



City Research Online

City St George's, University of London

Citation: Parmar, D. & De Allegri, M. (2014). Operationalizing impact evaluations: from theory to practice. In: Radermacher, R. & Roth, K. (Eds.), A Practical Guide to Impact Assessments in Microinsurance. (pp. 243-266). Luxembourg: Microinsurance Network and Micro Insurance Academy. ISBN 978-99959-864-8-3

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/4594/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

A Practical Guide to Impact Assessments in Microinsurance

Edited by Ralf Radermacher and Katja Roth



A Practical Guide to Impact Assessments in Microinsurance

Edited by Ralf Radermacher and Katja Roth



This publication aims to guide academic scholars and microinsurance practitioners, donors and policy makers in designing and carrying out high-quality impact evaluations in microinsurance. Since microinsurance is a fast-growing sector providing innovative tools for risk protection of low-income people, impact evaluations are key to gaining more insight into the effects of microinsurance and the underlying causal relationships which can help optimise microinsurance schemes in future, improving their success, and ultimately contribute to the welfare of the poor.



micro
insurance
network



This book aims at closing a gap: Not only does it show what needs to be considered for rigorous impact evaluations in microinsurance, but it discusses the specific strengths, weaknesses, objectives and limitations of different designs to attain this purpose, including randomised controlled trials, non-experimental designs, qualitative and participatory designs and mixed methods. Moreover, the book addresses the challenges of turning study designs into practice and how to best draw conclusions from given results, create reports and disseminate findings. Last but not least, the book proposes a set of core impact and outcome indicators which can provide a standard framework for all impact evaluations in microinsurance.

If impact evaluations in microinsurance were dishes, this book would be the cookbook!

**A copy of this book can be requested by writing to
info@microinsurancenetwork.org.**

This book is a joint publication of the Microinsurance Network and the Micro Insurance Academy. It is a project of the Microinsurance Network's Impact Working Group. The publication is supported by the International Initiative for Impact Evaluation (3ie), the International Labour Office's (ILO) Microinsurance Innovation Facility and the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH on behalf of the Federal Ministry for Economic Cooperation and Development (BMZ) of Germany.

**More about the Microinsurance Network at
www.microinsurancenetwork.org
and the Micro Insurance Academy at
www.microinsuranceacademy.org.**

© 2014 by Microinsurance Network. All rights reserved.

Picture on left: M. Moniruzzaman/talkativepictures.com

ISBN 978-99959-864-8-3

Acknowledgements

This book was prepared on behalf of the Impact Working Group (IWG) of the Microinsurance Network. We are extremely grateful to the many contributors for their help and support throughout the process.

Published by the Microinsurance Network and the **Micro Insurance Academy (MIA)**, we thank Véronique Faber, Executive Director of the Microinsurance Network, for her ongoing and energetic support during the development of this book. She gave invaluable advice, intelligent suggestions, creative support, and, most importantly, waited with great patience while this book evolved.

Special thanks to Annalisa Bianchessi, from the Microinsurance Network's secretariat, who coordinated the layout and production of the book and shared her genuine interest in the project. Interns, Tiffani Saverimuthu and Olivia Caverly, from the MIA and the Microinsurance Network respectively, were also a great pleasure to work with and provided critical administrative support. Heidi McGowan was deeply involved in the planning and editing process from the beginning. Our work benefited greatly from the foundation that she laid.

In addition we particularly thank Yvonne Deblon and Thomas Wiechers (GIZ), Michal Matul (ILO's Microinsurance Innovation Facility) and Craig Churchill (ILO's Microinsurance Innovation Facility & Microinsurance Network), and Howard White (3ie), who gave us honest feedback and thoughtful advice.

Financial support came from our partners the German Federal Ministry for Economic Cooperation and Development (BMZ), the International Labour Organization (ILO), the International Initiative for Impact Evaluation (3ie) and the Luxembourg Government. This book is the product of their support.

More than 20 experts from the microinsurance arena have contributed to this book by (co-)authoring chapters. Their rich experiences are reflected in their contributions and make this book such a useful resource for all interested in impact assessment and microinsurance. Furthermore, 15 microinsurance experts functioned as reviewers for the chapters. Their critical view and constructive feedback has contributed a lot to shaping this book. Our thanks to all of them!

One of our objectives was the proposition of core impacts, outcomes and appropriate indicators for microinsurance impact assessments. As will be explained in the course of this book, in order to identify those core effects and indicators, we have conducted a Delphi process in which almost 30 experts took part (although there was some attrition from round to round). We thank all of them for their participation and most valuable input.

A lot of other people have contributed to this book in many ways. All of the collaborators on the book, whether named or anonymous, spent valuable personal time and energy to share their expertise. We are deeply appreciative.

Enjoy!

Ralf Radermacher and Katja Roth

Table of content

Foreword

Craig Churchill

9

I. Introduction to Impact Assessments in Microinsurance

13

1. In a nutshell: impact assessments in microinsurance

15

Ralf Radermacher and Katja Roth

2. What is microinsurance?

23

David M. Dror and David Piesse

3. Why and how does impact happen?

41

Simon Quinn, Stefan Dercon, and Michal Matul

4. What can we learn from impact assessments?

61

Jonathan Bauchet, Aparna Dalal, and Jonathan Morduch

II. Approaches to Impact Assessments in Microinsurance

83

5. Experimental designs

85

Tomoko Harigaya and Dean Karlan

6. Non-experimental methodologies for quantitative analysis

111

Markus Frölich, Andreas Landmann, Markus Olapade, and Robert Poppe

7. Qualitative designs

131

Tara Sinha, Christina May, and Jeremy Holland

8. The case of mixed methods for impact evaluation in microinsurance

173

Howard White

III. Planning and Conducting Impact Assessments in Microinsurance

185

9. Defining indicators for impact assessments

187

Conor Doyle and Pradeep Panda

10. Core outcomes, impacts, and indicators for microinsurance

219

Katja Roth, Ralf Radermacher, and Pradeep Panda

11. Operationalising impact evaluation: from theory to practice

243

Divya Parmar and Manuela De Allegri

IV. Results of Impact Assessment Studies

269

12. Drawing conclusions

271

Karljin Morsink and Peter Geurts

13. Reporting and disseminating findings

293

Heidi McGowan

14. Systematic reviews

317

Howard White

Publishers and Partners

329

Foreword

Microinsurance is a rapidly evolving development strategy to extend efficient risk management solutions to low-income households and small businesses. As noted by Amartya Sen, the Nobel Prize-winning economist, crisis has a “class-dependent character”. Lower socioeconomic classes, many of whom work in the informal economy, are more vulnerable to risks than others, and yet they are the least able to cope when crises occur. Microinsurance holds the promise of breaking this perpetuating cycle of vulnerability and poverty whilst making important development contributions, not just at the household level, but also within the community and across the country.

At the household level, microinsurance can potentially help to break the cycle of poverty through both protective and productive contributions. On the protective side, insurance can shield policyholders from the financial consequences of various risks. Small, regular premium payments are more affordable than the large expenses that accompany crises. On the productive side, through life insurance policies, the poor can amass savings and build assets. Alternatively, insurance can help facilitate access to productive inputs such as credit, by covering risks that lenders do not want to carry.

There is also the peace-of-mind effect where the working poor may feel less compelled to set aside unproductive contingency funds “under the mattress” if they are insured, and may make investments in higher-risk, higher-return activities.

Beyond the household level, studies have demonstrated a causal link between the development of the insurance industry in general – not specifically microinsurance – and national economic development. This is accomplished, for example, by stimulating entrepreneurship and enabling businesses to operate with less volatility. Since insurers and reinsurers have an incentive to reduce claims, they also contribute to development by promoting risk reduction measures. And by mobilising long-term savings, insurers are an important source of long-term finance that can be invested in initiatives such as infrastructure improvements, as well as acting as a significant stimulator for the development of debt and equity markets.

But in many countries, the insurance industry is not achieving its development potential. The insurance sectors in many developing economies evolved in the second half of the twentieth century, and focused largely on corporate clients, with little effort expended to build the infrastructure required for personal lines. The emergence of

microinsurance provides the insurance industry with an opportunity to build from the bottom up and create a foundation of retail insurance, ultimately making a stronger contribution to the country's general economic development.

The contribution of microinsurance to a community and a country extends beyond its involvement in deepening the insurance industry. As microinsurance lies at the intersection between social protection and financial inclusion, the contribution to social and economic development will be greatest where these forces are well coordinated. For example, public private partnerships seem to be an important way to leverage market expertise to achieve public policy objectives. By fusing social protection with financial inclusion, it is possible to increase the effectiveness of both, enhancing the ability of workers in the informal economy to cope with the costs associated with illness or death of breadwinners, the theft of productive assets, and the destruction wrought by disasters.

All of these advantages of microinsurance should be clearly articulated as potential benefits. In theory, insurance is an efficient way for the poor to manage certain risks, and there is considerable anecdotal evidence to support the theory. Moving from theory into practice, we now need evidence

from rigorous research to ensure that microinsurance actually benefits poor households – and what design features of microinsurance are likely to yield the largest impact to those it protects.

Microinsurance is not easy, and significant innovations are required to overcome the challenges of viably extending valuable insurance coverage to large numbers of low-income households. During the early days of microinsurance, the focus was primarily on understanding how it worked, the operational tricks of the trade, and improving access. Now that we are seeing significant outreach, and perhaps half a billion low-income persons have some coverage, more attention is being paid to assessing whether the poor actually benefit from insurance or not.

This impact question is not just of interest to academics, although a growing number of academics appear to be interested in trying to answer it. The expansion of academic interest in microinsurance can be partly attributed to the powerful public policy implications if insurance is proven to be a cost effective means of reducing the vulnerability of low-income households. Indeed, donors and policy makers are keen to understand impact. If, for example, they are going to provide subsidies, they want to know whether these interventions really benefit

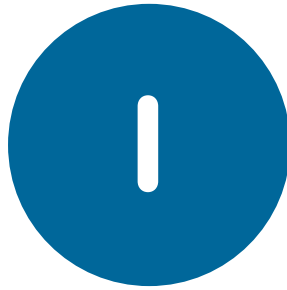
those most in need of protection. Further, they might want to know whether investing in microinsurance is more efficient than investing in other forms of risk protection.

Impact is also of great interest to microinsurance providers. This may seem obvious for microinsurers with a development agenda, like non-governmental organisations (NGOs) and community groups, but even commercially oriented insurers are keen to understand how, and to what extent, their policyholders benefit from their products. The interest of insurance companies to understand impact has two dimensions. Firstly, such research can be built into a process of continuous improvement and enable insurers to identify ways in which they can improve their products, and expand market share, by enhancing the value that their clients derive from insurance. Secondly, many companies take their corporate social responsibility (CSR) obligations seriously and are therefore interested in evidence to validate their efforts. In that same line of thinking, social investors are also keen for impact evidence to justify their investments in microinsurance.

However, it is certainly not easy or straightforward to demonstrate the impact of insurance. It is not possible to look at the social and economic characteristics of policyholders before and after insurance to assess the impact because if there are improvements, for example to their incomes or health status, they could be attributed to many causes. To produce credible results, it is necessary to follow certain research standards and norms.

To advance these efforts, the Microinsurance Network's Impact Working Group has developed these guidelines, sometimes known as the impact cookbook, to provide clear guidance on how to conduct impact studies properly, using both qualitative and quantitative research methods. We expect that these guidelines will lead not only to more valid studies, but also promote common frameworks, facilitating meta-analyses across studies so that we will be better positioned to demonstrate whether microinsurance theory really does translate into practice. I believe this publication will make an indispensable contribution to the campaign to prove and improve the value of insurance for low-income households and small businesses.

Craig Churchill,
ILO's Microinsurance Innovation Facility



Introduction to Impact Assessments in Microinsurance



In a nutshell: impact assessments in microinsurance

Ralf Radermacher and Katja Roth



1.1. Introduction

Everyone is confronted with various risks in life – but the consequences of such risks are often more severe for the poor than for better-off households. When confronted with the financial consequences of calamities like illness, accidents, death, or loss of agricultural or other produce, poor households often have to take actions that harm their productivity and prosperity prospects. These actions could include: borrowing at high interest rates, selling productive assets, taking children out of school or sending them to labour, compromising on a balanced diet, or leaving illnesses untreated. Such responses to risk impede households on their way out of poverty or may even force them deeper into poverty.

Over the past 15 years, microinsurance has increasingly been seen as an alternative and potentially more efficient way for the poor to manage their risks. Stated simply, microinsurance is insurance targeting the poor. It is meant to be offered at an affordable price, and so comes with some limitations in risk coverage. Microinsurance can be offered for all kinds of risks and by all kinds of providers. Often, communities or non-government organisations (NGOs) arrange these programmes themselves, but commercial insurers are increasingly engaging in the provision of microinsurance (see chapter 2

“What is microinsurance?”). They recognise that the largest untapped insurance market in terms of clients is amongst the poor, with an estimated market size of four billion people (Swiss Re 2010). Governments can also be providers of microinsurance – or take an interest in the promotion and regulation of the market, either as part of a financial sector development or as part of a social protection approach.

It is essential to learn more about the effects of insurance on the insured and the communities in which they live.

This is particularly true for many governments and funders of development, for whom the provision of insurance is not an end in itself, but rather a tool to achieve certain development objectives like, for example, reduction of poverty, improved access to health care and, through this, fewer maternal and infant deaths.

Understanding impact is key for sustainable market development

The effect microinsurance has on the lives of the poor or the larger society is what we call its *impact*. These changes can be positive or negative, direct or indirect, and intended or unintended. Furthermore, they can happen before (i.e., ex-ante) or after (i.e., ex-post) the occurrence of insured events and

Impact: end in itself, results take longer
 Outcome: bridge to impact, immediate results

can affect not only the insured, but also their household members, their communities, or other populations (Radermacher et al. 2012; OECD 2002). Hence, microinsurance can change the life of people even if they never make a claim or – even more – the lives of people that are not even insured.

Regarding these effects, one needs to differentiate between impacts and outcomes. Impacts are an end in itself, whilst outcomes are the bridge expected to lead to the desired impacts. Impacts are usually expected to be long-term effects of an intervention, whilst outcomes are normally achievable in a short term (OECD 2002).

Measuring the impact of a single microinsurance scheme can provide information about its *absolute impact* or its *effectiveness*, i.e., whether the insurance makes a difference for those covered and whether these differences meet the objectives one was aiming at with the insurance. It is, however, also of key interest to know whether one approach to microinsurance has a larger or different kind of impact than another approach – or how

microinsurance compares to other risk management techniques. Such comparative information, called *relative impact* or *efficiency*, can be immensely helpful for the design of future microinsurance programmes and investment decisions of donors and governments. Furthermore, collecting information on the costs of the implementation can provide insights into the cost-efficiency of programmes with similar goals.

However, to use the information to compare cost-efficiency between different schemes or microinsurance and other risk management approaches, it is important that the programmes are at a similar stage of implementation. On one hand, pilot projects may be comparatively more expensive than larger roll-outs, as economies of scale and learning effects are likely to reduce costs per insured. On the other hand, the quality of implementation is usually higher in the pilot, potentially leading to higher impacts per insured compared to programmes taken to scale. Generally, effects of scale and scope need to be considered when comparing programmes with regard to their efficiency.

Check for different effects
on different subgroups
of the sample

When examining the effects of microinsurance, it is useful to understand the effect on the insured. Sometimes information on the whole community offered the insurance, no matter whether they chose to enrol or not, is also helpful. Moreover, effects can differ for different subgroups of a sample, such as women versus men, children versus adults, or extremely poor versus better-off households. Understanding these differences can provide insight on equity and equality effects, which might also be the desired effects of an insurance programme. Donor organisations might particularly aim to improve the well-being of specific vulnerable subgroups. Hence, their prosperity compared to other subgroups might be in their focus. For example, it might be of interest to know how a health microinsurance programme helps to improve the health of girls compared to that of boys, since boys often get better access to care.

Understanding the impact of microinsurance is useful and necessary. The big question, however, is *how* to measure this impact.

1.2. Measuring impact

Microinsurance can have various effects on its clients and their community. Defining the aspects of interest is the first step in measuring impact. A core set of topics will be derived from the initial

objectives for which a microinsurance scheme was set up. This set of topics should also include potential effects that are unintended or negative. A causal chain is established to describe how the insurance is expected to unfold the impact in focus. Such a causal chain is called *theory of change* and is described in more detail in chapter 3 “Why and how does impact happen?” Formulating such a theory should be based on a detailed understanding of the context and functioning of the insurance scheme. Not only do potential – intended and unintended – outcomes and impacts need to be included, but also the inputs and outputs of the scheme, specifically on aspects of insurance processes and the insurance product, as these are likely to influence the kind and magnitude of the impact seen.

Take, for example, a health microinsurance scheme aimed at reducing maternal and infant deaths by promoting institutional deliveries. A theory of change for this scheme should take into account as its *inputs*: all scheme design decisions with respect to benefits covered, educational activities, and how the scheme is delivered (and to whom). The *output* of this is then the particular insurance scheme with a benefit

Product and processes influence impact

package taken up by a certain number of women in a particular region. It is assumed that women understand the scheme. As direct financial barriers for accessing these services are removed or lowered by the scheme, the utilisation of these services is expected to go up (outcome), which assumes that health services are available and service providers work with the insurance agency as required. This higher utilisation is expected to lead to a reduction in maternal and infant deaths (impact), which assumes the health services are provided in a timely manner and are of sufficient quality. **Differences in the input, i.e., scheme design and execution, might lead to different output, outcome, and impact.** For example, including the women's spouses into an information campaign on the importance of obstetric care might lead to much better outputs, outcomes, and impacts as compared to targeting the information campaign on the women alone. **Processes thus matter a lot for understanding impact and must be considered in a good impact assessment.**

Some outputs and outcomes measurable on the day-to-day level are captured in social performance indicators (Sandmark 2013), which

complement the financially oriented key performance indicators (Wipf and Garand 2010). With their detailed and comparable information on scheme design and results, such indicators can be very helpful in impact research design and interpretation of results. Particularly in comparative studies they can provide a set of structured criteria for comparing observed differences in outcomes and impact.

When it has been decided what to measure, a careful research design needs to ensure that the impact of insurance is actually captured, separating it from impact of other interventions or external events. This cannot be done with simple before and after comparisons. Instead, the situation (of the insured households, the communities, etc.) *with* microinsurance needs to be compared to the counterfactual, i.e., how the situation would be for the same persons (or households, communities, etc.) *without* microinsurance. Since the same household cannot be simultaneously insured and uninsured, a comparison group, as similar to the treatment group as possible, needs to be established. Chapter 4 "What can we learn from impact assessments?" elaborates these research design aspects in more depth.

Several approaches are possible for obtaining a reliable comparison group. Ideally, impact evaluations of programmes should be planned before the programme implementation starts. This can allow for the use of randomised controlled trials, described in detail in chapter 5 “Experimental designs”. In such a design, the insurance is offered to a randomly selected set of households or individuals and not offered to a control group of comparative nature. Based on the assumption that observable and unobservable differences between individuals in these groups are equally distributed, the effects observed after the insurance implementation can be attributed to the insurance itself. However, such a randomised trial, considered the most robust evaluation approach, requires not only planning the evaluation before the rollout of the insurance, but also requires close cooperation between evaluator and implementer – and often more management and coordination effort for the implementer. This might not be possible in every case, and frequently evaluations are also commissioned when a project is already running for several years. Hence, a randomised trial is not possible in all situations, or may not be the solution of choice. Other statistical methods can also be used to enhance comparability of data on the insured and uninsured. Chapter 6 “Non-experimental design methodologies for quantitative analysis” explains these approaches.

In their core nature, the techniques described above are mainly built on quantitative data. Such data are crucial for making any statement about the magnitude of effects. However, quantitative approaches alone frequently fall short of explaining some of the phenomena observed and confirming the causal chain described in the theory of change or in exploring aspects which might have escaped the attention of the researchers. Qualitative research approaches, described in chapter 7 “Qualitative designs”, can fill these gaps and contribute to answering questions on *why* (explanatory) and *what* (explorative). An ideal impact evaluation would combine both quantitative and qualitative techniques and is thus referred to as “mixed methods” (see chapter 8 “The case of mixed methods for impact evaluation in microinsurance”).

