pp. 372-378.
- Stigler, G. J., Stigler, S. M., & Friedland, C. (1995). *The journals of economics.* Journal of Political economy, 103(2), pp. 331-359.
- Storey, D.J., Greene, F.J. (2010) *Small Business and Entrepreneurship.* Harlow, Prentice Hall.
- Strickland A. (23/02/2017). *Astronomers discover 7 Earth-sized planets orbiting nearby star.* CNN. Retreived: [www.edition.cnn.com/2017/02/22/world/new-exoplanets-discovery-nasa](http://www.edition.cnn.com/2017/02/22/world/new-exoplanets-discovery-nasa) [Accessed 27 February 2017]

- Sutter, D., & Pjesky, R. (2007). *Where would Adam Smith publish today? The near absence of math-free research in top journals.* Econ Journal Watch, 4(2), pp. 230-240.
- Van Rooyen, C., Stewart, R., & De Wet, T. (2012). *The impact of microfinance in sub-Saharan Africa: a systematic review of the evidence.* World Development, 40(11), pp. 2249-2262.
- Verschuren, P. J. (2001). *Holism versus reductionism in modern social science research.* Quality & Quantity, 35(4), pp. 389-405.
- Wardle L. (2014). *The Universal Standards for Social Performance Management - Implementation Guide.* Social Performance Task Force (SPTF). Retrieved: [https://sptf.info/images/usspm%20impl%20guide\\_english\\_20141217.pdf](https://sptf.info/images/usspm%20impl%20guide_english_20141217.pdf) [Accessed 13 April 2017]
- Weintraub, E. R. (2002). *How economics became a mathematical science.* Duke University Press.
- Wooldridge, J. M. (2013). *Introductory econometrics: A modern approach.* Nelson Education.
- Ziliak, S. T., & McCloskey, D. N. (2004). *Significance Redux.* The Journal of Socio-Economics, 33(5), pp. 665-675.

### ***Interview***

Skype interview with Giampietro Pizzo, 09/02/2017.

### ***Courses***

ISPI School - Istituto per gli Studi di Politica Internazionale. “International cooperation for development” Intensive Course, 31/01/2014 and 01/02/2014.

### ***Website list***

Archivio di Stato di Bologna, Ministero dei beni delle Attività culturali e del turismo. [www.archiviodistatobologna.it/it/bologna/attivit%C3%A0/mostre-eventi/laudace-furto-al-sacro-monte-ovvero-lamaro-caso-conte-lucchini/2](http://www.archiviodistatobologna.it/it/bologna/attivit%C3%A0/mostre-eventi/laudace-furto-al-sacro-monte-ovvero-lamaro-caso-conte-lucchini/2)

AVERT. [www.avert.org](http://www.avert.org)

CERISE. [www.cerise-spi4.org](http://www.cerise-spi4.org)

CGAP platform – Consultative Group to Assist the Poor (World Bank). [www.cgap.org](http://www.cgap.org)

CGAP – Portal of microfinance. [www.microfinancegateway.org](http://www.microfinancegateway.org)

CNN. [www.edition.cnn.com](http://www.edition.cnn.com)

European Microfinance Network. [www.european-microfinance.org](http://www.european-microfinance.org)

Imp-Act Consortium. [www.ids.ac.uk/project/imp-act](http://www.ids.ac.uk/project/imp-act)

IMPACT FINANCE Management. [www.impact-finance.com](http://www.impact-finance.com)

ISPI School – Istituto per gli Studi di Politica Internazionale. [www.ispionline.it](http://www.ispionline.it)

KARMAX – Impact Monitoring System. [www.kharmax-impact.com](http://www.kharmax-impact.com)

Malaria Consortium. [www.malariaconsortium.org](http://www.malariaconsortium.org)

M-CRIL - Inclusive Microeconomics. [www.m-cril.com](http://www.m-cril.com)

Microfinanza. [www.microfinanza.com](http://www.microfinanza.com)

MicroFinanza Rating. [www.microfinanzarating.com](http://www.microfinanzarating.com)

MicroFinance Transparency. [www.mftransparency.org](http://www.mftransparency.org)