With the research focus and method decided, the evaluation can be further operationalised. A key question here is how to measure the research aspects of interest, i.e., defining indicators that adequately describe the aspects of interest. All indicators need to be precisely described and understandable to respondents to ensure that the collected information can be useful. Chapter 9 “Defining indicators for impact assessments” explains aspects linked to the choice of indicators and provides examples of indicators commonly

used. As the number of evaluations of microinsurance schemes is still limited, the indicators used to date may not necessarily correspond to the main areas of interest in the microinsurance space. Yet, including these indicators in the evaluation makes it easier to compare studies. Agreeing on core indicators for the microinsurance space is thus of key importance, so that future studies can include these indicators and enhance the body of evidence on topics considered of key importance. Chapter 10 “Core outcomes, impacts and indicators for microinsurance” suggests a number of core indicators, based on a Delphi study amongst microinsurance experts.

The selected indicators constitute the basis for the questions to be included in data collection instruments (also chapter 10). Transforming indicators into questions includes careful consideration of the targeted respondents, the precise formulation of the question, and a balancing of what eventually can be included in the research. The design of the research tool and selection of the respondents has direct influence on the planning of field work. For certain respondents, seasonality will affect availability, e.g., due to harvest or temporal migration. Further, social and cultural norms must be taken into account when selecting the research team and planning a smooth entry into the target area.

The amount of training and capacity building will need to be decided to ensure the research tools are properly administered – and we recommend planning for close supervision, enabling the evaluator to react to any problems encountered. Such close involvement and supervision can help increase data quality, which is essential for obtaining reliable answers to the research aspects in focus. Chapter 11 “Operationalising impact evaluation: from theory to practice” develops these aspects further.

*Use of core indicators
can enhance
comparability*

The primary answer provided by any given study is whether the specific insurance scheme under scrutiny has resulted in a certain impact. As chapter 12 “Drawing conclusions” describes, the data needs to be checked for validity and related back to the theory of change. Also, conclusions should include practical recommendations for the specific scheme. However, generalisation of the insights or their transferability to other settings, like other regions or target

groups, is often intended as well. High external validity is required for this. It is also desirable to make the insights of the evaluation available publicly, to allow others to learn from a scheme's results—regardless of whether the results are desirable or undesirable. To allow such learning, information on the context in which the scheme operates and how the evaluation was conducted is essential. A common reporting standard can help in achieving this, as described in chapter 13 “Reporting and disseminating findings”, which also provides an overview about dissemination options.

A growing body of evidence increases the reliability of insights gained by different studies, providing a basis for systematic reviews. When these studies use common indicators and provide a comparable set of background information, **conclusions on what works and why can be derived from an analysis across studies**. Such analyses are called **systematic reviews**, which may include a **statistical synthesis using meta-analyses** (see chapter 14 “Systematic reviews”). Reviews provide answers on specific questions by drawing on the entire body of evidence available. They can be a key tool for providing reliable guidance for the development of the microinsurance sector, but need robust studies as a foundation.

This book is written to support the creation of more and better

quality evidence on the impact of microinsurance – evidence which in turn should then be used to create an even more impactful microinsurance sector.

References

Organisation for Economic Co-operation and Development (OECD). 2002. *Glossary of key terms in evaluation and results based management*. Paris: OECD Publications. <http://www.oecd.org/dac/evaluation/2754804.pdf>.

Radermacher, R., H. McGowan, and S. Dercon. 2012. What is the impact of microinsurance? In: *Protecting the poor: A microinsurance compendium*. Volume II, ed. C. Churchill and M. Matul. 59-82. Geneva: International Labour Organization.

Sandmark, T. 2013. *Social performance indicators for microinsurance – A handbook for microinsurance practitioners*. Luxembourg: Appui au Développement Autonome (ADA). http://microinsurancenetWORK.org/publication/fichier/Social_Performance_Indicators_web_2.pdf.

Swiss Re. 2010. Microinsurance—risk protection for 4 billion people. *Sigma* 6. http://media.swissre.com/documents/sigma6_2010_en.pdf.

Wipf, J. and D. Garand. 2010. *Performance indicators for microinsurance – A handbook for microinsurance practitioners*. 2nd ed. Luxembourg: Appui au Développement Autonome (ADA). http://microinsurancenetWORK.org/publication/fichier/KPI_MI_Handbook_v2_EN.pdf.

Systematic reviews and meta-analyses provide robust evidence across studies

What is microinsurance?

David M. Dror and David Piesse



2.1. Introduction

Latest estimates suggest that there are some 500 million microinsurance clients throughout the developing world. According to Craig Churchill, this could pass the one billion mark by the end of the decade (Microinsurance Network 2013). The reason for this large growth has been the increasing activity of governments, insurance companies, and providers worldwide broadening the geographic scope and range of insurance services available to low-income people.

About 70% of microinsurance schemes are operated in Asia. For historical reasons, schemes have been concentrated in India and West Africa. In West Africa, “mutuelles” (i.e., community-based microinsurance schemes) developed after some governments instituted user fees for health care services during structural adjustment programmes. In India, microinsurance schemes arose after implementing the obligations of insurers to rural and social sectors by the Indian Insurance Regulatory and Development Authority (IRDA) in 2002. But recently, microinsurance has expanded to all the developing countries and also to many affinity sections of developed nations.

A range of products cover a variety of risk including health, life/funeral, disability, agriculture (crop-based weather

index and livestock), property, credit life, and disaster (natural and man-made)¹. The prevalence of each type of product varies to some extent by geography and by available technology. According to local risks and cultural norms, for example, funeral coverage is widespread in South Africa. In terms of the technology and expertise available, for example, weather index coverage require data that is usually collected by meteorological equipment, and in some cases microinsurance projects include a component to install such equipment (e.g., Zambia, Malawi, etc.).

2.2. What is microinsurance?

The definition of microinsurance can be split into its two aspects: Firstly, what constitutes *insurance* and secondly, what is *micro* in microinsurance.

2.2.1. Definition of insurance

Insurance is a concept involving a contract under which an insurer shall pay specific pre-defined compensation when financial damages are caused by pre-defined cost-generating events, in exchange for up-front payments of a premium by the insured. In principle,

¹ There are also developments in the catastrophe landscape where numerous public private partnerships are in place for natural disaster protection. A catastrophe joint venture, “MICRO”, is underway in Haiti for earthquake cover and the Philippines has set up an earthquake insurance pool via the Asian Development Bank (ADB).

the premium should reflect the fair cost of the risk transferred from insured to insurer, and the calculation should be based on the frequency and severity. According to theory, insurance offers a trade-off between an unaffordable (or large) loss, which is uncertain, and an affordable loss, which is certain (the premium). This theorem dates back to Friedman and Savage (1948). The net effect of this trade-off is to “smooth” fluctuations in the income of the insured that are caused by exogenous changes, such as different “states of nature”² rather than by autonomous explanations, such as bad choices of consumption in a given set of supply and demand, or bad behaviour in risky situations. The assumption underlying this smoothing is that the insured gains utility from experiencing two years of average consumption rather than experiencing one year of starvation plus one year of excessive consumption. A common explanation for the utility gain is that excessive consumption does not increase happiness, or what economists call utility, as much as starvation lowers it (Gruber 2007, 317).

In most high-income countries, where the public at large could be required to cover the costs of large-scale uninsured events, governments often

require all persons to be insured, and such mandatory insurance is implemented through deductions either from income at source, or inseparably attached to the most common activities of daily life. Examples include: gainful employment with mandatory insurance covering several risks like health, unemployment, old-age pension, or workman’s compensation; owning or driving a car with mandatory third-party liability insurance; and financing of a house with credit life insurance of the borrower and earthquake insurance of the house. In low-income countries, where, on the one hand, governments rarely provide comprehensive disaster relief, and, on the other hand, are often unable to identify all the population or to apply universal tax collection or mandatory insurance to all, the decision to insure is mostly voluntary and individual. When affiliation to insurance is voluntary and individual, the theory suggests that people who estimate their risk exposure to be higher than average would be more likely to insure (*adverse selection*), whilst those estimating their risk exposure to be lower than average would be less likely to insure. The flip side of this phenomenon is when an insurance company agrees to insure only those individuals that it estimates to be exposed to the risk below average (“cherry picking” or “cream skimming”). Both these phenomena affect the insurance market negatively.

2 A finite set of alternatives that might occur, of which only one actually occurs, e.g., real world outcomes like health vs. sickness, abundant harvest vs. bad harvest, normal rainfall vs. drought, etc.

Moreover, in many low-income countries, the three basic conditions for the creation of an insurance market are absent: solvent demand, relevant supply, and reliable governance ensuring that contracts will be enforced. Due to these factors, regular commercial insurance, as is common in industrial countries, is difficult to establish in low-income countries. Therefore, other approaches are needed.

2.2.2. Defining the micro in microinsurance

There are different approaches about how the term micro in microinsurance can be understood.³ Firstly, micro can be understood as a characteristic of the financial situation of the clientele, i.e., an insurance targeted at low-income (and financially marginalised) people in developing countries. Secondly, micro can be understood as characteristic of the product, i.e., an insurance offering limited benefits for small premiums. Thirdly, micro can be understood as characteristic of the process by which the schemes are created and administered. All three ways of interpreting the term micro lead to different definitions of microinsurance and to different answers about what microinsurance is.⁴

Approaches to defining micro differ

2.2.2.1. Micro as characteristic of the target group

Churchill (2006, 12-13) defines microinsurance as follows:

Microinsurance is the protection of low-income people against specific perils in exchange for regular premium payments proportionate to the likelihood and cost of the risk involved. This definition is essentially the same as one might use for regular insurance except for the clearly prescribed target market: low-income people. However, as is demonstrated in this chapter and throughout this book, those three words make a big difference. How poor do people have to be for their insurance protection to be considered micro? The answer varies by country, but generally microinsurance is for persons ignored by mainstream commercial and social insurance schemes, persons who have not had access to appropriate products. Of particular interest is the provision of cover to persons working in the

³ For comparison and further explanation, see Churchill and McCord (2012, 9-10) who use a similar (although not fully identical) description of ways of defining microinsurance.

⁴ For further distinctions, compare also Ingram and McCord (2011).

informal economy who do not have access to commercial insurance nor social protection benefits provided by employers directly, or by the government through employers.

Churchill's definition clearly draws upon the financial situation of the clientele (low-income people). So does the definition of the insurance industry in the Philippines when it defines microinsurance as

providing the poor access to a basket of insurance products, support and services in pursuit of poverty reduction and to provide holistic insurance protection to the stakeholders of the microfinance industry (Martinez 2012).

Although targeting low-income people is usually considered a core characteristic of microinsurance, in practice, this feature raises operational problems. Measuring the household's income (and defining it as low) is not only complex, but also costly and, moreover, not required per se for establishing a microinsurance scheme, particularly if the scheme is not subsidised (Dror 2014).

2.2.2.2. Micro as characteristic of the product

Since microinsurance products and related services are aimed at meeting the risk protection needs for the

low-income and financially-excluded sector, affordability of the premium payments is a paramount consideration for defining microinsurance. The resulting approved microinsurance products are the solution that meets the needs of the target group. Hence the micro in microinsurance can also be understood as characteristic of the product, i.e., of the premiums and the benefits.

India was the first country to seriously define microinsurance products within its regulatory framework, referring to microinsurance as insurance offerings with claim payments less than Rs 50,000 (IRDA 2005). Although targeted towards low-income (and informal sector) people in India, this definition referenced the product offering, not the targeted sector. Based on small premiums and proportionately small benefits, microinsurance products have emerged in India with low-cost premiums that are underwritten in advance. This approach references regulatory microinsurance definitions as they are expressed in terms of regulatory frameworks or charters. These types of frameworks are growing in number across developing countries.

The International Association of Insurance Supervisors (IAIS) and Microinsurance Network combine aspects of the target group and the regulations associated with the product in their

definition of microinsurance (IAIS and Microinsurance Network 2007, 10):

Microinsurance is insurance that is accessed by low-income population, provided by a variety of different entities, but run in accordance with generally accepted insurance practices (which should include the Insurance Core Principles). Importantly this means that the risk insured under a microinsurance policy is managed based on insurance principles and funded by premiums. The microinsurance activity itself should therefore fall within the purview of the relevant domestic insurance regulator/ supervisor or any other competent body under the national laws of any jurisdiction.

2.2.2.3. *Micro as characteristic of the process*

The micro in microinsurance may also relate to the process of designing, introducing, and administering the insurance schemes. In their paper first introducing the term “microinsurance”, Dror and Jacquier (1999) characterise it as voluntary, group-based, self-help insurance. Consequently, the micro in this definition relates to the locus of decisions. With this definition, the main feature of microinsurance is that the schemes are governed directly to some degree by the insured members,

who are somewhat involved in operating the insurance locally.

Stated differently, if an entire country could be described as the *macro* level of society, and a province or district would be *meso*, then the group, village, or neighbourhood would be the micro level. This does not mean that microinsurance units (MIUs) cannot have a large outreach. Consequently, micro does not imply that it cannot be replicated to very large numbers, but it does imply the application of the principle of subsidiarity.⁵

At their inception, MIUs are typically launched, designed, implemented, and administered by and for groups of people without “access to the resources and financial techniques of commercial insurance” (Vaté and Dror 2002, 126), yet they have access to the target population, and can adapt the insurance business process to prevailing, often intricate, informal methods of risk management and financial intermediation. According to this description, successful microinsurance programmes are structured and managed in several fundamentally different ways

⁵ Subsidiarity is an organising principle whereby matters ought to be handled by the smallest, lowest, or least centralised competent authority. Subsidiarity conveys the idea that a central authority should have a subsidiary function, performing only those tasks which cannot be performed effectively at a more immediate or local level. Subsidiarity is, ideally or in principle, one of the features of federalism, where it asserts the rights of the parts over the whole. The concept is applicable in the fields of government, political science, management, military, and, metaphorically, in the context of microinsurance as well.

from commercial insurance offerings, or from social insurance schemes organised by the government in some countries.

Main requirements for MIUs, as understood in this sense, are that they are simple, affordable, and located close to its members.

Dror (2014, 420) builds upon these characteristics whilst taking into account characteristics of the target group and the product when defining health microinsurance as

insurance contextualized to the WTP [willingness to pay], needs and priorities of people in the informal sector who are excluded from other forms of [...] insurance. The schemes are voluntary, with premiums suited to people with low incomes. Although [...] microinsurance is independent of the size of the insurer, the scope of the risk covered, and the delivery channel, it is essential that the scheme is designed to benefit the insured. For practical intents and purposes, this definition implies a central role for the community in at least the design of the scheme, and possibly its operation and governance.

Although originally established for health microinsurance, this definition

can be easily adapted to other fields of microinsurance.

It may be tempting to argue that microinsurance organisations can achieve better renewal rates than commercial insurers because they are driven by demand, not profit, and are based on the needs of the community. In fairness, the evidence for this is, for the time being, still not sufficiently conclusive in low-income countries. However, with extremely low penetration of for-profit commercial insurance, the prospect of delivering microinsurance as a low-cost and low-value product seems even less promising.

2.2.2.4. What micro is not

Regardless of the definition used, experts agree that micro does not refer to the size of a scheme's membership or the total value of premiums amassed or assets insured. Larger microinsurers, including India's Yeshasvini, have millions of clients and collect millions of dollars in premium payments annually (Yeshasvini 2011).

2.2.2.5. Common characteristics of microinsurance definitions

Although the described definitions are very different in their basic approach to microinsurance, they have important characteristics in common. As Dror (2014) describes, the most common

features of microinsurance definitions include:

- Microinsurance is insurance and applies principles of risk pooling
- Microinsurance is suited for people on low incomes
- Microinsurance targets people in the informal sector
- Microinsurance is independent of the class of risk (life, health, crop, livestock, assets, etc.)

Moreover, in most understandings, coverage is always contributory, i.e., never fully subsidised, and, as will be described in detail below, microinsurance can be delivered by different channels, including community-based schemes, insurance companies, or service providers (Dror 2014).

2.3. Business models for microinsurance

To date, there exist at least four operating business models to deliver microinsurance.⁶

Partner-agent model. Whilst in this model the insurance company, i.e., the *partner*, takes responsibility for designing, pricing, and underwriting of products as well as for the scheme's solvency in the long-term. An intermediary, i.e., the *agent*, takes over certain local

tasks like distribution and marketing, premium collection, and product servicing. These responsibilities are delegated to the agent by the partner, i.e., by the insurance company. In rural settings where it is usually costly and time-intensive to assess potential customers, the role of the agent is often taken over by non-governmental organisations (NGOs) or microfinance institutions (MFIs), which, on one hand, are already in contact with the population and, on the other hand, might have identified a need for microinsurance and are thus willing to cooperate. Acting not only as agents of the insurance agency, but also as counsel of the population, they can help in designing suitable products and pressuring the providers for reasonable prices and services.

Provider-driven model. In this model, the policyholders pay premiums directly to the service provider, which, for example, in health microinsurance

There are different ways to deliver microinsurance

⁶ Compare for this and the following explanations Dror 2014, amongst others.

may be a hospital or certain physicians. They are, in turn, allowed to use the services of this provider according to the conditions that have been agreed upon in the insurance policy for free or with a copayment.

Charitable insurance model. In this model, an external charitable organisation is supplementing the scheme financially and, moreover, takes over basically all responsibilities of the “insurer”. By subsidising the scheme, its long-term sustainability is granted, at least as long as the charitable organisation is supporting this.

Mutual/cooperative insurance model. In this model, the community of members is responsible for all aspects of the scheme, hence taking over the role of the insurer mutually. Herewith, the insured are at the same time the insurer. By this, the needs of the members can better be mirrored in the benefit package. Often, mutual societies are not only cooperating in the field of insurance, but also in other fields of interest, functioning as broader mutual-interest organisations.

In reality, microinsurance schemes are often built as combinations of these models and can also change over time. For example, Yeshasvini Trust in India was originally founded by health-care providers. It was provider-driven, but is currently receiving subsidies and,

thus, is also run as a charitable insurance model, having features of both models (Dror 2014).

2.4. Why is microinsurance important?

Microinsurance is potentially an important new risk management tool for low-income people in developing countries for several reasons. Vulnerability inordinately affects poor people and reinforces or exacerbates their poverty. Regarding susceptibility to risk, for example, poor people

- typically live and work under more crowded, unsanitary, stressful, or unsafe conditions
- suffer from higher rates of malnutrition (which make them more susceptible to illness and injury)
- lack the education necessary to make informed preventative or reactive choices (or the money to implement those choices), and
- frequently hold beliefs (for instance, in the importance of dowries or ostracising widows) that aggravate their situations when risks (like the death of the breadwinner of a household) actually materialise

When these risks materialise, low-income people are, furthermore, frequently less able to cope. Regular solutions (like medical care) are, for example, often inaccessible

(particularly given the general shortage of facilities in rural areas), unaffordable (considering the costs of treatment, transportation, bribes, medicines, waiting, aftercare, and missed work), or of terrible quality (with provider absenteeism, poor staff training, corruption, patient abuse, and chronic shortages of equipment, supplies and medicine endemic in health-care facilities which predominantly serve low-income clients in developing countries). Social safety nets (like national health insurance schemes) are — due to narrow tax bases, corruption, inefficiency, and other government priorities like indebtedness — commonly weak or non-existent. Moreover, the low-income people's plentiful and creative array of informal risk mitigation techniques is often insufficient, particularly when risks covary or repeatedly occur over a short period.

These informal arrangements, which are classified into self-insurance and shared insurance, bear examination because of their prevalence within poor communities (Morduch 2003). Of these, self-insurance is the most widespread and significant form of informal risk mitigation. It consists of ex-ante and ex-post approaches, which are implemented before and after the hazards in question occur. The predominant ex-ante approach is savings. Because appropriate savings facilities

and products are typically unavailable in low-income areas, however — and because monetary savings are vulnerable to inflation — asset build-up and drawn-down are particularly popular means of self-insuring. However, low and volatile incomes (and threats to the accumulated assets themselves) make it difficult to amass large enough sums to adequately militate against shocks. Other ex-ante strategies include calculated, and often conservative, employment, production, and social-familial strategies, such as

- diversifying occupations and crops
- working for less, but under more secure arrangements (such as “tied labour”) using less effective, but cheaper combinations of production inputs (such as less high-cost fertilisers and more low-cost labour)
- migrating to places with uncorrelated income patterns, and
- tactically selecting marriage partners, fostering children, and cultivating friends

Whilst lowering uncertainty, many of these approaches lead to inefficient outcomes through which people sacrifice more profitable, but riskier activities and the adaptation of potentially valuable new technologies to achieve some degree of income with certainty and “limit exposure only to... shocks that can be handled with the means available” (Morduch 1995, 104). For

example, Walker and Ryan estimate that households sacrifice up to 25% of their average income to reduce exposure to shocks in certain parts of India (1990, 197).

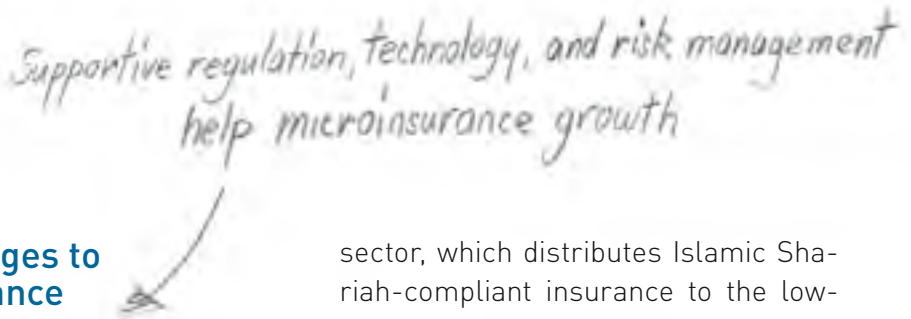
In contrast, ex-post approaches include borrowing (though constraints often exist on the availability and cost of credit), changing consumption patterns (for example, eating less or withdrawing children from school) and adjusting labour supply (working longer hours or employing children). Because these strategies are implemented reactively by households under duress, they typically have less favorable terms and prove more exacting on family finances.

Shared insurance, on the other hand, includes reciprocal loan- and gift-giving practices and participation in rotating savings and credit associations (ROSCAs), through which a group's members regularly contribute equal sums of money and sequentially receive the proceeds. Shared insurance schemes are typically organised amongst families, neighbours, or other groups of people with the ties necessary to identify and curb moral hazard amongst participants. Information asymmetries and enforcement problems may exist even amongst close-knit people (Morduch 1999, 189). Moreover, the financial capacity of intra-family lending within poor

families is in any case limited. Therefore, borrowing with interest from professional lenders and liquidating saving are more important coping mechanisms than shared insurance approaches (Morduch 1999, 189; Binnendijk et al. 2012).

Whilst these risk mitigation instruments collectively enable real and significant consumption-smoothing, they do not provide complete coverage — and ironically prove costly for households in terms of everything from foregone profits to intensified gender problems, since very often women bear the brunt of strategies like migration, fostering, and strategically-arranged marriages. Besides bridging the gap, microinsurance schemes can be created to complement or crowd out the best and worst of these approaches whilst enabling low-income people to pursue more profitable income-generating activities and more gratifying personal relationships.

Supportive regulation, technology, and risk management help microinsurance growth



2.5. Challenges to microinsurance

The success of the microinsurance sector is based on the three guiding principles of outreach, sustainability, and proving benefits for all. In order to achieve these milestones and to increase the penetration, a combination of regulation, technology, and risk management is required.

When there is market and demand identified, distribution diversity is one of the key factors to success. In South America, microinsurance has successfully increased penetration, as well as enlarged into the middle class, by using a variety of retail distribution channels across Columbia, Mexico, Peru, Guatemala, Bolivia, and especially Brasil. Across Asia, retail distribution has not been facilitated to a similar degree. However, this seems to be changing as new approaches are developing. These include, for example a dengue fever insurance sold in supermarkets in Indonesia, various microinsurance products in 7/11 stores in Thailand, and multi-level marketing schemes in the Philippines, just to name a few. Cooperative selling of insurance is a large portion of the distribution market for microinsurance and well illustrated in the microtakaful

sector, which distributes Islamic Shariah-compliant insurance to the low-income sectors in the Islamic regions of Asia, Africa, and the Middle East, encompassing a charitable component by Shariah law.

It can be argued that one of the reasons the penetration of microinsurance over the past decade has not been as fast as many hoped, is that many of the assets in the sector are deemed intangible, non-liquid assets embedded in projects and internal processes, and are, therefore, invisible to boards of large corporations and government organisations. These assets are employee skills, information technology infrastructure, corporate culture, and innovative ideas, and can be viewed in the future as the new intellectual property and patents of the organisation. Measuring the value of these intangible assets is an accounting function and moves the asset value to the balance sheet, and, therefore, to the attention of stakeholders, where the availability of financial support depends. As the intangibles are an important part of a new business, this increases the chance of new ideas becoming reality. Models that only value physical assets are less able to take advantage of new markets (Dror 2012).

Expanding the discussion of the importance of embedded intangible assets may be out of scope in this book, but the way projects are viewed in terms of value creation from the base of the pyramid will have a correlation with impact assessments of the sector. These assets are the distillate of decades and centuries of social and fiscal investment by developed societies.

2.5.1. Regulation

Microinsurance is a commercially viable market with 2.6 billion people living in the range between Int\$1.25 and Int\$4 per day⁷, giving rise to a US\$33 billion market. Coupled with that is the government and aid-supported microinsurance market with 1.4 billion people on less than Int\$1.25 per day⁷, giving rise to a US\$7 billion market (Swiss Re 2010, 9). The principles of good governance apply to microinsurance the same way as other sectors. Regulation is required and will have an impact on how insurance is sold, bought, and distributed at base of the pyramid. A balance needs to be sought, as too little or too much regulation can negatively impact the sector. The ability of governments to move from constitution to charter to commercial implementation and to create adequate legal frameworks will greatly facilitate the prac-

tice of microinsurance. Recent success in this area has been seen in the Philippines as they published a charter based on their constitution.

2.5.2. Technology

Technology is the key, because without it there is no outreach to rural areas and no capability to adequately analyse, on a macroeconomic and risk management basis, the sustainability and profitability of the schemes designed to help the base of the pyramid. It is these automated techniques that will allow the value creation of assets to appear on balance sheets and raise attention to the right stakeholders in public and private partnerships required to educate and finance the microinsurance sector. These would be simple procedures allowed by regulation. Good data is vital to successful impact measurement at all levels. The microinsurance sector needs to leverage the advent of the global cloud computing networks and the security of data offered within them. The rise of Internet and mobile technology is a game changer.

2.5.3. Risk management

Regulation and technology capabilities dovetail with the importance of risk management and risk transfer in the microinsurance sector. Insurance is the risk industry and risk transfer is a trodden path that has enabled the

⁷ An international dollar has the same purchasing power as a U.S. dollar has in the United States. The Int\$ is adjusted over time by reference to gross national income and exchange rates of local currencies to US\$.

industry to survive in over 300 years of trading. Reinsurance is one of the risk transfer mechanisms used in the industry and is vital to the microinsurance sector as a capital base can be offered to primary insurers or, indeed, direct to communities to handle their risk transfer affairs. Evaluating and measuring the assets identified will greatly assist the flow of capital from reinsurance and capital markets (alternate risk transfer) to microinsurance projects. A very important aspect is to show that the internal processes that handle claims have a tangible value for impact assessments.

Regulation needs to allow for lower capitalisation as an entry point for the microinsurance sector and allow the additional premium assets to be included in an impact assessment for the computation of solvency on microinsurance schemes. This is especially important as more complex health, weather index, and innovative climate change (linked to food shortage) products are introduced in the market place. There is a need to measure the key indicators around the solvency ratio and the expense ratio, which is the cost of distribution and the cost per transaction. This, in turn, generates a set of official performance standards established by regulatory authorities for effective delivery of microinsurance and impact analysis.

The next stage of this process leads to the stochastic — or actuarial — measurement of microinsurance, using Dynamic Financial Analysis, or measuring risk mathematically using probability theory, which needs good historical data to achieve (Piesse, in preparation). This process will align the intangible assets to the company strategy and align to the regulator strategy. A key alignment here is that of literacy, health care, financial inclusion, and risk transfer, thereby liquefying the intangible assets and appearing on a balance sheet. This includes risk from non-cost effective information technology projects whose costs prohibit microinsurance schemes and prevents them from going into production. However, impact models and good data are not sufficient in isolation. There is a need for independent risk quantification that is capable of bringing stakeholders together for sustainable risk transfer solutions based on the increase in natural disasters and climate change that mostly affect newly penetrated microinsurance bases. This is an important correlation of catastrophe risk with market, credit, underwriting, macroeconomic, and insurance risk combined in one holistic risk analysis. The development of effective catastrophe microinsurance needs reinsurers, catastrophe modelers, insurers, governments (public private partnerships), regulators, World Bank/United Nations/ADB,

and development rural banks to step up to form alliances to protect people from natural disasters and other mega risks. There is no such thing as micromodeling, and microinsurance is subject to the same scenario simulation as other insurances. Pioneer work has been done in two pilot locations in India by the Micro Insurance Academy (MIA) in collaboration with the Asia Risk Center (ARC), an affiliate of Risk Management Solutions (RMS), the world's leading risk modeling company, on crops risk assessment and quantification of climate change contributions. A climate-vulnerability mapping program, based on an extensive household survey and weather data, has been established (Sharma and Jangle 2012; Sharma 2012).

In summary, sustainability of microinsurance, or the ability to create long lasting renewable products in the sector that leads to benefits for all, is achievable by a confluence of regulation, technology and impact analysis, and the various methods of risk management that lead to that impact analysis. We must look at the holistic picture and not risks in isolation. Right now, microinsurance and its differing models make it a younger cousin of the larger insurance industry. As microinsurance grows and more people become included, microinsurance and its differing models will become the status quo.

2.6. Conclusion

This chapter has illustrated three ways of how to approach microinsurance: one focuses on the target group, another on the product, and the third one on the processes. All approaches involve the potential input of reinsurance capacity and the approach taken will have a different impact on the assessment process. There is a big difference for policyholders paying a fixed premium set by the industry in contrast to willingness to pay, which is the community approach. Whatever the approach, the guiding principles of outreach and sustainability remain the same, as access needs to be gained to remote areas and the programmes that are created must be renewable over time to make sure sustainable access to insurance coverage is received. These guiding principles, along with technology, good product design, and flexibility are required for microinsurance schemes to fit their community.

In order to assess the impact of catastrophe, market, credit, insurance, underwriting, and operational risk on microinsurance projects and communities, scenario analyses are required with a definite shift in the direction of dynamic financial analysis, where the whole process is managed mathematically through actuarial models. This process requires good historical data,

such as claims history and weather pattern data from weather stations, and measurable indices. In weather-based indices, the impact of basis risk needs to be understood to make sure that the right claims are paid to the right claimants.

Barriers to penetration are the lack of available capital pointed at microinsurance projects and the governmental support on education. We identified that this required a set of committed and willing stakeholders who have a business interest in the project creating a tangible asset of the project and move it to the corporate balance sheet for recognition. This step should assist in providing adequate risk transfer facilities to the microinsurance schemes.

Natural catastrophes and non-correlated risks are major threats to the success of microinsurance, as it is the policyholders of these schemes that are most exposed to these risks. Reinsurance catastrophe pools established via public private partnerships would mitigate this threat.

Finally, there is infrastructure and political will. Without that will and balanced regulation, microinsurance will not live up to its full potential.

References

- Binnendijk E., R. Koren, and D. M. Dror. 2012.** Hardship financing of healthcare among rural poor in Orissa, India. *BMC Health Services Research* 2012: 23-36. doi:10.1186/1472-6963-12-23
- Churchill, C., ed. 2006.** *Protecting the Poor: A Microinsurance Compendium*. Geneva: International Labour Office.
- Churchill, C. and M. J. McCord. 2012.** Current trends in microinsurance. In: *Protecting the poor: A microinsurance compendium*. Vol. II, ed. C. Churchill and M. Matul, 8-39. Geneva: International Labour Organization.
- Dror, D. M. 2012.** Investing in embedded intangibles to enhance solvent demand at the base of the pyramid. Presentation held at the Global Risk Forum (GRF) One Health Summit, Davos, Switzerland. February.
- Dror, D. M. 2014.** Health microinsurance programs in developing countries. In: *Encyclopedia of health economics*. Vol 1, ed. A. J. Culyer, 412-421. San Diego: Elsevier.
- Dror, D. M. and C. Jacquier. 1999.** Microinsurance: Extending health insurance to the excluded. *International Social Security Review*. Geneva, ISSA, 52(1): 71-97. <http://ssrn.com/abstract=180448>
- Friedman, M. and L. J. Savage. 1948.** Utility Analysis of Choices Involving Risk. *Journal of Political Economy* 56(4): 279-304.
- Gruber, J. 2007.** *Public Finance and Public Policy*. 2nd ed. New York: Worth Publishers.
- Ingram, M. and M. J. McCord. 2011.** Defining "Microinsurance": Thoughts for a journey towards a common understanding. MicroInsurance Centre discussion paper presented at Microinsurance Network member meeting, Königstein. June.

Insurance Regulatory and Development Authority (IRDA). 2002. Obligations of insurers to rural or social sectors. Regulations, Reference number 3218_GI_2002_ENG. <http://www.irda.gov.in>

Insurance Regulatory and Development Authority (IRDA). 2005. Insurance regulatory and development authority micro-insurance regulations. http://www.irdaIndia.org/regulations/IRDA-Mirco-Ins_reg-2005.pdf

International Association of Insurance Supervisors (IAIS) and Microinsurance Network. 2007. Issues in regulation and supervision of microinsurance. Issues Paper. <http://www.iaisweb.org>

Martinez, W. H. 2012. Role of the state and regulation in promoting microinsurance - The Philippine experience. Presentation held in Ho Chi Minh City, Vietnam. July.

Microinsurance Network. 2013. Microinsurance. 1 billion covered by 2020. Press Release, January 30.

Morduch, J. 1995. Income smoothing and consumption smoothing. *Journal of Economic Perspectives* 9(3): 103-114.

Morduch, J. 1999. Between the market and state: Can Informal insurance patch the safety net? *World Bank Research Observer* 14(2) (August): 187-207.

Morduch, J. 2003. Consumption smoothing across space: Testing theories of risk-sharing in the ICRISAT study region of south India. In: *Insurance against poverty*, ed. S. Dercon, 38-56. New York: Oxford University Press.

Piesse, D. In preparation: Dynamic financial analysis for microinsurance.

Sharma, B. 2012. Assessing vulnerability to climate change using GIS. *Geospatial World Weekly*. 31 December.

Sharma, B. and N. Jangle. 2012. Assessment of vulnerability to climate risks in Ahmednagar (Maharashtra) and Vaishali (Bihar) using GIS techniques. Proceedings of 13th ESRI India User Conference.

Swiss Re. 2010. Microinsurance—risk protection for 4 billion people. *Sigma* 6. http://media.swissre.com/documents/sigma6_2010_en.pdf

Vaté, M. and D. M. Dror. 2002. To insure or not to insure? Reflections on the limits of insurability. In: *Social reinsurance: A new approach to sustainable community health financing*, ed. D. M. Dror and A. S. Preker, 125-152. Washington, D.C.: World Bank.

Walker, T. S. and J. G. Ryan. 1990. *Village and household economies in India's semi-arid tropics*. Baltimore, MD: Johns Hopkins University Press.

Yeshasvini Trust. 2011. Achieve-ment since inception of the scheme. <http://yeshasvini.kar.nic.in/achieve.htm>

Why and how does impact happen?

Simon Quinn, Stefan Dercon, and Michal Matul



3.1. Risk and its implications for poverty¹

When we think of *risk*, we often think of *poor outcomes* – for example, a failed harvest, or the death of a family breadwinner, or falling victim to a crime. However, the welfare costs of risk arise even if such outcomes do not happen; the costs of risk are greater than merely the possibility of a poor outcome. This is the sense in which risk itself – rather than merely the realisation of poor outcomes – deserves recognition as a policy priority.

Specifically, risk can affect decisions and welfare in two distinct ways:

Firstly, poor households incur costs through the adoption of *ex-ante risk management strategies*.

Secondly, households suffering negative shocks incur costs through *ex-post risk coping strategies*.

The sheer presence of risk reduces welfare and, thus, needs management

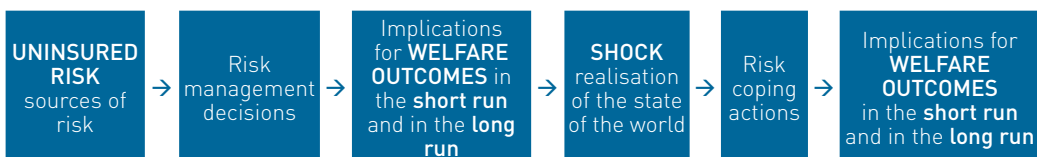
The distinction is summarised in the following diagram, from Dercon and Kirchberger (2008).

Ex-ante risk management strategies

Ex-ante risk management strategies are strategies adopted by poor households in order to reduce their exposure to future risk. Such strategies generally involve trading off higher average consumption in order to reduce the likely variability in consumption. This can be done, for example, by limiting specialisation in production or by building up *precautionary savings*.

There is a useful analogy here between risk management strategies and the

Diagram: Risk and outcomes



¹ The discussion here draws upon the review in Dercon and Kirchberger (2008).

problem of self-sufficiency. In general, economists view self-sufficiency as an inefficient objective, because a household trying to be self-sufficient must engage in economic behaviours for which it does not have a *comparative advantage*. For example, a household whose members are skilled at making clothing should not generally try to grow all of its own food; conversely, a household that owns a small farm should not generally try to produce its own clothing. So it is with risk management: quite obviously, a poor household has no comparative advantage in providing insurance products, and it is particularly costly for such a household to be forced to do so. But this is precisely what happens when poor households have limited access to microinsurance: households are forced to diversify away from their preferred activities in order to minimise risk. So, for example, farmers facing higher risks may plant a variety of less productive crops in order to limit the consequences of any single crop's failure; similarly, households may under-invest in their children's education in order that the children's labour may generate precautionary savings.² In

this sense a lack of adequate insurance facilities is a *missing market* problem, and microinsurance products can be understood as helping markets work better for the poor.

Ex-post risk coping strategies

Ex-post risk coping strategies are strategies adopted by poor households in order to mitigate the effects of poor outcomes after they have occurred. That is, risk coping strategies are strategies for coping with the aftermath of a risk that occurred, rather than coping with the risk itself. Nonetheless, the strategies require households to incur costs that could be reduced if risks were better insured. Risk coping strategies typically involve smoothing the consequences of costly events – and this typically occurs both over time and over a network of other economic actors. For example, a household suffering the death of a breadwinner will typically respond by drawing upon household savings, whilst reducing its own future consumption and by receiving transfers from other households in its community. Even though both strategies are described as ex-post, both typically require substantial ex-ante planning, and incur substantial ex-ante costs. For example, a household must build up precautionary savings before a breadwinner's death. Similarly, informal insurance arrangements amongst a network of households can generally

2 The discussion here has characterised the poor as "risk averse". In cases of extreme poverty, we may expect to find "risk-loving" behaviours, where poor households effectively choose to "gamble for resurrection". For example, there were reports from rural Zimbabwe in 2009 that starving Zimbabweans were panning for gold. However, this kind of behaviour is generally limited to cases of extreme poverty—that is, cases in which the immediate policy response would likely be emergency aid rather than the introduction of a microinsurance product.

only be implemented as an agreement to insure future shocks.

These kinds of strategies serve to limit the consequences of risk for poor households. Certainly, it would be a mistake to describe any poor household as *uninsured* merely because the household does not hold any formal insurance products. In insurance, as in most aspects of life, poor households can use a remarkable variety of strategies in order to deal with difficult circumstances. However, risk coping strategies are far from perfect. Households may be dissuaded from amassing adequate precautionary savings, for example, because of *sharing norms* that demand such wealth to be distributed amongst a family or community network. Similarly, informal insurance arrangements may be difficult to implement because their participants may lack a formal or credible way of committing to a contract. Additionally, participants may worry that other members may join only if they anticipate making a claim (adverse selection), and may then have inadequate incentives to mitigate risks themselves (moral hazard). Further, community insurance arrangements may be useless in responding to highly correlated shocks – for example, the risk that an entire village may lose its harvest in the same season.