MicroRate. [www.microrate.com](http://www.microrate.com)

MIX Platform – Microfinance Information Exchange. [www.themix.org](http://www.themix.org)

Planet Rating. [www.planetrating.com](http://www.planetrating.com)

Progress out of Poverty. [www.progressoutofpoverty.org](http://www.progressoutofpoverty.org)

RITMI – Rete ITaliana di Microfinanza. [www.microfinanza-italia.org](http://www.microfinanza-italia.org)

SPMTF platform – Social Performance Management Task Force. [www.sptf.info](http://www.sptf.info)

Smart Campaign. [www.smartcampaign.org](http://www.smartcampaign.org)

UNAIDS – United Nations Programme in HIV and AIDS. [www.unaids.org](http://www.unaids.org)

WHO – World Health Organization. [www.who.int](http://www.who.int)



## Appendix

### **PART A – Cross referencing**

The following table contains: the 15 papers reviewed in section 3.1, the 3 academic reviews on microfinance IE (Bauchet et al. 2011; Duvendack et al. 2011; Van Rooyen et al. 2012) from which some publications were retrieved and the Armendáriz et al. (2010) handbook. Each cell contains a symbol: **B** indicates the publication appears in the bibliography; **R** indicates the paper has been specifically reviewed (note that being reviewed implies also being in the reference list); **X** pinpoints that the paper is neither in the review or the bibliography. The grey shaded cells with a small **X** indicate that a paper could not be included in the reference list because of the publication date, whereas the diagonal is represented by the grey shaded cells with a big **X**. Finally the **\*** symbol specifies that a previous version of the paper appears either in the review or bibliography. The final row accounts for number of times a paper (identified with a number) is cited in the other publications, while the final column indicates how many of the publications included in the table are at least included in each paper's bibliography.

**Table A. 1:** Cross reference overview

PUBLICATIONS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	TOT
1. Angelucci et al. (2012)	X		X	B	B*	B	B	X	X	B*	X	B	B	X	X	X	X	X	7	

2. Ashraf et al. (2009)	X		X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	0
3. Attanasio et al. (2011)	X	X		X	B*	B	B	X	X	B*	X	X	B	X	X	B*	X	X	X	6
4. Augsburg et al. (2015)	X	X	X		B	B	X	X	X	X	X	X	B	X	X	X	X	X	X	3
5. Banerjee et al. (2013)	R	X	R	R*		R	B	X	X	X	X	X	B	X	X	B	X	X	X	7
6. Crépon et al. (2011)	X	X	X	X	B*		X	X	X	X	X	X	B*	X	X	B	X	X	X	3
7. De Mel et al. (2009)	X	X	X	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X	0
8. Dupas et al. (2013a)	X	X	B	X	B	B	B		B	X	X	X	B*	X	X	X	X	X	X	6
9. Dupas et al. (2013b)	X	X	X	X	X	X	X	B		X	X	X	X	X	X	X	X	X	X	1
10. Fafchamps et al. (2014)	X	X	X	X	X	X	B	X	X		X	X	X	X	X	X	X	X	X	1
11. Field et al. (2013)	X	X	X	X	B	X	B*	B	X	X		X	B	X	X	B*	X	X	X	5
12. Karlan & Valdivia (2011)	X	X	X	X		X	B	X	X	X	X		X	X	X	X	X	X	X	1
13. Karlan & Zinman (2011)	X	X	X	X	B*	X	X	X	X	X	X		X	X	B	X	X	X	X	2
14. Pronyk et al. (2008)	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X	X	X	X	0
15. Ssewamala et al. (2010)	X	X	X	X	X	X	X	B*	X	X	X	X	X	X		X	X	X	X	1
16. Armendáriz et al. (2010)	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>B*</b>	<b>X</b>	<b>B</b>	<b>B*</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>B*</b>	<b>X</b>	<b>X</b>		<b>X</b>	<b>X</b>	<b>X</b>	<b>4</b>

17. Bauchet et al. (2011)	X	X	X	X	R	R	R	R*	X	R*	R*	R	R	X	X	B*	X	X	X	9
18. Duvendack et al. (2011)	X	X	X	X	R*	X	B	B*	X	X	X	X	R*	X	X	B	X	X	X	5
19. Van Rooyen et al. (2012)	X	R	X	X	B*	X	X	R*	X	X	X	X	R	R	B*	X	B	X	X	7
TOT	1	1	3	2	11	6	9	7	1	3	1	2	10	1	1	8	0	1	0	68



**PART B – SPM Universal Standards**

The following tables briefly sketch the standards and essential practices encompassed in D2, D4, D5 and D6 of the SPMUS.