For these reasons and more, empirical studies tend to find that informal risk coping mechanisms achieve only partial insurance; negative shocks do cause reductions in household wealth and household consumption, even if risk coping and risk management strategies dampen the shocks' effects.

From poor outcomes to poverty

There is a difference between having a bad harvest and living in poverty: a bad harvest is a particular outcome from a particular process, whereas poverty is a prolonged way of life. There are several reasons that exposure to risk in poor communities may push households into poverty. The most obvious reason that a temporary negative shock might cause a permanent welfare loss is that a shock may be so severe that a household is unable to return to its previous circumstances. This might be the case, for example, for a household facing a prolonged drought. But there is a more subtle reason, too: many of the risk management and risk coping strategies discussed earlier are relatively more costly for households that are poorer. For example, it may be the case that poorer and wealthier households both choose to accrue precautionary savings, but that only the poorer households need to withdraw their children from education in order

Exposure to risk
may be a cause for poverty traps

to do so. For this reason, exposure to risk may be a cause for *poverty traps*, meaning that poor households may face a vicious cycle precisely because of the particularly high costs that they incur for risk management and risk coping strategies.

In summary, on one hand, we should expect an effective microinsurance product to bring substantial benefits to poor households – not only by mitigating the consequences of poor outcomes, but also by shifting behaviour away from costly risk coping and risk management strategies. We should expect important short-term gains, but may even anticipate permanent gains in welfare for households otherwise facing a risk induced poverty trap. However, there are likely to be several mechanisms by which these benefits accrue – and those mechanisms may vary in their importance across different contexts. On the other hand, insurance, as any intervention, can also have unintended effects that can reduce the welfare of insured members or others from the same communities. It may not be an effective mechanism for all groups under all circumstances. For these reasons, it is important that

both the development and the evaluation of microinsurance products rely upon a clear framework for a *theory of change*.

3.2. Theory of change

The importance of means and mechanisms

This section discusses some of the likely mechanisms by which microinsurance may operate; that is, the section considers not only what kind of final *impacts* we might expect from microinsurance, but also the intermediate steps by which these impacts may come about. There are at least two reasons why it is important to consider such mechanisms:

- Firstly, any microinsurance product is likely to be refined and improved over time. In order to understand what kinds of refinements are likely to be valuable, a researcher needs to understand not only what the average effect of a particular product is, but also how that effect occurs. For example, suppose that a particular microinsurance product is found to be effective, but only

amongst a small number of households choosing to adopt it. This could, for example, be the result of insufficient advertising and promotion, or of households not understanding the product, or of households understanding the product well, but deciding that its costs are too great. Any policy maker wishing to improve the product would need to understand not only the average effect of the product, but also the mechanisms determining its adoption and implementation.

- Secondly, successful microinsurance products are likely to be implemented in different communities and different contexts. For example, suppose that a particular microinsurance product is found to work well in rural western Kenya. Without understanding how and why it has been successful, it is difficult to have any confidence that the product's success could be replicated in, for example, urban India.

This point has been noted several times by microinsurance researchers. For example, Radermacher et al. said this about the notion of impact in the microinsurance context:

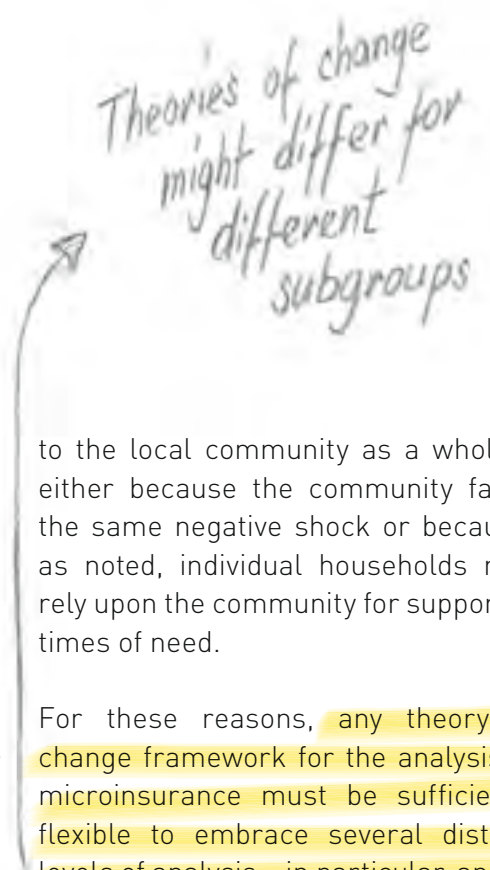
Impact encompasses the changes that microinsurance makes to the economic or social circumstances of insured people or their households, enterprises or communities. It can

be positive or negative, affect both insured and uninsured populations, occur either before, ex-ante, or after, ex-post, insured events happen, and have micro-, meso- and macro-level implications, often in ways that are linked. For example, livestock cover can provide payouts that smooth household consumption ex-post after animals become sick or die, but can also pre-emptively encourage households to reallocate money they may have saved for such emergencies to other more profitable ends ex-ante, before any problems occur. Similarly, health microinsurance can improve policyholders' health through increased access to care, which can additionally reduce local disease burdens and thus improve the health of nearby uninsured people too. *As impact is multifaceted and manifests itself in different ways, we need to be aware of each intervention's myriad potential effects and their relationships to each other.*" (2012, 59-60, emphasis added)

A clear theory helps understand the reasons for impact

In short, one size does not fit all and a careful analysis of mechanisms is important to ensure that policy makers learn effectively from past practices. This section structures a basic theory of change for microinsurance, in order to identify some of the important mechanisms by which microinsurance products may affect the welfare of poor communities.

Theories of change
might differ for
different
subgroups



Level of analysis

Risk is a phenomenon that operates at many levels of community structure. Typically, the primary unit of analysis is the poor household. As noted, negative shocks are likely to be felt most directly by the household collectively – for example, through the destruction of a household's crop or the death of a household breadwinner. Similarly, the household may respond to risk collectively – for example, by withdrawing children from school as part of a risk management strategy to build precautionary savings. However, the household is not the only useful level for analysis. Research in development economics increasingly emphasises the divergent interests of different decision makers within the household; it may be that different household members respond in different ways to particular risks (this is sometimes termed the “non-unitary model of the household”). Conversely, risk may have consequences that apply

to the local community as a whole – either because the community faces the same negative shock or because, as noted, individual households may rely upon the community for support in times of need.

For these reasons, any theory of change framework for the analysis of microinsurance must be sufficiently flexible to embrace several distinct levels of analysis – in particular, analysis at the levels of the individual, the household, and the community. Similarly, any empirical analysis of a microinsurance product should consider the possible role of *spill-overs*, allowing the possibility that the consequences of microinsurance adoption are felt even by those not directly involved.

A framework for a theory of change

Any analysis of microinsurance – whether theoretical or empirical – must consider at least four key steps:

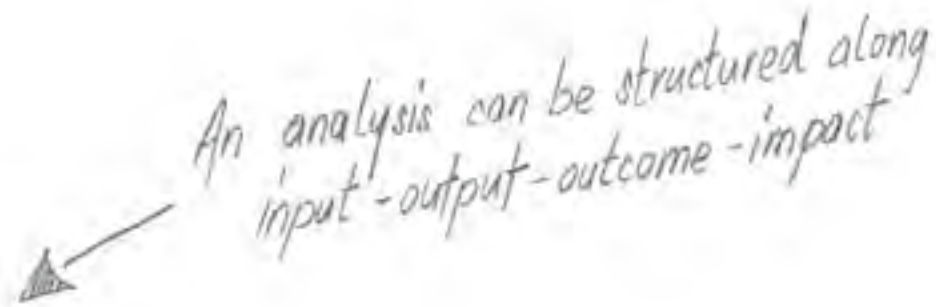
- Firstly, some provider – whether a government, an NGO, or a private company – must offer some kind of microinsurance product. This can be framed as the input³ to the microinsurance process.
- Secondly, the target community must decide whether to adopt the product. Depending upon the product and the context, this may be either an individual decision (for example, in the case of a household agreeing to a health insurance plan), or a decision for an entire community (for example, in the case of a group of farmers agreeing to adopt index insurance for rainfall). This decision may involve consideration of the demand for the product from prospective clients, the costs involved in the product to both to the provider and the clients, and the subsequent use of the product. All of these aspects can be considered part of the output of the microinsurance process.
- Thirdly, the adoption of microinsurance may change the types of services available to poor communities, and the types of behaviour that members of those communities choose. For example, microinsurance clients typically receive a bundle of contractual rights, often accompanied by new information and new services (for

example, rights to claim specific sums in case of particular losses); clients may respond to this by changing their financial planning and their management of risk. As Radermacher et al. (2012, 59) have argued: “Insurance is not an end in itself. Households purchase (and donors support the development of) microinsurance because they want to manage risks better.” This process of changing available services and client behaviour can be considered as the *outcome* of the microinsurance process.

- Fourthly, the outcomes of microinsurance may ultimately improve client welfare. If a microinsurance product is working, this should occur through a reduced exposure to risk. However, this may have long-term benefits (the earlier discussion noted the possibility of a risk induced poverty trap), and it may be that a well-designed microinsurance product reduces a household’s general vulnerability to poverty and even breaks a *poverty cycle*. All of these consequences should be considered as microinsurance impact.

One of the main reasons for using a comprehensive theory of change framework is to avoid wasting resources on evaluating outcomes or impacts of badly designed schemes. Hence, the analysis of inputs is crucial to select the right scheme to be evaluated. The Product, Access, Cost

3 There are many other definitions of the terms input, output, outcome, and impact. See, for example, the classification proposed by the Organisation for Economic Co-operation and Development (OECD) Development Assistance Committee (DAC).



and Experience (PACE) tool, presented later in this chapter, can be used to conduct this initial assessment.

The table below illustrates the four steps used to consider the mechanisms by which health microinsurance may improve households' welfare. The table shows how potential innovations may contribute to the process *input*, how pragmatic implementation and careful advertising or information lead to rational client adoption (the process *output*), how adoption of such a product may improve health services, as well as clients' risk management practices and the general health of their community (the process *outcome*) and, finally, how such a product may lower vulnerability to poverty and, potentially, break a risk induced poverty trap: a valuable microinsurance impact. This example also highlights some potential unintended, negative outcomes.

Outcomes: what should we look for?

The theory of change framework complements the earlier discussion by suggesting several specific ways in which

microinsurance may shift the behaviour of clients and their communities. It suggests several important potential outcomes that we should look for in empirical analysis. The following discussion flags several important potential outcomes. Of course, the lists of outcomes (and impacts) provided here are indicative rather than exhaustive.

Better risk management practices

As mentioned earlier, informal methods are not sufficient and often costly in helping low-income households to manage risks. Selling assets or taking high-interest loans from money-lenders can have devastating consequences on household welfare in the long-term. Microinsurance promises to provide a more cost-effective solution for certain risks and for certain groups. However, by relying on formal insurance products that cover only certain risks, insured individuals might invest too little in other risk management strategies (e.g., savings or the social network) to deal with the risks not covered by the formal insurance.

Table: Example of a four-step theory of change framework for health microinsurance

Input	Output	Outcome	Impact
- Expanded benefits (product innovations) - Comprehensive - Outpatient to complement inpatient - Drugs - Defined benefit, incl. maternity - Benefits to non-claimants, incl. free health checkups - Prevention campaigns - Alternative financing mechanisms - Savings accounts - Subsidies - Conditional cash transfers (CCT), remittances - Credit - Process innovations - Third party administrator (TPA) and other back-end - Front-end - Smart controls for moral hazard, adverse selection, fraud; managing provider networks - Consumer education / marketing - Business models - Public private partnership (PPP) - Managed care model - Insurer-led model - Hybrid model - Mutual	- Higher appeal to clients, higher client satisfaction, higher demand and renewals (demand/scale) - Trimmed supply costs and better control over costs (costs) - Clients and their families use insurance payouts and health-care services in a rational way (use)	- Access to health microinsurance improves health-seeking behaviours - Access to health microinsurance improves quality of health services - Access to formal microinsurance products improves clients' financial capabilities and makes their risk management behaviours more cost-effective (e.g., less high-interest borrowing, reduced out-of-pocket expenses) - Upper poor and vulnerable non-poor benefit more from microinsurance than other groups - Access to microinsurance has positive or negative spill-over effects on non-clients living in the same communities (e.g., improved quality of health care for all or weakened informal risk sharing networks for uninsured members).	Use of microinsurance products leads to improved well-being at household / individual level; lowers vulnerability to poverty and breaks the poverty cycle

This can leave them more vulnerable than before.

Additionally, if many people from a community take insurance (and invest less in informal risk sharing), it can weaken the reciprocal social

networks to the extent that individuals without insurance might be worse off than before. But one can also argue that in the event of covariant risk that affects the entire community, premium payouts can financially strengthen the community's social networks.

More investment in productive activities

The earlier discussion noted that poor households facing risk are likely to shift their efforts towards less productive activities in order to reduce their expected variation in consumption; that is, risk management strategies are likely to discourage households from specialising in the most productive activities possible. It follows that an effective microinsurance product is likely to induce households to shift time and resources towards more productive activities. For example, we might expect a farming household to plant crops with higher expected yields, even if those yields may be more variable. We might expect urban households to consider starting new enterprises, even if such an enterprise may be a risky undertaking. And, we might expect poor households generally to invest more in their children's education, even if they therefore forego some opportunities to accrue precautionary savings. On the flip side, the *peace of mind* triggered by being insured can push some households into taking more uninsured risk than before.

Changes health-seeking behaviours and access to health care

Many microinsurance products are designed to protect clients against the costs of health shocks. Self-evidently,

any evaluation of such a product must measure its effects upon clients' standard of health. However, there is a more subtle way in which microinsurance may improve health: it may encourage clients to place greater value on their health. This would be the case if, for example, a microinsurance product was to increase the effectiveness of households' own health investments. For example, a client may be less likely to smoke if he or she expects to live longer. In some health microinsurance schemes, clients are also "forced" to change their health-seeking behaviours through a specific scheme design, such as obligatory free health checkup every year or distributing mosquito nets for malaria prevention. Health microinsurance can also create distortions. For example, when implemented without good controls to manage providers' networks, it can

Effects
can be positive or
negative, and affect
the insured as well
as the uninsured

contribute to an increase in health-care prices, which can make it more expensive to use the services by uninsured members of the community. Or, if the local health-care system has limited capacity, the health microinsurance can make it even more difficult for uninsured members to access health care when insured patients are served first. Through these sample channels, the equity of accessing health care can be lowered at the community level, because the poorest rarely can afford insurance.

Improved financial capability

There are several reasons that we might expect microinsurance products to improve clients' capability to manage financial decisions and the assessment of risk. Firstly, the very process of considering, negotiating, and agreeing to a microinsurance contract may improve clients' financial capability. For example, a household saving for a rainy day may consider why it saves so much, and how a microinsurance product may reduce that savings incentive. Secondly, microinsurance may actually serve to shift households' preferences over risk and over decisions about future events. Research in economics has increasingly emphasised the importance of *behavioural biases*, by which – for example – households may have inconsistent preferences over future events (weighing

present events too highly relative to the future) and inconsistent preferences over risky outcomes (weighing tiny risks too highly relative to larger risks).⁴ Microinsurance may help to overcome these biases by attaching a specific price to a particular set of future risky outcomes. Microinsurance clients may weigh future – and risky – events differently if they can weigh those events by using a price that is immediate and certain.

Impacts: what should we look for?

Reductions in vulnerability

The most important potential impact of microinsurance is a reduction in the variability of income and consumption – and, therefore, a reduction in clients' vulnerability. If a microinsurance product can achieve this result, most clients will surely consider it a success; if it does not achieve this result, any other benefits will likely seem incidental at best. It follows that any evaluation of a microinsurance product must evaluate whether the product has reduced clients' vulnerability and how it impacted vulnerability of non-clients living in the same community. Ideally, this should be measured both in terms of actual changes in the variability of consumption and income, and whether clients perceive any such changes.

⁴ See Kahneman and Tversky (1979) and Tversky and Kahneman (1992) for more on prospect theory.

Changes in asset accumulation

Asset accumulation may or may not be a desirable outcome of a microinsurance programme. The earlier discussion noted that precautionary saving – a form of asset accumulation – is one of several inefficient risk coping strategies that poor households use. If a particular community or household retains an inefficiently high level of savings for precautionary purposes – for example, by placing children into work rather than into school, or by passing up valuable opportunities to open new enterprises – then microinsurance may lower the rate of asset accumulation, and this may represent a welfare gain. Conversely, it may be that a community holds inadequate savings because of a high risk of loss or appropriation. In that case, we might hope that microinsurance would encourage asset accumulation. For these reasons, the effect of microinsurance upon asset holdings is a critical question for any future empirical research; however, it is also a question that must be intimately linked to other impacts – for example, whether households are more likely to open a microenterprise, or whether households are more likely to invest in their children's education.

Changes in educational levels⁵

There is a critical relationship between risk and education, as noted earlier. Because education is a costly investment in future welfare, it is particularly vulnerable to be purloined to offset present risk. Households adopting such a strategy are likely to forego substantial future earnings; further, this kind of strategy may be one reason for a risk induced poverty trap. For these reasons, households' investment in education and subsequent change in educational levels is a critical impact that any evaluation of microinsurance should cover. Further, such an evaluation should measure the progress and effects of the children's education at different stages. It may be, for example, that microinsurance has limited effect upon whether children attend primary school, but a critical effect upon their attendance at secondary school (i.e., when they reach an age where their labour becomes more valuable, and hence a more attractive alternative option).

5 One can classify health and education impacts as outcomes if just the investment in education or health is considered. For example, investing more in health seems like changing behaviours (outcome) that will lead to better health status (impact). The distinction between outcomes and impacts is somewhat fluid. We have decided to keep them under impacts because they involve long-term changes in behaviour (i.e., in the sense of a permanent shift in the value placed upon health, or a permanent shift in human capital, etc.).

Changes in health status

The earlier discussion of microinsurance outcomes noted two important pathways by which a microinsurance product may improve health investment: by insuring some health costs directly, and by increasing the relative value of good health to poor communities. These pathways hold obvious relevance for any microinsurance programme in the area of health; however, they are also relevant potential impacts for microinsurance in other areas (for example, rainfall index insurance). Effective microinsurance may free funds from inefficient risk coping and risk management strategies for higher investments in health – even in the case of microinsurance products far beyond the health sphere. As mentioned earlier in the outcome section, some undesired outcomes of health microinsurance can lead to negative impacts at the community level in the long-term. For example, the existing inequalities in health status can be further increased if no mechanisms are created for the poorest to access microinsurance.

Higher self-reported well-being

Self-reported well-being is certainly not a definitive outcome measure for assessing the impact of microinsurance. It is well known, for example, that respondents may shift their well-being

reference point over time. However, it is clearly an important impact nonetheless. There are at least two reasons for this. Firstly, self-reported well-being may act as a useful summary statistic, in order to weight and combine the various effects of different impacts. For example, suppose that a rainfall insurance product causes households to increase their investment in health, but somehow causes a reduction in their children's education. Should this be considered an overall improvement in welfare? Self-reported well-being measures may help to answer this question. Secondly, self-reported well-being may capture the psychological effects of risk and its mitigation – something that standard impact measures may overlook. Economic policy makers and researchers typically measure outcomes in terms of tangible, economic impacts. This is an eminently sensible approach, but may overlook – for example – a general sense of stress or helplessness associated with living under substantial risk. An amelioration of these effects may be no less important than traditional measures for considering the value of microinsurance. Carefully framed questions on self-reported psychological effects may, therefore, be extremely valuable in understanding the broader consequences of any microinsurance programme.

One effect does not fit all: different impacts on different clients

To this point the discussion has considered microinsurance clients generally, without drawing any explicit distinctions between different types of clients or non-clients living in the same communities. However, any given microinsurance product will likely be adopted by clients in quite different circumstances – for example, by men and by women, by wealthier and by poorer, by those with large households and by those with small, etc. It is likely that any product will affect these clients differently. There are three key reasons for this kind of heterogeneity.

- Firstly, different clients are likely to face different risk – different both in its origin and its seriousness. For example, the health risks of choosing to have children are obviously much greater for women than for men; similarly, clients in different occupations will likely be exposed to different risks.
- Secondly, different clients are likely to have different options available

for risk coping and risk management. For example, wealthier farmers may have more scope to diversify their crop, and households with more children may have more scope to use their children's labour to build precautionary savings.

- Thirdly, different prospective clients are likely to engage differently with the microinsurance product. This may be the case at the point of contractual agreement. For example, more educated prospective clients may take a more nuanced view of the terms of any contract, or different cultural or religious backgrounds may promote different attitudes towards issues of risk and interest accrual. Additionally, this may be the case at the point of insurance payout. For example, different clients may use a microinsurance payout for different purposes, or may face different social expectations to share their gains (*sharing norms*).

For all of these reasons, it is critical that any analysis of microinsurance considers the different ways that different clients are likely to be affected.

*Different clients, different impact:
spell out expected differences
in effects!*

This ought to occur throughout the microinsurance process – from initial planning phases to empirical evaluation and product refinement. Importantly, as the earlier discussion has emphasised, it is not enough for analysis to consider whether different groups are affected differently. In order to better design and adapt a product for client needs, it is vital to consider the means and mechanisms by which such heterogeneity may arise. For example, suppose that larger families are less willing to agree to a microinsurance product. This could be because they have better risk coping and risk management options to rely upon, or because the specific terms of the contract may penalise larger households (for example, if the premium scales linearly with household size), or because they have a larger social network that would expect to share in any insurance payout, or some other reason yet. However, without understanding why such a phenomenon is observed, a microinsurance provider may struggle to adapt and improve its product appropriately.

3.3. Towards a general microinsurance theory of change

The preceding discussion of likely outcomes and impacts provides a set of general proposals for relevant variables of interest: “What we should look for?” This section situates those

proposals within a suggested framework for a general microinsurance theory of change. The framework builds directly upon the theory of change just considered, as well as upon the International Labour Organization’s (ILO) Microinsurance Innovation Facility’s (MIFs) client-value assessment framework and tool called PACE (see box below).⁶ A general microinsurance theory of change might rely upon the following motivating questions.

Input

What is being offered? Analysis of input should focus on evaluating added value of microinsurance in relation to alternatives (i.e., savings, credit, informal groups, social security schemes, etc.) within the four broad categories below:

- **Product** - describes appropriateness by reviewing coverage, benefit level, eligibility criteria, and availability of value-added services
- **Access** - focuses on accessibility and simplicity by investigating choice, enrolment, information, client education, premium payment method, and proximity
- **Cost** - measures both affordability and value for money, whilst looking at additional costs to keep down overall costs of delivery

⁶ See more at www.ilo.org/microinsurance. Matul, Tatin-Jaleran, and Kelly (2012) present the tool and main results from its pilot testing. Matul and Kelly (2012) provide a technical guide on how to do PACE analysis.

- Experience - assesses responsiveness and simplicity by looking at claims procedures and processing time, policy administration, product tangibility, and customer care

Other considerations

- What is the subject matter of the insurance contract?
- In relation to the relevant subject matter, what kinds of specific risks do clients want to have insured?
- What specific events trigger the insurer's liability? Does the set of triggering events adequately cover the main risks related to the relevant subject matter? Does the set cover the kinds of specific risks that clients want covered?
- What remedies are available to the microinsurance provider should the client default upon its payments? Are these formal legal remedies, or informal remedies?
- What remedies are available to the client should the microinsurance provider default upon its payments? Are these formal legal remedies or informal remedies?
- Would a prospective client find the agreement credible? Would the prospective client believe that the microinsurance provider will be willing and able to payout in response to any triggering event?
- What alternative products are available to prospective clients? What strategies for risk coping and risk management are likely?
- Has there been a similar product offered before? Why was it successful or not successful? Is there a general distrust in insurance (e.g., badly designed schemes offered before in the area)?

Output

- What determines client participation? How many clients are likely to participate? What are their likely characteristics?
- Why do they buy or not buy the insurance? (scale)
- What steps can the microinsurance institution take to reduce its costs, whilst still ensuring its ability to meet the terms of the insurance agreement? (costs)
- Do prospective clients clearly understand the terms of the insurance agreement? (use)

Outcome

- What are the effects upon client behaviour?
- How do other actors respond to those effects?
- Relevant considerations for outcomes have been discussed in the earlier section, "Outcomes: What should we look for?"

Impact

- What are the effects upon client well-being?
- Are there spill-over effects for the community as a whole?
- Relevant considerations for impacts have also been discussed in an earlier section, “Impacts: What should we look for?”

Box: The ILO’s Microinsurance Innovation Facility *client-value* framework and the theory of change

The ILO’s Microinsurance Innovation Facility has recently developed a framework for defining and assessing the concept of client-value and the process of creating value in a microinsurance context. The framework is described by Matul, Tatin-Jaleran, and Kelly (2012) as encompassing an iterative process of four distinct steps:

1. Product and process design - “How do products meet client needs in relation to alternatives?”
2. Demand - “What factors influence the choices of low-income households?”
3. Product use - “What is client satisfaction, loyalty and feedback to improve products?”
4. Impact - “To what extent and how microinsurance improves risk management and reduces vulnerability?”

The general theory of change for microinsurance proposed earlier embraces this client-value framework at its foundation. For example, issues of product and process design are highlighted in considering the demands and desires of likely clients, as are likely alternative products and strategies (input). Issues of demand are coincident with the discussion in output, and also overlap the contractual issues covered in input. Product use and impact together cover the discussion of outcome and impact. As mentioned earlier, the PACE tool that the ILO’s Facility developed to analyse and improve client-value can be used to conduct the input evaluation within the suggested theory of change.

Finally, note that the entire client-value framework is designed as an *iterative process*; it is designed so that analysis of product impact feeds back naturally into product and process design. This iterative structure is inherent in the proposed theory of change, which structures a process of impact evaluation; that is, a process designed to allow microinsurance products to be developed, refined and improved over time.

3.4. Defining research questions

This chapter concludes by considering some specific research questions to capture outcomes and impacts discussed earlier. Dercon and Kirchberger (2008, 7) have suggested some important questions for further research, together designed to allow a better evaluation of the *overall impact on household welfare*. The questions were:

- To what extent do low-income households adopt more efficient risk-management strategies when they start using microinsurance?
- How do consumers use insurance payouts?
- Does insurance coverage promote undertaking higher-risk, more productive economic activities?
- Does health insurance contribute to more efficient health seeking behaviours?

The authors went on to acknowledge several subsidiary questions:

Any of these impacts will need to be unpacked further to address questions such as which segments of low-income households benefit the most? What are the intra-household dynamics? How does insurance impact women, men, other household members? How do they

benefit (e.g. is it through more efficient behaviours, stronger asset or human capital position, more asset accumulation, etc.)? Are there any externalities at the community level? For example, does it affect local health care provision, does it crowd out informal schemes, does it affect credit markets? Finally, questions arise about which products provide the highest impact: what is the best product for particular risks in particular circumstances? Can this be related to pricing of products (e.g. low premiums with low protection compared to high premium for higher protection)? Or comparing the impact of single versus composite products (for example, combining health and agriculture insurance products, or mandatory versus voluntary products)? Product comparison should not limit itself to insurance alone: a key concern when studying the impact of insurance will have to be more work on comparing the impact of insurance with other complementary financial services (such as savings, consumption or emergency credit) as well as safety nets and social protection (including social security and cash transfers) (Dercon and Kirchberger 2008, 7).

This concludes the discussion of what is being evaluated and how to develop a framework of theory of change.

References

Dercon, S., M. Kirchberger, J. W. Gunning, and J. P. Platteau. 2008. Literature review on micro-insurance. *Microinsurance paper no. 1*. Geneva: International Labour Organization Microinsurance Innovation Facility.
<http://www.ilo.org/public/english/employment/mifacility/download/litreview.pdf>

Kahneman, D. and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47:263-91.

Matul, M. and E. Kelly. 2012. How to conduct a PACE client value assessment: A technical guide for microinsurance practitioners. Geneva: Microinsurance Innovation Facility.

Matul, M., C. Tatin-Jaleran, and E. Kelly. 2012. Improving client value: insights from India, Kenya and the Philippines. In: *Protecting the poor: A microinsurance compendium*. Vol. II, ed. C. Churchill and M. Matul, 300-330. Geneva: International Labour Organization.

Organisation for Economic Co-operation and Development (OECD). 2002. Glossary of key terms in evaluation and results based management, from the Development Assistance Committee (DAC) Working Party on Aid Evaluation. <http://www.oecd.org/dac/evaluation/2754804.pdf>

Radermacher, R., H. McGowan, and S. Dercon. 2012. What is the impact of microinsurance? In: *Protecting the poor: A microinsurance compendium*. Vol. II, ed. C. Churchill and M. Matul, 59-82. Geneva: International Labour Organization.

Tversky, A. and D. Kahneman. 1992. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5:297-323.

What can we learn from impact assessments?

Jonathan Bauchet, Aparna Dalal, and Jonathan Morduch



4.1. Introduction

How can we determine that an intervention is making a real difference?

At age 40, Feizal was supporting his family in rural northern India. He earned a living selling aluminium pots, which he strapped on to his bicycle and took from village to village. The sales provided the lion's share of his household's US\$36 average monthly income. But one day Feizal had a bad fall from his bicycle and broke his leg. Initially, he relied on the care given by traditional doctors – at a cost of US\$33. After three months, the leg showed no improvement, but Feizal's family could not afford modern treatments. It took several more months, and the resources of his extended family, to pull together US\$250 to pay for a hospital treatment. The family had to draw on a wage advance from Feizal's son's employer and deplete the family's savings, which had been reserved for Feizal's daughter's wedding. In the end, Feizal's leg was treated in a modern hospital and he recovered. But he had spent eight months with no income, his family's savings were gone, and the family was US\$100 in debt.¹

What would Feizal's situation have been if he had access to health insurance? Would Feizal have gone to a modern doctor sooner, thereby receiving better treatment and minimizing recovery time? Could proceeds from an insurance policy have helped him avoid falling into debt? Would Feizal have been able to protect his family's consumption levels? These questions are at the heart of impact evaluations.

With certain assumptions, evaluators can establish that the difference between Feizal's situation and that of insured individuals was *caused by* having a microinsurance policy, i.e., it is the impact of microinsurance. The rough notion of *making a difference* can be translated into a precise question that should be at the heart of every impact study: "How have outcomes changed with the intervention *relative to what would have occurred without the intervention*?"²

What would have occurred without the intervention?

1 The story of Feizal and his family is part of the financial diaries project collected by Orlanda Ruthven and described in *Portfolios of the Poor* (Collins et al. 2009), in which authors provide a window into the financial lives of 250 households in India, Bangladesh, and South Africa.

2 This approach to impact evaluation, based on a comparison with hypothetical outcomes, is often called the Rubin Causal Model after its originator, Donald Rubin, a statistician at Harvard. See Rubin (1974) for the very origins of this model.

Yet, even if that question could be answered in Feizal's community in northern India, the particular answer would likely be different elsewhere—and those other places might be of even greater interest to policy makers or investors. The question about what we can learn from one place to understand another place – in technical jargon, the problem is *external validity* – deserves much more attention, and we return to it at the end of this chapter. But since we can't learn from *any* place unless a study is credible and free from important biases, we devote most of our attention to the problem known as *internal validity*: Are we *actually* measuring what we want to?

4.2. What quantitative impact assessments measure

Impact evaluations try to measure and understand the change in a participant's life that occurred because of an intervention. The *intervention* could be a policy, a project, an insurance product, or a specific feature of a product. For instance, the intervention could relate to a particular product feature, such as the extent of coverage, a change of pricing structure, or variations in the distribution channel.

Impact evaluations focus on the *outcomes evaluations*. Process evaluations measure programme processes, inputs and outputs. They answer

questions like: How many insurance education sessions did trainers conduct? How many farmers attended the sessions? How many households purchased a given insurance policy? These indicators are a measure of the efficiency of the intervention. Whilst they are useful in estimating the outreach of the programme, they give little information about how the programme affected household well-being.

Impact evaluations, on the other hand, look to answer questions like: Did farmers invest in high-grade seeds because they had insurance coverage? Did the change in investment result in higher income? Did insurance make a notable difference in coping with a drought?³ Sometimes the likely answers seem obvious, but well-designed evaluations have the power to surprise. During droughts, for example, families often get help from neighbours or relatives, and sometimes from the state. Families borrow, draw from their own savings, and many even migrate. The impact of insurance might be large for some families but not for others.

The focus on outcomes and impacts, rather than inputs and outputs,

3 These questions illustrate individual- or household-level outcomes. Possible indicators for such outcomes include income, asset ownership, nutrition, education levels, health status, or the cost of medical treatment. Micro-insurance impact evaluations can also focus on institution-level outcomes.

distinguishes impact evaluations from process and impacts of the intervention.

Impact evaluations are critical in quantifying the intervention's true value and understanding the underlying mechanisms.

This chapter is mostly focused on quantitative impact evaluations, estimating the amount of change caused by an intervention for a population of interest. Qualitative impact evaluations are also used in some settings and are particularly useful for gaining insight into *how* interventions generate impacts. They proceed from a different logic, however, and it is beyond the scope of this chapter to explain them in detail.

4.2.1. The greatest challenge of quantitative impact evaluations: addressing selection bias

The ultimate goal of quantitative impact evaluations, and their greatest challenge, is to establish credibly that the intervention *caused* a difference in the lives of the participants.⁴

The challenge is to separate the change that was caused by the intervention from the change that would have happened anyway without the intervention.

In other words, how can the evaluator establish that the outcomes have been caused by the intervention, and not by other concurrent events, underlying trends, or characteristics of the participants? For instance, evidence shows that richer and more educated households are more likely to sign up for health microinsurance (Giné and Yang 2007). In other words, these households *selected* themselves into this intervention. If insured households happen to have good health outcomes, is it due to the insurance itself, or to their capacity to afford better hospitals and better understand doctors' recommendations, even without health insurance? These characteristics can have a significant effect on impact estimates. In a study of microfinance clients, for example, McKernan (2002) found that not isolating the effect of microcredit from other concurrent effects can lead researchers to overestimate the impact by 100%. When measuring impact in microinsurance, not separating the impact of the intervention from that of other confounding factors could lead to underestimating or overestimating the impact of the insurance, depending on the situation. It is possible, for instance, that households that suffer from a preexisting illness are more likely to buy health insurance - the classic *adverse selection* problem.⁵

4 Impact evaluations can be improved by exploring how differences were caused, sometimes by adopting a mixed method approach and conducting qualitative research in parallel to the quantitative evaluation.

5 For further discussion of the adverse selection problem, see Churchill (2006) in the context of microinsurance or Armendáriz and Morduch (2010) in the context of microfinance generally.

In this case, an impact evaluation that compares health outcomes of these households with another set of households could find a (mistakenly) negative impact of microinsurance on health condition since the insured individuals appear to be in worse health than uninsured patients. On the other hand, richer farmers are more likely to have enough disposable income to contract a rainfall microinsurance product, so comparing these farmers to those who did not sign up for insurance will make it seem like the insurance was successful at helping them deal with a drought, when part of the measured impact is, in fact, due to the better initial situation of the insured farmers.

This potential bias in the results is called *selection bias*. Disentangling the influence of individuals' characteristics from that of the intervention (i.e., addressing the selection bias) is surprisingly difficult to do. Some individual characteristics can be observed, measured, and controlled in a statistical analysis.

For instance, gender, age, and residence location are likely to influence both the decision to contract insurance and the outcome from having insurance. Most of these kinds of factors are easy to measure and their influence on the outcomes can be separated out by statistical means.

The big challenge arises with unobservable factors. Attributes like an individual's propensity to fall sick, organisational ability, or access to social networks, are far harder – and others are impossible – to measure. But they can create big biases. Not all hope is lost: certain evaluation methodologies make it possible to recover the net impact of the intervention, free of selection bias. We highlight their principles in the next section.

Impact evaluations of microinsurance present a specific challenge. Whilst the most fundamental benefit of insurance is that it offers households protection when facing emergencies, having access to insurance can also provide important benefits in the *absence* of

Individuals' characteristics can influence results and cause bias

Counterfactual: what would have happened without microinsurance?

adverse events. Economic theory suggests, for example, that having insurance can allow farmers to take more risk, altering their investment and employment strategies, which could have an impact on their well-being. Impact evaluations of microinsurance that capture the impact of these altered strategies can help us understand the full effect of the access to insurance. This is not a methodological consideration, but is a practical challenge when designing microinsurance impact evaluations.

4.3. Getting credible answers (internal validity)

4.3.1. Control groups are essential

Whilst it might seem that researchers would spend most of their time trying to capture what happens when people have insurance, they end up spending even more time trying to capture what happens when they *do not* have insurance. This is the *counterfactual*, and it is the key to credible evaluations. The

question is: What would have happened to the participants had they not received the intervention?

Unfortunately, we cannot ever know what would have actually happened to an insured individual had she not had access to insurance because people can only be in one circumstance at a time. But with the right design, an impact evaluation can form a credible estimate of the counterfactual for a group of participants taken together.

The counterfactual is usually estimated by measuring impacts for individuals who do not participate in the intervention, but who are similar to those who do, in as many respects as possible. The group of individuals who participate in the intervention is commonly referred to as the *treatment group*, and the group of those not participating is referred to as the *control group* (Shadish et al. 2002).⁶

⁶ The methodology has roots in the medical procedures used to test the effectiveness of drugs.

Once treatment and control groups are formed correctly, the quantitative impact of the intervention is measured by comparing outcomes in the treatment and control groups. Statistical techniques are typically used to increase the confidence that the results are not spurious – that is, that they would also be likely to occur in other samples. The difference in outcomes between the two groups is a good measure of the causal impact of the programme if, and only if, the groups are truly comparable.

To demonstrate causal impact we need truly comparable treatment and control groups

4.3.2. Control groups need to be truly comparable

Having a control group is a necessary, but not a sufficient condition, to eliminate the selection bias. The way that treatment and control groups are constituted is fundamental. Whilst having a control group eliminates the effect of general trends such as national macro-economic conditions on the measured impact, it does not necessarily eliminate the influence of individuals' characteristics. In fact, the selection bias will always exist when individuals are allowed to self-select to participate in the intervention. This is because their observed and unobserved characteristics influence both their decision to participate and their outcomes. To use the same example as above, richer households will be more likely to be able to afford insurance premiums as

well as to cope with unexpected financial shocks.

The only sure way to eliminate the selection bias is to let an event or rule external to the intervention – an *exogenous* event or rule – determine who participates in the intervention. In this situation, individuals are *assigned* to the treatment and control groups, they do not form the groups themselves.

Two main types of exogenous events have been used by evaluators. Firstly, a national or state policy that changes access to programmes. For example, the adoption of a national policy of large-scale school building has been used to estimate the impact of education on wages (Duflo 2001). In this methodology, receiving more education was decided by the Indonesia

legislature, and, as such, was not related to any personal characteristic of the students benefitting from the policy. We are not aware of any such study on microinsurance. Secondly, an exogenous event could be a lottery that decides who receives microinsurance and who cannot. Here again, the assignment to treatment and control groups is not related to characteristics of the households or individuals who are being insured – it is random.⁷ This randomised controlled trial methodology has become a gold standard in quantitative impact evaluation (Bauchet and Morduch 2010), and can be used to estimate the impact of microinsurance.

In addition to these types of events, in some cases, exogenous rules can be used to eliminate the influence of participants' characteristics on measured impacts. In regression discontinuity designs, for example, an eligibility rule with a clear cutoff point can be used to create credible inference.

Some microfinance institutions in Bangladesh, for example, had a rule that they only served households owning under a half-acre of land. A potential study design is to compare the outcomes of households just below the half-acre cutoff (who thus get access to the treatment) to households just above the cutoff. This method requires

additional assumptions. The most critical assumption is that participants ranked just above and just below the cutoff are similar in observable and unobservable characteristics.

4.3.3. Control groups need to not have access to the intervention during the evaluation

Whilst great care must be devoted to creating truly comparable groups, even greater care is often necessary to maintain the separation of the treatment and control groups over the course of the evaluation. Obviously, allowing participants to switch group would reintroduce the selection bias that the initial exogenous assignment aimed to eliminate. However, more subtle threats exist.