**Table B. 1:** Dimension 2

D2 - Ensure board, management and employee commitment to social goals	STANDARD	ESSENTIAL PRACTICE
	Members of the Board of Directors (BoD) hold the institution accountable to its mission and social goals	<ul style="list-style-type: none"> <li>&gt; BoD is provided with an orientation on the social mission, goals and their responsibilities related to the SPM</li> <li>&gt; BoD reviews SP data: mission, compliances, human resources, SP risks, client protection, growth &amp; profit allocation</li> <li>&gt; BoD uses SP data to provide strategic direction, considering both social &amp; financial goals</li> <li>&gt; BoD uses SPM criteria to evaluate CEO performance</li> <li>&gt; BoD has a strategy to prevent mission drift during changes in ownership structure and/or legal form.</li> </ul>
	Senior management (SM) oversees implementation strategies for achieving social goals	<ul style="list-style-type: none"> <li>&gt; SM integrates the SP into business planning</li> <li>&gt; The FP raises awareness and concern about fair and responsible treatment of clients</li> <li>&gt; SM analyzes and addresses SP related risks</li> <li>&gt; SM analyzes SP data to compare actual performance with stated social targets</li> <li>&gt; The CEO holds SM accountable for making progress toward FP's social goals</li> </ul>
	Employees are recruited, evaluated, and recognized based on both their social & financial performance	<ul style="list-style-type: none"> <li>&gt; Candidates are screened and hired for their commitment to social goals and their ability to carry out SP related responsibilities</li> <li>&gt; Employees are trained &amp; evaluated on both their SP &amp; financial performance responsibilities</li> <li>&gt; FP implements policies to promote ethics &amp; prevent fraud</li> <li>&gt; The institution incentivizes quality portfolio</li> </ul>

**Table B. 2:** Dimension 4

D4 - Treat Clients Responsibly	STANDARD	ESSENTIAL PRACTICE
	Prevention of over-indebtedness	<ul style="list-style-type: none"> <li>&gt; FP conducts appropriate client repayment capacity analysis before disbursing a loan</li> <li>&gt; FP uses credit bureau and competitor data for client assessment (when feasible)</li> <li>&gt; SM &amp; BoD are aware of and concerned about the risk of over-indebtedness</li> <li>&gt; FP's internal audit department monitors that policies to prevent over-indebtedness are applied</li> <li>&gt; FP avoids dangerous commercial practices</li> </ul>
	Transparency: Clients should be able to make informed decisions	<ul style="list-style-type: none"> <li>&gt; FP fully discloses cost and non-cost information about products, services and contract terms.</li> <li>&gt; FP communicates proactively with clients in a way that clients can easily understand</li> <li>&gt; FP uses a variety of disclosure channels (verbal &amp; written communication, radio, web, television, etc.)</li> <li>&gt; FP dedicates adequate time for client review &amp;</li> </ul>

		<p>discloses information frequently  &gt; FP provides accurate and timely information</p>
	Fair and Respectful Treatment of Clients	<p>&gt; FP has defined in specific detail appropriate debt collection practices  &gt; FP's human resource policies on recruitment &amp; training are aligned around fair and responsible treatment of clients  &gt; In selection &amp; treatment of clients, FP does not discriminate against some categories of customers  &gt; Collections staff are expected to follow the same practices as the FP's employees  &gt; FP informs clients of their rights</p>
	Privacy of Client Data	<p>&gt; FP has a privacy policy &amp; appropriate technology systems  &gt; FP informs clients about when and how their data is used &amp; gets their consent</p>
	Mechanisms for Complaint Resolution	<p>&gt; FP's clients are aware of how to submit complaints  &gt; FP's employees are trained to handle complaints  &gt; FP's complaints &amp; resolution system is active &amp; effective  &gt; FP uses client feedback to improve practices &amp; products</p>