Households in the control group, for example, might have opportunities to sign up for microinsurance products, maybe from a competing insurer. These households self-select to purchase insurance, which, in addition to reducing the estimated impact, would reintroduce a selection bias.

Households in the control group might also be acquainted with households in the treatment group, and benefit from their relationships, for example, through help in times of need. This spill-over of benefits from the treatment to the control group *contaminates* the assignment.

⁷ See Bauchet and Morduch (2010) for a more detailed exposition of the theory.

The threat of spill-overs can be mitigated by implementing specific designs. The level of assignment is the single most powerful way to address the threat of contamination. Rather than assigning individuals to treatment and control groups, families, households, or entire communities can be assigned to each group. In evaluations of microinsurance, for instance, members of the same family should be assigned together to either treatment or control. At the village level, weather insurance could encourage some farmers to adopt riskier and more productive crops and techniques, which in turn would have positive impacts on the entire community. Recognising this possibility, the evaluator might need to assign entire communities to treatment or control.⁸

4.3.4. Studies need to be big enough to reveal the impact of uncommon events (power)

Since asking all clients how the insurance affected them is generally too costly, a sample of clients is surveyed and statistical methods are used to determine whether conclusions based on the sample can be generalised to all clients.

How big of a sample is needed? This question is particularly important for

studies investigating risk, such as microinsurance impact evaluations, because most events are uncommon. To observe the effect that microinsurance has on households' ability to cope with adverse events, a sufficient number of these events need to happen in both the treatment and control groups.

The need for big samples also arises from the presence of *noise* in all measurements, due to natural variations in the data and measurement errors. This noise might even be particularly loud when measuring outcomes and indicators of social processes, such as risk, vulnerability, or income. But with a large enough sample, the impact of noise can usually be minimised and the impacts of interventions emerge clearly. If the sample is too small, the noise may mask the intervention's real effects: measured impacts may be positive and large, but conventional measures of statistical significance would not be able to establish that the measured impacts are nothing other than noise. Intuitively, the larger the sample, the more confident one is that findings based on that sample are valid for all clients. But, when data collection is required, large samples can be expensive. Evaluators are always trying to balance sample size with budget constraints.

The statistical concept of *power* refers to the ability to detect the impact of an

⁸ See Bauchet and Morduch [2010] for more details on the level of assignment and other responses to spill-overs.

intervention with statistical methods. Power calculations are used to determine the sample size that is required to detect the programme's effect.⁹ Statistical power generally improves with larger sample sizes, but it is not as simple as that. The design of the evaluation matters as well. Power calculations are based on the size and variation of the impact, the size of the sample that is used to measure the impact, and the desired level of statistical significance.

The important point is that impact evaluations need to consider sample size issues seriously and carefully to ensure that the study is able to capture the impacts whilst keeping budgets under control.

This section has emphasised the need to adopt rigorous evaluation designs, based on exogenous assignment to treatment and control groups, to estimate the causal impact of an intervention in an unbiased manner. Achieving a high degree of internal validity is necessary for all impact evaluations, and influences the way findings can be interpreted.

4.4. Interpreting results and understanding the change

Even in evaluations using the design that best establishes internal validity (i.e., the degree of confidence that impacts are caused by the intervention), interpreting results requires stepping back and critically considering the evidence. At least four broad questions should be asked: What impacts are being measured? How did these impacts come to be? How cost effective is the intervention at producing these impacts? How confidently can the evidence from one evaluation be extrapolated to other contexts?

Questions of what intermediate steps have led to these impacts or what pathways theoretically underlie these changes are important as well. As we will explain in detail below, it is important to understand the underlying theory of change, in particular when programmes are planned to be transferred to other contexts. Therefore, researchers are not only interested in quantifying the impact, but also in knowing why and how the impact occurred. To get at these kinds of questions, qualitative or participative studies can help probe the underlying mechanisms.

⁹ This section focuses on how power calculations are used to determine a sample size pre-study. Power calculations can also be used post-study to estimate the level of power obtained with a given sample size (see Bauchet and Morduch [2010] and Duflo et al. [2008] for more technical introductions and references).

Put microinsurance
into context

4.4.1. Interventions are not implemented on a blank slate

Books such as *The Poor and Their Money* (Rutherford 2000) and *Portfolios of the Poor* (Collins et al. 2009) reveal how active the financial lives of poor households are. Households use an array of formal and informal saving, loan, and insurance products and maintain financial relationships with their friends and relatives in order to manage their irregular income, finance large expenditures, and smooth consumption.

Whether evaluating microinsurance as a whole or a specific feature of a microinsurance product, evaluators must carefully define their intervention and place it in a larger context. The impact of introducing a microinsurance product in a new market, for example, is a marginal impact, that is the impact of *adding* the product to the mix of informal mechanisms and formal products already available to households. These include the informal strategies described by Collins et al. (2009), as well as insurance products offered by semiformal organisations such as microfinance institutions,

social insurance schemes offered by the government, and interventions that other insurers or their partners might be implementing. In most cases, the counterfactual is not the absence of insurance mechanisms. The new insurance product will most likely supplement rather than replace the strategies previously used by households. Thus, the challenge is to parse the net impact of the new product, and, ideally, document its complementarities and exclusivities with existing strategies. Interpreting impact estimates accurately requires an understanding of the intervention's precise effect.

4.4.2. The intervention needs time to produce impacts, but long-term impacts are more difficult to attribute to the intervention

Most impact evaluations measure the outcomes of the intervention one or two years after it was implemented. These (relatively) short-term impacts might be smaller than, or different from, expected impacts, which often require time and multiple exposures to the intervention to emerge. In microfinance, for example, borrowers may



not experience improvements in their business and livelihood until after they complete several loan cycles. Similarly, the impact of microinsurance might appear long after households have contracted their first insurance product: households might not adopt new crops, for example, until they have personally benefitted from rainfall insurance during a drought.

Moreover, interest in the results of the evaluation is often high, particularly when the intervention is popular or seems promising, and policy makers and businesses often can't wait. Estimating short-term impacts satisfies a rightful desire to learn how policy and programmes can be improved. But budget permitting, additional surveys should be conducted to estimate both short- and long-term impacts.

Ideally, researchers would wait three to five years before measuring the impact of an intervention. In some cases, more time might even be necessary.

The intervention needs time to have an impact

Waiting that long, however, makes it very difficult to maintain the separation of treatment and control groups and prevent spill-overs, which is a fundamental requirement to be able to claim that the intervention caused the observed impacts. In addition, the risk of attrition, i.e., the drop out of participants, is higher the longer researchers wait before following up to measure post-intervention outcomes. At the least, attrition requires the evaluation to be initiated with a larger sample, but it can also introduce bias in the estimate of impact if participants with specific characteristics drop out of the study.

4.4.3. The impact of access may be as important as the impact of use

Evaluators must decide whether they plan to estimate the impact of *access* to microinsurance, or the *use* of microinsurance when designing the evaluation. This choice influences the design of the evaluation and determines how findings should be interpreted.

Insurers, for example, are certainly interested in evaluating the impact that their products have on the well-being of households who sign up for them. Concern for insured households' well-being and good business practices would also recommend evaluating the impact of adding or modifying specific features of insurance products on

households who use these products. Finally, evaluating the impact of using specific products or features is fundamental, since, if they are not effective amongst those who use them, they should probably not be offered, at least in their current form.

Policy makers and funders, however, are also particularly interested in the impact of offering an intervention such as microinsurance, knowing that not all households who are eligible for it will use it. Many policies, particularly aiming to promote development in a broad sense, are interventions offered to individuals who are not required to participate. The impact of having access to the intervention, rather than actually using it, is, therefore, more relevant to policy makers deciding on which policy to support, or how to improve a given policy.

Evaluations must be designed specifically to measure the impact of access to, or use of, the intervention.¹⁰ Intuitively, the method of assignment to treatment or control groups must mirror the type of impacts in which the evaluator is interested. To measure the impact of *access to* an intervention, the treatment group must be constituted of individuals who are exogenously given access to the intervention. Some

of these households will decide not to participate. To measure the impact of *using* an intervention, the treatment group must be made of individuals assigned to use the intervention.

The impact of having access to the intervention is typically lower than the impact of using the intervention, since some of those offered the intervention do not participate, but must still be considered part of the treatment group. The treatment group here is constituted from those having access to the intervention (i.e., being offered to participate), regardless of whether they use it or not. In an evaluation of the impact of access to an intervention, the internal validity provided by a natural or random experiment does not extend to comparing only those who use the intervention, since individuals or households can choose whether to participate or not. That reintroduces a selection bias.

4.4.4. The distribution of impacts can be (at least) as important as the average impact

Impact evaluations, particularly ones based on exogenous assignment into treatment and control groups, are typically designed to determine the average impact of a programme.¹¹ In many

10 Being able to decide which type of impacts is measured is most common in randomised experiments. In many natural experiments, the exploitable source of exogeneity dictates which type of impacts is measured.

11 The theory underlying the rigorously of randomised evaluations applies to a comparison of average outcomes in the treatment and control groups, but does not extend to comparison of medians or other measures of distribution such as percentiles.

cases, however, organisations care about the distributional impacts of an intervention and not just the average impact.

Imagine an organisation which offers a microinsurance product to a randomly-selected group of households and temporarily denies access to the product to another group. The first group is the treatment group and the second is the control group. The difference between the average outcome of the treated group and the average outcome of the control group is an accurate estimate of the intervention's average impact (see, notably, Bauchet and Morduch 2010 and Duflo et al. 2008 for technical details). This is the causal impact of the microinsurance programme. The average impact is an important parameter, and is often what social investors and practitioners want to know.

But what if half of the treated population gains by 100 percent, and the other half lose by 100 percent? In this case, the average impact is zero. Zero is a clean estimate, but it hides the action. Thus, practitioners and investors might care about who is gaining and losing, so that they can target the programmes appropriately.

A clean estimate of impact for specific subgroups can be estimated through clever designs. For example, stratifying the treatment and control groups by gender allows one to estimate the impact for men and for women separately. Stratifying means dividing the sample along one or more observable characteristics (such as gender), and performing the assignment to treatment and control for each subgroup separately rather than for the entire sample at once. One limitation of this method is that dividing into subgroups generally requires a larger sample. To have the greatest credibility, subgroups should be identified before the evaluation is started (based on expectations of the way that impacts are likely to vary in different parts of the population) and built into the survey design.

*Stratifying
treatment and
control groups to
measure impact
for subgroups*

4.4.5. Cost-benefit calculations are critical companions of impact evaluations

Well-designed impact evaluations will provide evidence about the difference that a given intervention makes in the lives of people and/or the performance of an organisation. Knowing the impact of a specific intervention is not the only guide for future action, however.

The costs of producing such impact must be factored in recommendations for replication or scaling-up.

Cost-benefit analyses are widely used tools of public policy and should also be systematic companions to impact evaluation. They allow policy makers, funders, and implementing organisations to compare different interventions, or different features of an intervention, and implement the one(s) that provide the best “bang for the buck”. For example, health microinsurers might want to know whether establishing a cashless payment system would provide additional benefits compared to the current mechanism that reimburses for health expenses incurred. The impact on both households and on the insurer of the change in coverage needs to be evaluated and compared to the increased (or decreased) costs for both insured patients and the insuring organisation.

Generalising results

4.5. Generalising from one place and time to another (external validity)

Most impact evaluations aim to improve the understanding of *what works*, both to determine whether the investments have been effective and to learn about possibilities for other places. The ability to generalise the findings from an evaluation is called *external validity*.

Learning from one context to another requires both external validity and internal validity. Some statistics-based evaluations exploit data coming from large geographical areas, varied contexts, and/or diversified populations, so their conclusions may be applicable to a wide range of situations (high external validity). But if those studies lack an exogenous determinant of participation into the intervention, such as a random assignment, they may perform less well in providing unbiased estimates of impact (low internal validity). It is then difficult to draw clear lessons.



Evaluations based on random assignment into treatment and control groups, on the other hand, do have high internal validity, but they are nearly always implemented with a specific partner in a particular context, which can reduce confidence that measured impacts would also extend to a different setting. For instance, a randomised evaluation of flip charts as teacher's aides in schools in Kenya (Glewwe et al. 2004) only tells us whether the flip charts helped raise test scores for these students in these schools in this region of Kenya. One could imagine that students or schools in other parts of Kenya, India, or Peru have different educational needs, and would benefit differently (or not at all) from their teachers' using flip charts.

We need to understand the specific context of the evaluation before drawing general conclusions. This means considering three big questions:

1. How does the population studied *there* differ from the population I'm interested in *here*? Are they better educated? Poorer? Healthier? Etc.
2. How do supporting inputs differ? Are there critical government programmes in place? Good roads and transport? Community institutions? Effective organisations to deliver the interventions in question?
3. How do alternative activities differ? Does the studied intervention mostly substitute for existing opportunities? Does it complement them? Morduch et al. (2013), for example, found that a very promising antipoverty programme in South India ended up having no net impact because alternative options were so good (and the control group availed themselves of those options). The same programme had bigger impacts in sites with very similar populations, but where, it seems, such good alternatives were lacking.

Some of these questions can be addressed with an eye to understanding how and why the intervention worked or not. Combining qualitative and participatory research designs with rigorous quantitative evaluations can be applied to increase the understanding of the mechanisms that produced impacts and to gain external validity.

4.6. Conclusion: Using evaluations to improve operations

It is tempting to view evaluations as mainly backward-looking assessments. But their greatest power is often as forward-looking guides to innovation and improvement. Businesses, donors, investors, and policy

makers often have to select between competing programmes when deciding how to allocate scarce resources. Rigorous impact evaluations are an indispensable tool for strategic planning. They inform choices that leaders must make. Knowing what difference a specific intervention makes also calls upon all stakeholders to improve the intervention, try alternative – and potentially better or cheaper – methods, and share the knowledge gained with other individuals and organisations.

Karlan et al. (2009), for example, identify several ways in which rigorous impact evaluations can help microfinance institutions increase both their sustainability and social outreach, including improving their borrower risk assessment techniques and learning about the impact of the price of the loans on demand. In microinsurance, impact evaluations can test the effectiveness of two different insurance products or test the effect of specific elements of the products, such as different marketing techniques, pricing structures, or distribution channels. Understanding the impact of their operations on client participation and well-being can enable practitioners to design better products and services, and thereby increase scale, sustainability, and social impact.

This is an exciting time for the microinsurance industry. The past few years

have seen an influx of interest from insurers and investors, and regulators are driving new initiatives to broaden access. As organisations make new investments and test innovations, they should pay attention to whether their products are having the impacts for which they hoped. When done right, impact evaluations are a tool to efficiently direct future allocations design better products, and improve operations.

Impact evaluations can be a tool to:

- efficiently direct future allocations
- design better products
- improve operations

Suggestions for further reading

Bauchet, J., and J. Morduch. 2010. An introduction to impact evaluations with randomized designs. A Research Framing Note for the Financial Access Initiative, New York Univ.

Collins, D., J. Morduch, S. Rutherford, and O. Ruthven. 2009. *Portfolios of the Poor: How the world's poor live on \$2 a day*. Princeton, NJ, and Oxford, UK: Princeton University Press.

Duflo, E., R. Glennerster, and M. Kremer. 2008. Using randomization in development economics research: A toolkit. In: *Handbook of Development*. Vol.4, ed. T. P. Schultz and J. Strauss, 3895-3962. Amsterdam and New York: North-Holland.

Karlan, D., N. Goldberg, and J. Copestake. 2009. Randomized control trials are the best way to measure impact of microfinance programs and improve microfinance product designs. *Enterprise Development and Microfinance* 20(3): 167-176.

Karlan, D., J. Morduch, and S. Mullainathan. 2010. Take-up: Why microfinance take-up rates are low & why it matters. A Research Framing Note for the Financial Access Initiative, New York Univ.

References

Armendáriz, B., and J. Morduch. 2010. *The economics of microfinance*. 2nd ed. Cambridge, MA: MIT Press.

Bauchet, J., and J. Morduch. 2010. An introduction to impact evaluations with randomized designs. A Research Framing Note for the Financial Access Initiative, New York Univ.

Churchill, C., ed. 2006. *Protecting the Poor: A Microinsurance Compendium*. Geneva: International Labour Office.

Collins, D., J. Morduch, S. Rutherford, and O. Ruthven. 2009. *Portfolios of the Poor: How the world's poor live on \$2 a day*. Princeton, NJ, and Oxford, UK: Princeton University Press.

Duflo, E. 2001. Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review* 91(4): 795-813.

Duflo, E., R. Glennerster, and M. Kremer. 2008. Using randomization in development economics research: A toolkit. In: *Handbook of Development*. Vol.4, ed. T. P. Schultz and J. Strauss, 3895-3962. Amsterdam and New York: North-Holland.

Giné, X. and D. Yang. 2009. Insurance, credit, and technology adoption: Field experimental evidence from Malawi. *Journal of Development Economics* 89(2009): 1-11.

Glewwe, P., M. Kremer, S. Moulin, and E. Zitzewitz. 2004. Retrospective vs. prospective analyses of school inputs: the case of flip charts in Kenya. *Journal of Development Economics* 74(1): 251-268.

Karlan, D., N. Goldberg, and J. Copestake. 2009. Randomized control trials are the best way to measure impact of microfinance programs and improve microfinance product designs. *Enterprise Development and Microfinance* 20(3): 167-176.

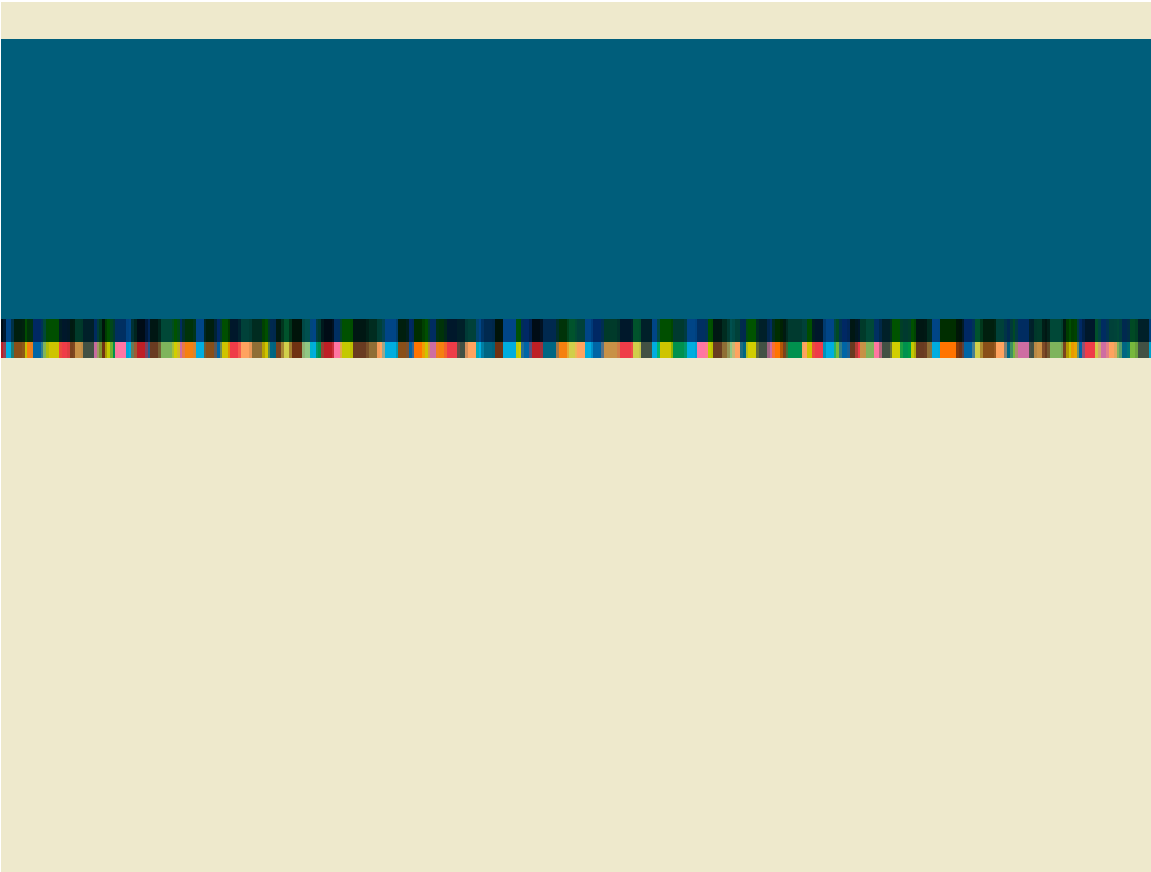
McKernan, S.M. 2002. The impact of micro-credit programs on self-employment profits: Do noncredit program aspects matter? *Review of Economics and Statistics* 84(1): 93-115.

Morduch, J., S. Ravi, and J. Bauchet. 2013. Failure vs. displacement: Why an innovative anti-poverty program showed no net impact. PRIMCED discussion paper, series no.32, Institute of Economic Research, Hitotsubashi University.

Rubin, D. B. 1974. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 66(5): 688-701.

Rutherford, S. 2000. *The poor and their money*. New Delhi and New York: Oxford University Press.

Shadish W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.





Approaches to Impact Assessments in Microinsurance



Experimental designs

Tomoko Harigaya and Dean Karlan



5.1. Introduction: What is an experimental design?

Does expanded access to health insurance improve the health status of the poor? How does weather insurance affect farm investment and output amongst maize farmers? These are typical questions asked in impact assessments. Implicit in these questions is the identification of the causal effect of a microinsurance programme. We are interested in determining how the health status of the poor changed because of access to health insurance, and how much farm investment and output increased amongst maize farmers because of the weather insurance. How can we best answer these questions and isolate the causal impact of a programme? In this chapter, we introduce an evaluation method that uses the random assignment of an intervention. Programme impact is measured by comparing outcomes between the intervention and non-intervention groups.

The main objective of any impact assessment is to isolate the causal effect of a programme. Many microinsurance programmes track the changes in the well-being of programme participants over time and call this change “programme impact”. However, many factors other than access to microinsurance affect the lives of programme participants

during the intervention. New economic policies may be implemented, more non-governmental organisations (NGOs) may enter in the area, or bad weather may affect farm outputs. Thus, this method of mere comparison of programme participants before and after the intervention is unreliable in determining the true impact of the programme. If it could, then we would find aging as an impact of access to microinsurance, since people were younger before getting microinsurance than afterwards!

Hence, to evaluate impact, we must compare programme participants to non-participants, not merely follow participants over time. Comparison with any non-participants, however, would also be insufficient since self-selected programme participants are inherently different from non-participants. They may differ in terms of their personality, eligibility, or risk attitude. If we observed differences between programme participants and non-participants over time, it would still be unknown whether those differences were caused by the microinsurance programme itself or other factors that led them to participate or not to participate in the programme. This is referred to as the attribution problem.

The experimental design offers a simple solution to the attribution problem:

it constructs a valid comparison group through a random assignment of an intervention. This chapter discusses why and how to evaluate microinsurance schemes using randomised controlled trials (RCTs), also called randomised evaluations (REs). In the simplest form of an RCT, we construct a comparison group that looks just like the target population of the programme before the intervention takes place. This is done by randomly assigning individuals from the target sample to either a treatment or control group. Randomisation ensures that the group that receives an intervention (the “treatment group”) and the group that does not (the “control group”) have comparable characteristics; therefore, we can confidently conclude that the difference in the outcome observed is the causal effect of the intervention.

In section 2, we discuss different types of evaluations and their purposes. Section 3 explains the key methodological issues in programme evaluation and why RCTs are often treated as the gold standard for evaluating microinsurance schemes. This section also reviews ethical considerations in designing evaluations. In section 4, we discuss the process of designing and implementing an RCT. Finally, section 5 summarises the key points and discusses external validity. Note that this chapter will not elaborate on the statistical methods and econometrics of

experiments, but will rather focus on the conceptual framework and highlight issues that need to be considered when planning evaluations.

5.2. Why do we evaluate microinsurance programmes?

5.2.1. Impact evaluation

Microinsurance offers protection against risks for low-income households. There are good reasons to think that access to insurance will benefit the poor. Firstly, it reduces the vulnerability to negative shocks by at least partially compensating for the economic loss due to uncontrollable events, such as bad weather or illness. This reduction in vulnerability could, consequently, address underinvestment in profitable – but risky – opportunities and increase individuals’ economic capacity. These are lofty goals, of course, and the impact of a particular programme depends on many factors, including the insurance product design, the targeting, and alternative mechanisms to cope with risks, amongst other things.

Microinsurance, like any other intervention, may also have unintended consequences. A primary consideration is simply thinking about its effect on pre-existing informal insurance networks.

Outreach
is not
impact!

Formal insurance may weaken informal safety nets if it causes some individuals to withdraw from informal insurance arrangements. In addition, insurance products may simply replace other household insurance mechanisms, such as savings. Thus, whilst microinsurance may increase individuals' economic capacity, its net effect is ambiguous. Rigorous evaluations can measure the exact impact of an insurance scheme on various aspects of the lives of the target population and assess the overall effect of insurance.

The microinsurance industry and donors should measure social impact to determine whether microinsurance schemes are worthy of subsidy, and, if so, by how much. It is also important for for-profit insurance providers to measure the social impact of microinsurance for several reasons. Firstly, they often benefit from subsidies indirectly, through subsidised microfinance programmes. Many private insurance providers partner with microfinance institutions (MFIs) to distribute insurance policies and compete with subsidised microinsurance programmes. Secondly, private insurance

providers attract social investors' money by claiming that microinsurance has a double bottom line, i.e., that it provides benefits both for the customers (social bottom line), as well as the insurance providers (economic bottom line). Investors should know whether this claim is true. Third, demand for insurance does not necessarily imply that the product is beneficial for consumers. Private insurance providers may argue that meeting the unmet demand for insurance itself is a sign of social impact. However, there are many products in high demand that do not enhance welfare of consumers. For example, despite the steady demand for cigarettes and alcohol, they may not be welfare-enhancing in the long-term. All of these reasons together make a strong case for impact evaluations, regardless of whether the scheme is subsidized or not.

It is important to note that social impact cannot be measured by programme coverage. Practitioners often conflate outreach (i.e., take-up of the insurance programme) with impact. Although outreach is a good intermediate indicator that tells us whether a programme has accomplished its operational

activities, it does not provide insight into the changes in the well-being of the target population. Suppose we observe a high take-up of a health microinsurance scheme marketed by a trusted organisation in the local community. The take-up may be driven by the reputation of the marketing organisation rather than by meeting the needs of the people and mitigating their risks. On the other hand, even if take-up is low, it does not automatically imply that the potential impact of the insurance, had the households purchased the product, is low. They may be credit-constrained and unable to purchase the product. They may simply lack product information. To assess the social impact of the programme, we must look at changes in household well-being beyond intermediate outcomes, such as household expenditures on health and health status.

5.2.2. Other benefits of evaluation

5.2.2.1. Product innovation

Well-designed evaluations tell us not only whether a given microinsurance programme works, but also why it works or does not work.¹ By under-

standing which components of the microinsurance programme work, we are better equipped to modify product designs to increase impact. Knowing why the programme is beneficial also helps us identify settings where the programme can be effectively scaled up. Even if the insurance product is known to be effective, there is always room for product innovation that increases outreach and impact.

An evaluation can also provide insights into how a microinsurance programme can be improved. When there is a new product idea, an evaluation could assess the relative impact of the new and existing products. If there are competing product ideas, we can evaluate all of them concurrently and compare the results to decide which idea works best.

5.2.2.2. Process improvement

Rigorous evaluations can also assess the effectiveness of the implementation of microinsurance schemes. Even the most effective insurance product would fail to bring positive changes in the absence of an effective implementation process.

Marketing and information dissemination is a critical component of any programme implementation. Recent RCTs suggest that this is particularly important for microinsurance

¹ For example, Giné and Yang (2009) evaluated a rainfall insurance with credit offered to maize and groundnut farmers in Malawi and found that the loan bundled with the rainfall insurance saw a lower take-up. The evaluation offers suggestive evidence for why the insurance reduced the take-up: because farmers have limited liability for loans, the insurance premium is effectively an increase in the interest rate.

schemes because many poor households are unfamiliar with the mechanisms of formal insurance. Lack of information on enrolment process, limited trust in insurance providers, and high premiums are found to be significant barriers to microinsurance programmes (Thornton et al. 2010; Cole et al. 2013). The importance of good marketing is also highlighted in a study in rural India, which found that the commission structure for insurance marketing agents distorts the information provided to potential beneficiaries and affects their decisions to purchase insurance (Anagol

et al. 2013). Evaluations on marketing pitches would help us identify the effective strategies to provide consumer protection in under-regulated markets and to deepen outreach.

Finally, the lack of information on product details and claims procedure could not only affect the take-up of the insurance product, but also result in non-usage of the insurance policy conditional on take-up. Providing insurance education on benefit coverage and claims procedures, beyond typical marketing campaigns, may affect the outreach and utilisation.

Box 1: FFH microinsurance education in Ghana (Schultz, et al., 2013)

In 2003, the Ghanaian government passed the National Health Insurance Act. This health insurance scheme was developed to serve as a core strategy for improving access to health care. Since the implementation of the act in March 2004, insurance coverage has quickly expanded: according to the World Health Organization (WHO), 55% of the population had registered in the National Health Insurance Scheme (NHIS) by 2008. However, insurance coverage has been much higher amongst the wealthy than amongst the poor, failing to reduce the disparity in access to health care.

To encourage NHIS expansion in low take-up regions, Freedom from Hunger (FFH) developed an education module that discusses the mechanics of the NHIS insurance scheme and its benefits. If the limited NHIS take-up observed in the Northern region is due to a lack of understanding of the insurance scheme, or general knowledge about health risks, FFH's education programme may effectively increase take-up amongst poor households. In order to measure the impact of the education programme, a randomised controlled evaluation was implemented in 2010 - 2011.

In this evaluation, 275 existing client groups, served by a microfinance institution based in Tamale, were randomly assigned into three groups: 1) a treatment group in which FFH offered the education module in six sessions ("short"); 2) a treatment group in which FFH offered the education module in one consolidated session ("consolidated"), and 3)

a control group in which no insurance education programme was offered (“control”).² The evaluation measured changes in NHIS enrolment and health and financial outcomes over 20 months.

The comparison across treatment and control centers revealed the following:

- Greater knowledge of the insurance registration process in short and consolidated groups than in control groups
- No significant increase in the insurance registration rate
- No change in the likelihood of health care utilisation or health insurance usage

These results suggest that the low take up of microinsurance products is not simply due to the barrier to information on how to enrol and process claims.

Note that some business processes are simpler than others to evaluate in an RCT. Customer-facing business processes, such as marketing materials, marketing incentives, and pricing, can be easily randomised across individual customers or marketing agents. In contrast, core business processes that cannot be randomly varied by individuals (e.g., enrolment and claims procedures) may require a large sample of insurance offices.

5.2.2.3. Demand for insurance

Assessing whether or not a microinsurance programme reaches its target population group is another important question. The poor may have a very different willingness to pay for insurance products than those with a regular flow of income (Siegel et al. 2001). A small increase in the price of the insurance policy might result in a large drop in demand amongst the poor.³

It's important to include business processes in evaluation - but some are difficult to assess

2 The study also tested the effect of offering a reminder session one year after treatment centers received the education modules. See Schultz, Metcalfe, and Gray (2013) for the full description and results of the study.

3 A number of randomised evaluations on pricing of health products and services suggest that the poor are price sensitive to a range of health products, including water chlorines, bed nets, and deworming pills (Kremer and Glennerster 2012).

If the local demand for insurance is price sensitive, the choice of premium would significantly affect the programme adoption and impact. Therefore, understanding the gradient of willingness to pay amongst the target population is useful in expanding outreach. The existing evidence for rainfall index insurance products, for example, suggests that the demand structure varies widely across contexts. On one hand, a study in rural India found the demand for the rainfall insurance to be low and highly price sensitive: even when the premium is set lower than the actuarially fair price, the demand

is way below 50% (Cole et al. 2013). On the other hand, a study in Ghana shows a relatively high demand — 40-50% of farmers purchased the insurance at the actuarially fair price and a 50% subsidy increases the take-up rate to 60% (Karlan et al. 2012).⁴ An evaluation of pricing strategies could help us understand the local demand structure for insurance and determine whether and how much premiums should be subsidised to improve targeting and expand coverage.

⁴ Both of these studies used an RCT design and randomly assigned the premiums during the initial marketing.

Box 2: Price sensitivity for rainfall insurance in Ghana (Karlan et al., 2012)

Underinvestment in agriculture amongst smallholder farmers has been documented around the world. In 2009, researchers set out to study the importance of credit constraint and vulnerability to weather shocks in farmers' investment decisions in rural Ghana. They developed a rainfall index insurance and randomly assigned 1350 maize farmers to receive 1) the rainfall index insurance, 2) capital grant, or 3) the combination of both. Amongst those who received the insurance offer, the prices were randomised from one eighth of the actuarially fair price (8-9.5 GHC/acre or US\$0.2) to the market price (14 GHC/acre). Figure 1 shows the declining demand over the offer price.

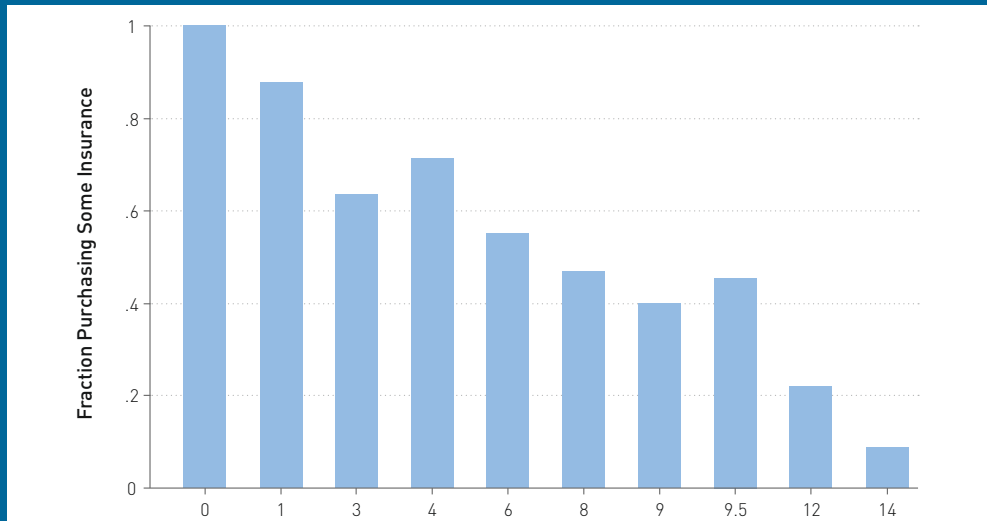
The demand for insurance is high: 42% of farmers purchased the insurance at the actuarially fair price (around 50% discount off of market price) and 67% purchased the insurance at a 75% discount off of market price.

The demand falls quickly over the prices beyond the actuarially fair price: 11% of farmers purchased the insurance at the market price.

Further analysis reveals that the demand curve looks similar for the wealthy vs. poor farmers. This suggests that the lack of liquidity does not explain the steep decline in demand over high prices.

The take-up of the insurance product in the following years shows that the experience of farmers within the close social network is an important determinant of the demand for insurance.

Figure 1: Insurance Takeup by Price per Acre (Cedis)



5.2.3. Take-away messages from this section

- Impact evaluations tell us what changes a microinsurance programme could bring to the target population.
- Knowing the size and extent of the impact of microinsurance programmes is important for all players in the microinsurance industry.
- Evaluations of microinsurance designs and processes help us understand why the programme works (or does not work) and identify opportunities for product and process improvement.

5.3. How do we evaluate?

The fundamental question being asked by an impact evaluation is: *how is the well-being of the target population different in the world with the microinsurance programme compared to that in the same world without the microinsurance programme?* This alternative world is called the *counterfactual* and is something which we do not observe. We cannot monitor the same household both with and without the microinsurance scheme over the same time period. The key to designing a good evaluation is creating the best possible

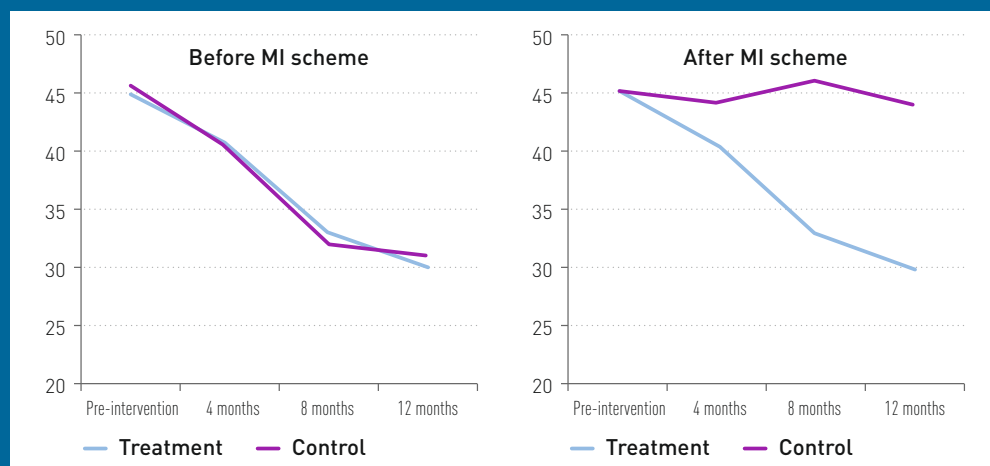
representation of this counterfactual. If we can construct a good representation of the counterfactual, we can infer what *would* have happened to the individuals and communities in the absence of the programme. A randomly assigned control group gets us this counterfactual.

5.3.1. Why do we need a control group?

A well-constructed control group gives us our best assessment of what would have happened in the absence of the programme. We can get a good estimate of programme impact by comparing the treatment and control groups.

Without a control group, we could only observe the outcome measures for the microinsurance programme participants. We could measure outcome changes over time amongst programme participants, but this is problematic because there could be any number of other factors influencing their lives at any point in time. Therefore, we cannot attribute any change observed amongst programme participants to the microinsurance programme. The example below illustrates this point.

Figure 2A & 2B: Change in the proportion of community members with moneylender loans before and after the implementation of a microinsurance (MI) scheme



Example: Suppose that we are interested in measuring the impact of a micro health insurance scheme on the likelihood of borrowing from moneylenders for health-care expenditures. The borrowing rate before the intervention in the sample communities was 45%. Twelve months after the introduction of the insurance scheme, we observed

the borrowing rate of 30% — a substantial reduction in the borrowing rate! Without a control group, we might conclude that the insurance scheme was effective in reducing the use of moneylenders. However, we cannot attribute the change in the borrowing rate to the insurance scheme because there are many other things that happened to the insurance beneficiaries in the same time period and might have influenced the use of moneylenders. The government might have increased its spending on health, which made the health care less expensive. New microfinance programmes might have improved access to credit, reducing the demand for moneylender loans. Or there could have been a steady improvement in health conditions in the communities, which resulted in a lower demand for health care. All of these factors could potentially reduce the borrowing rate. If we only observe the outcome measures for insurance beneficiaries, the effect of the insurance scheme is unclear.

Now, suppose we had a control group of communities that had similar characteristics with the treatment communities but did not receive the microinsurance scheme. The baseline borrowing rate in the control group was also 45%. If we observe a similar decrease in the borrowing rate amongst the control communities over the 12-month period, the 33% decrease in the borrowing rate in the treatment communities must be due to something that affected both treatment and control communities rather than the insurance scheme (figure 2A). On the other hand, if we observe no or only little change in the borrowing rate of the control communities, we are more certain that the insurance scheme caused the reduction in the borrowing rate, given that the only difference between the treatment and control communities was the presence of the microinsurance scheme (figure 2B).

The presence of a valid control group helps us weed out the non-programme effects and isolate the programme impact. Without a control group representing what would have happened in the absence of the programme, we cannot disentangle the effect of the programme from other concurrent factors affecting programme participants.