**Table B. 3:** Dimension 5

<b>D5 - Treat employees responsibly</b>	<b>STANDARD</b>	<b>ESSENTIAL PRACTICE</b>
	The FP follows a written Human Resources (HR) policy that protects employees and creates a supportive working environment	<p>&gt; A written HR policy is available, complying with national law &amp; explains employees' rights  &gt; Staff compensations constitute a living wage  &gt; FP accepts &amp; responds to employee complaints through a formal and confidential grievance system  &gt; FP neither employs nor benefits from forced or compulsory labor or illegal child labor  &gt; FP assesses employees' health and safety risks and takes steps to mitigate them  &gt; FP documents, reports, and investigates all occupational accidents, injuries or diseases</p>
	The FP communicates to all employees the terms of their employment and provides training for essential job functions.	<p>&gt; Each employee receives a written job description &amp; an employment contract.  &gt; Employees receive job-specific training and/or skill development.  &gt; Each employee knows how his/her performance will be evaluated &amp; rewarded by the FP</p>
	The FP monitors employee satisfaction and turnover	<p>&gt; FP analyzes employee satisfaction  &gt; FP monitors &amp; understands employee turnover  &gt; FP takes action to correct problems leading to employee turnover and dissatisfaction.</p>

**Table B. 4:** Dimension 6

<b>D6 - Balance Financial and Social Performance</b>	<b>STANDARD</b>	<b>ESSENTIAL PRACTICE</b>
	The FP sets and monitors growth rates (GR) that promote both financial sustainability and client well-being	<ul style="list-style-type: none"> <li>&gt; FP establishes a policy on sustainable target GR, approved by BoD, considering institution's growth capacity &amp; targeted markets</li> <li>&gt; FP analyses GR &amp; market saturation to assess whether growth policies ensure both financial sustainability &amp; client well-being.</li> <li>&gt; FP monitors whether its internal capacity is keeping pace with growth &amp; enhances it if required</li> </ul>
	Equity investors, lenders, board, and management are aligned on the FP's double bottom line and implement an appropriate financial structure in its mix of sources, terms, and desired returns	<ul style="list-style-type: none"> <li>&gt; FP has clear strategies, consistent with its social goals, its desired level of returns &amp; its profit allocation policies</li> <li>&gt; FP engages with funders whose expectations for financial returns, timeframe &amp; exit strategies are aligned with its mission &amp; stage of development</li> <li>&gt; When deciding on funding sources, FP considers how cost of capital is passed on to the client</li> <li>&gt; FP minimizes financial risk as it relates to its obligations to clients</li> <li>&gt; FP has a transparent financial structure, as reflected in its annual audited financial statements that incorporate any off-balance sheet sources of funding into leverage ratios</li> </ul>
	Pursuit of profits does not undermine the long-term sustainability of the FP or client well-being	<ul style="list-style-type: none"> <li>&gt; Products and services offer value to the client for the price</li> <li>&gt; FP offers market-based, non-discriminatory pricing</li> <li>&gt; FP has efficiency ratios aligned with competitors</li> <li>&gt; FP does not charge excessive fees</li> <li>&gt; BoD monitors whether FP's pricing levels are consistent with its own policies on returns</li> <li>&gt; FP establishes a field-officer-to-client ratio that promotes high service quality for clients</li> </ul>
	The FP offers compensation to SM that is appropriate to a double bottom line institution.	<ul style="list-style-type: none"> <li>&gt; BoD ensures that compensation of the CEO &amp; other senior staff is in line with FP's social goals</li> <li>&gt; If SM compensation is in part incentive-based, the incentives are based on the SP evaluation.</li> <li>&gt; Upon request, FP transparently discloses compensation to regulators, auditors, raters, donors, lenders, and investors</li> <li>&gt; FP calculates the difference between the average compensation of its top level executives and its field employees, &amp; analyzes if this spread is consistent with its mission</li> </ul>