5.3.2. Why do we randomise?

An ideal control group should have the same observable and unobservable characteristics as the treatment group, with the exception of the presence of the microinsurance scheme. One frequently used approach is to use people who decide not to participate in the insurance scheme as a control

group. However, even if the non-participants have similar demographic characteristics to the programme participants, the very fact that they choose not to participate in the scheme makes non-participants different from the participants. They may be exposed to lower levels of health risks. They may have different attitudes toward health risks and value differently the

Randomisation of treatment distributes observable and unobservable characteristics equally

long-term health benefits. If these potential differences exist between participants and non-participants, we can no longer attribute the observed differences in outcomes between the treatment and control groups to the microinsurance scheme alone.

The advantage of randomised controlled trials is that they can construct groups that have no observed or unobserved differences on average.⁵ When individuals or communities are randomly assigned to treatment and control groups, so long as we have a reasonably large group of people, we can expect the average characteristics of both groups (such as income level, health risks, and attitude and beliefs toward health care) to be similar. We can also expect that the two groups face an equal probability of encountering external events (such as policy changes and natural disasters). Thus, any difference between the two groups is then the presence of the microinsurance scheme. Any difference observed

over time between the groups can be attributed to the insurance scheme.

5.3.3. Ethical considerations

Ethical considerations for evaluations are twofold. If we are uncertain about the true effect of access to insurance on the poor, it could be unethical to scale up the insurance programmes. After all, those are funds that could be used for other interventions that are known to have some benefit, so there is always an ethical trade-off between conducting research to measure impact and using those funds to provide more services. The trade-off depends greatly on how much knowledge we already have about the intervention and the alternatives, and how many resources will be allocated in the future based on the results of the research. The only way to ensure that a programme has the intended impact on the target population is by conducting a rigorous evaluation.

There are several other ethical concerns frequently raised with randomised controlled evaluations:

⁵ Bauchet and Morduch (2010) provide an analytical framework with more technical details on randomisation.

- *Randomly denying access to microinsurance scheme is unfair*

Microinsurance schemes, and development programmes in general, are operating under limited-resource settings. Allocation of public programmes and selection of beneficiaries are often influenced by political factors (e.g., national insurance policies may be more accessible for politically connected individuals). If the available resources permit the programme to reach only a certain number of the poor, randomisation may actually be the fairest approach for selecting beneficiaries. Clearly this depends on the specific context, but one fact is typically true: firms, government and non-profits have limited resources, and an evaluation that embraces the true constraint of the organisation but finds a way to allocate randomly does not lead to a reduction in individuals receiving services. Rather, it just allocates the service in a way that allows for an evaluation. Naturally this is not true when resources or the ability to offer a service are truly unconstrained. In such situations, RCTs typically use *encouragement* designs (see Section 5.4.) in order to avoid having to deny access to a service to any individual who comes forward and asks for it.

- *The need for insurance is urgent: holding back the implementation is unethical*

Holding back the programme implementation is only unethical when its positive impact is known relative to other alternatives, and when further research will not help direct future resources any better.

- *RCTs are costly—the money spent on the evaluation could be used to serve more people*

This is a common misperception. Conducting household surveys is costly, but conducting a randomised controlled trial is not, i.e., the decision to randomise versus not randomise is not a costly one. In fact, randomising assignment is often cheaper than not randomising assignment, holding all else constant. The costs of evaluation range widely depending on the types of outcomes measured, expected effect sizes, and other factors related to the evaluation design. A rigorous evaluation should be considered as an investment for greater impact of future programmes. It is more costly to continue offering an ineffective programme than to invest in a rigorous evaluation and improve programme impact in the long term.

5.3.4. Take-away messages from this section

- To assess the impact of a programme, or a relative impact of two or more programmes, we need

to construct a valid control group that would tell us what would have happened in the absence of the programme.

- The best way to construct a valid control group is to randomly assign individuals, households, or communities to treatment and control groups. By its construction, randomisation ensures that treatment and control groups are statistically comparable.

5.4. Designing randomised controlled trials

A common mistake in designing a programme evaluation is to consider evaluation as an activity that takes place after a programme is completed. Planning an evaluation after programme completion is problematic. Firstly, the remaining participants at the end of the programme are not representative of the target population. Secondly, it

is hard to identify a comparison group that looked just like the programme participants before the intervention. Because these challenges are difficult to address in the post-intervention analysis, the evaluation design needs to be built into the programme design. In this section, we will go through the process of designing a randomised controlled evaluation.

5.4.1. Identifying a problem and potential solutions

Development programmes are typically designed to address market failures. Designs evolve over time depending on the response from potential beneficiaries and implementing agencies. But how do we identify the right programme designs and potential solutions to improve existing programmes?⁶

Existing knowledge of programme design and impact is always a good place to start. There is an increasing number of RCTs on microinsurance programmes. These evaluations that explain why the programme worked could provide insight into key features of an effective product design. The input from the management, the implementation staff, or expert researchers can also help identify problems

*Evaluations
must be planned
before the
intervention!*

⁶ Burns and Dalal (2010) document the process of designing an insurance product in India in the aftermath of tsunami in 2004.

and potential solutions specific to the setting.

In addition, qualitative data can be useful when searching for potential solutions and brainstorming ideas for product innovation. Interviews with non-enrolees could help identify existing barriers to accessing insurance. Focus group discussions with enrolees can help reveal the needs of the potential beneficiaries and what they value about the insurance. Enrolee satisfaction surveys may also point to product features that need reassessment. Note that the qualitative data would not give you clear answers to the questions on the programme impact. However, it provides information on how people perceive the values of insurance and existing products — an important insight for product innovation.

5.4.2. Piloting the intervention ideas

Once a potential solution is identified, the new programme design needs to be tested in a small-scale pilot. The purpose of the pilot test is to ensure the operational feasibility of the new programme. The pilot phase allows the implementer to solve problems with the implementation procedure and modify it as needed so that the full-scale evaluation will not be affected by unexpected operational glitches. When operational problems arise during the

It's best to resolve all operational issues of the insurance before starting the evaluation

full-scale implementation, they may hinder delivery of the programme as designed. As a result, the evaluation may not be able to accurately measure the impact of the programme. It is therefore critical to work out the kinks of the operational details before implementing a full-scale evaluation.

The length of the pilot phase depends on the complexity of the programme implementation and operational target set by the implementing agency. If the implementing agency has a target to achieve a certain level of outreach within a short period of time, extending the small-scale pilot may not be favorable. However, it is important to remember that changing the implementation procedure or programme design after the launch of the full-scale evaluation could affect the scope of the evaluation analysis and limit what can be concluded from the

evaluation. The evaluators and implementing agency should ensure that all operational issues are resolved before moving on to the full-scale evaluation implementation.

5.4.3. Selecting impact indicators

After all the operational issues are resolved in the programme pilot, there are several steps to take before launching the full-scale evaluation. Firstly, the implementing organisation selects a set of impact indicators as main outcome measures. In any development programme, the ultimate goal is to improve the well-being of programme beneficiaries. Ideally, the outcome measures should capture different stages and mechanisms through which the intervention works. By selecting a set of indicators that reflect each of these potential impact mechanisms, the evaluation can assess why a particular insurance product works better than others. Piecing out different mechanisms of the microinsurance scheme can help identify the strong/weak impact channels and areas for product or process innovation in the future.

5.4.4. Selecting a sample frame and sample size

For every RCT it is necessary to select a sample frame — the pool of all potential “objects” to be included in the study

— *before* implementing the programme. These can consist of individuals or an entity of units like households and communities. When selecting the evaluation sample frame, you should consider the following questions: Who is the micro-insurance scheme designed for? Would the impact be different for different population groups (by gender, age, income level, or geographic region)? Is there any interest in measuring the programme impact for a particular segment of population? Donors are often interested in measuring programme impact for the poorest households. Then, you might define the target sample for the evaluation by the poverty status of households and communities.

The final sample size (i.e., the number of observations to be included in both treatment and control group) depends on a number of parameters, including the expected effect sizes (i.e., the magnitude of the expected impact) of main outcome indicators and take-up rate of the insurance product. Statistically, smaller impacts are more difficult to detect. Suppose we observe a 1% difference in the household expenditure in health between treatment and control groups. The probability that this difference is caused by the programme and not by random error or chance would be smaller than when we observe a 10% difference in the same outcome between the same treatment and control groups. To detect a 1% reduction in

household expenditure, the evaluation needs a relatively large sample compared to detecting a 10% reduction. Or to put in other words: If we want to be sure that very small changes we observe have actually been caused by the insurance, we would have to observe the same very small change for lots of people. Similarly, if the insurance take-up is low, the initial sample size needs to be large because the number of people who will experience the direct impact of the insurance scheme is smaller. The sample size also depends on how frequently the outcome event occurs. Suppose the insurance benefits only cover rare events. There would be few people in the sample who face those rare events and have the opportunity to use the insurance policy during the evaluation period. In such case, the evaluation needs a larger sample size for the same reason that the evaluation of a low take-up programme needs a large sample. Once these parameters are estimated, the evaluator can calculate the sample size needed to detect a given effect size using statistical methods. Similarly, for a given sample size and a defined significance criterion, one can calculate the minimum effect size that can be measured. All these processes are called “power calculation”.

5.4.5. Randomisation

Many critics of randomised controlled evaluations argue that randomly

assigning programme offers is often infeasible in reality. It might be politically difficult to randomise access to the government insurance programmes. The insurance policy that covers families cannot be randomised across members within the household. Whilst a randomised approach is certainly not feasible in every setting of development programme evaluation, there are many randomisation techniques that could resolve operational challenges.

5.4.5.1. Encouragement design

In a basic form of RCT, the treatment group receives an intervention and the control group is withheld from receiving the intervention. Under the encouragement design, the treatment group receives an “encouragement” to enrol in a programme. The control group can also participate, but does not receive any encouragement. The examples of an encouragement include additional marketing materials, a visit by a marketing agent, or financial incentives to enrol in a microinsurance scheme. Naturally, we expect the group receiving the encouragement to have a higher enrolment rate in the programme. This difference in the programme enrolment is used to estimate the programme impact. The encouragement design is particularly useful when withholding the intervention from the

control group poses an ethical concern or a logistical difficulty.

5.4.5.2. Phase-in design

When the implementing agency has a constraint in rolling out the programme at once, the randomisation will determine the order of intervention implementation, rather than who receives the intervention. The individuals randomised to receive the intervention later in time serve as a control group. This method is particularly useful when evaluating a nationwide or region-wide programme where withholding a programme for an undefined period of time is not feasible, but the capacity for implementation is limited.

5.4.5.3. Cluster randomisation

There are many evaluation settings in which individual-level randomisation is not practical. Many insurance schemes cover all members of the household, thus individuals from the same household must be assigned to the same group. Similarly, when a microinsurance scheme is introduced to selected households in a community, others often quickly find out, creating an operational challenge to keep the insurance from control households in the same community. The unit of randomisation can be adjusted depending on the evaluation setting. Individuals, households, communities, and schools

are some of the most commonly used units of randomisation, but any cluster of target individuals can be treated as one unit of observation and receive the same random assignment as long as the number of randomisation units satisfies the power calculation. (Power calculations must account for the clustered design.)

5.4.6. Things to be considered⁷

5.4.6.1. Externalities

Externalities (or spillover effects) refer to indirect effects of a programme on non-programme participants. For example, access to a microinsurance scheme could have unintended consequences for the non-participants. If enough number of programme participants in the treatment group, start withdrawing from informal insurance arrangements, the informal insurance mechanism in a community may no longer function. Thus, non-participants in the control group may end up without any insurance — this is a “negative spillover effect” of the microinsurance programme. On the other hand, a health microinsurance programme might improve the health status and capacity for economic activities amongst the participants. If they

⁷ For a technical discussion of methodological issues around randomised evaluations, see Duflo et al. (2008). Goldberg and Karlan (2007) also offer a comprehensive discussion on methodological issues of randomised evaluations in the context of microfinance.

remain engaged in the informal insurance arrangements, non-participants might benefit from stronger safety nets – this would be a “positive spillover effect” of the programme. If non-participants who are affected by the spillover effect (positively or negatively) are in the control group, programme impact cannot be accurately measured by comparing the treatment and control groups. Take the example given above with the informal insurance arrangements. If the negative spillover effect described above occurred and non-participants were not able to use informal insurance arrangements anymore, for both the treatment and control group the use of informal insurance arrangements would be reduced (let us say by 30% for the treatment group and by 20 % for the control group). For the treatment group, this would be due to the use of microinsurance. For the control group, this would be due to the breakdown of informal insurance structures, a development which is also based on the introduction of the microinsurance programme. Thus, both effects would be due to the microinsurance programme. However, when comparing the results of treatment and control group, one would only look at the *difference* between the two effects and interpret this as result of the microinsurance programme for the treatment group (i.e., a reduction of 10%). Since the control group is not participating in the programme, one

would assume a zero effect for this group (which is in fact the basis of the calculations). But in reality, it would have been a reduction of 30% for the treatment group and a reduction of 20 % for the control group.

The simplest way to minimise the spillover effect amongst the control group is to employ cluster randomisation. In the example of informal insurance arrangements above, cluster randomisation at the community level could prevent the negative spillover from affecting the control group because everyone in the same community receives the same assignment. More generally, cluster randomisation reduces the risk of spillover effect on the control group by treating a group of people who may affect each other through the intervention as one unit of randomisation and give everyone in the group the same assignment.

In some cases, we are also interested in measuring spillover effects amongst non-participants as part of programme impact. Because insurance providers in the formal sector would not generally enrol everyone in the community where they enter, how the introduction of formal insurance affects vulnerability of non-participants is in itself an important question. In a cluster randomisation design, we can measure the spillover effect by comparing the outcomes between non-participants

in the treatment group and potential non-participants in the control group — the trick is to identify potential non-participants in the control group who look just like the non-participants in the treatment group. For example, if there is a clear and observable eligibility criteria for the programme, outcome comparison of non-eligible individuals between treatment and control groups could capture the spillover effect.

We emphasise the importance of identifying potential spillover effects before implementing an evaluation because they can cause potential biases on impact estimation. A small pilot of the intervention could often help identify the potential spillover effects and simple solutions that can be incorporated into the evaluation design.

5.4.6.2. Contamination of random assignment

One of the biggest methodological threats to RCTs is the potential contamination of random assignment. However, sometimes contamination is

actually a good thing to have, as long as one is aware of its possibility and measures it. If the contamination is natural, i.e., a result of externalities from the intervention, then this is part of what one ideally sets up the evaluation to measure. Random assignment of the intervention is essential to rigorous evaluations because it creates statistically comparable groups that have on average the same observed and unobserved characteristics. When this assignment is contaminated (i.e., people in the control group receive the treatment), our estimate of the effect of the programme could no longer be accurate if not taken into account. Contamination may happen because of the information spillover to the control group. A lack of supervision and operational error could also result in contamination. Again, there are analytical methods that adjust for any contamination of random assignment, but the estimation would become less precise. The evaluators and implementing organisation should make every effort to minimise contamination through proper monitoring of the implementation.

It's important to make efforts to keep the treatment group separate from the control group

5.4.6.3. Sample attrition

Sample attrition becomes problematic in RCTs when individuals who drop out of the study in the treatment and control groups are systematically different. Suppose that the presence of a microinsurance programme in rural communities discourages migration amongst the poor by reducing vulnerability to economic shocks. When the researchers follow up with the individuals in the sample after a year, they find that a large proportion of the poor in the control group has migrated to unknown locations and cannot be interviewed. In another case, the sample for the follow-up survey in the control group could be systematically wealthier than that in the treatment group. Because the impact of the microinsurance programme on the wealthy may be different from the impact on the poor, the comparison between the two groups no longer gives an accurate estimate of the average programme impact.

The impact estimate can be adjusted in the analysis if the factors that explain the differential attrition are known and measurable. For example, if we knew who in the treatment group *would have* migrated away, had the microinsurance programme not been offered, we could remove them from the analysis sample and estimate the programme impact amongst those

who would have stayed in the community in the absence of the intervention. The estimated impact here is not the average effect of the microinsurance programme amongst the entire initial sample frame. However, it gives us an unbiased impact estimate for a particular population group.

In reality, explaining the differential attrition between treatment and control groups is not easy. In order to minimise the attrition problem, it is useful to collect additional contact information (on neighbours, relatives in the home province, etc.) from the study participants before intervention. Note that not all sample attritions pose a threat to the validity of the evaluation. If the intervention does not affect the likelihood of remaining in the study at the end of the evaluation period, the comparison between the treatment and control groups gives an accurate estimate of the programme effect.

5.4.6.4. Presence of other NGO and government programmes

Whilst random assignment ensures that the treatment and control groups are balanced before programme implementation, there is always a small chance that other policies and programmes affect treatment and control groups differently during the programme implementation. When these external interventions are relevant

The baseline survey is useful but not a prerequisite for RCT

to the programme being evaluated, it threatens the integrity of the evaluation design. For example, other insurance providers may decide to enter the market and offer similar insurance products to individuals in our evaluation sample. The evaluation can still tell us the marginal impact of one additional insurance provider as long as the aggregate insurance coverage in the treatment group is higher than the coverage in the control group. In fact, given increasing competition in the microinsurance market, it is difficult to find a setting where there is a large enough demand for microinsurance and where we can plausibly expect no insurance provider, other than the programme being evaluated, to enter whilst the evaluation takes place. Whilst we need to recognise that the interpretation of the marginal impact of one additional microinsurance programme is different from that of expanding access to microinsurance, it is still a useful question for microinsurance providers to learn about the marginal impact that their programmes can make.

5.4.7. Data collection and analysis

5.4.7.1. Data collection

In a typical RCT, the evaluator collects both baseline and follow-up data. Baseline data, however, is not strictly necessary in RCT. The random assignment of intervention ensures that the treatment and control groups have comparable characteristics — this is an assumption we could make as long as the proper randomisation was carried out. That said, there are three advantages to collecting baseline data. Firstly, baseline data can be used to validate randomisation and ensure the balance across treatment and control groups. Even though we expect the proper randomisation to create comparable treatment and control groups, there is always a small chance that something goes wrong. It is reassuring to be able to confirm the balance between the treatment and control groups in the data. Secondly, baseline data allows the evaluator to compare the *change* in an outcome measure over time, rather than the *level* of an outcome measured after the programme implementation. Comparison of the change in an outcome captures the programme impact more precisely.

Finally, baseline data allows the evaluator to explore differential programme impacts for various population groups. For instance, to compare the programme impact amongst the poor and non-poor, the evaluator needs the data on the poverty status before the programme implementation.

Basic outcome measures, such as demographic information and insurance take-up, could be collected from insurance providers and implementing organisations. When assessing the impact of insurance on household well-being, however, additional data may need to be collected through field interviews. Ideally, these field interviews are conducted by a third party which is not involved in programme implementation. Separating implementation and survey activities ensures that the collected data is not affected by people's expectations on what services and products they might receive in the future from the implementing organisation.

5.4.7.2. Analysis

One advantage of RCTs is the simplicity of its analytical framework. When there is no methodological issue arising during programme implementation (such as contamination of random assignment and spillover), the programme impact can be measured by simple comparisons of outcome

measures between treatment and control groups. The statistical test tells us the likelihood that we observe the mean difference between the treatment and control groups by luck (rather than the programme). When this likelihood is small enough ($< 5 - 10\%$), we call the result "statistically significant".

Note that this comparison must be conducted between *everyone* in the treatment group and *everyone* in the control group regardless of who actually enrolled in the insurance scheme and dropped out. The point of random assignment was to create two statistically comparable groups. If we select different groups of individuals from the treatment and control groups for the analysis (e.g., only those who remained in the insurance scheme in the treatment group vs. everyone in the control group; only those who ever enrolled in the insurance scheme in the treatment group vs. those who never enrolled in the insurance scheme in the control group), the two groups are no longer comparable.

5.4.8. Take-away messages from this section

- Evaluations must be planned *before* the programme implementation and incorporated in the implementation so that the methodological issues can be addressed in the implementation design.

- Pilot the intervention before conducting the evaluation and resolve all operational issues.
- Randomised controlled designs can be adjusted to the constraints of operational settings and feasibility.
- Close monitoring of the evaluation implementation is critical in minimising the methodological issues that may arise during the implementation and preserving the power to detect programme effects.
- Analysis must be conducted with the entire evaluation sample regardless of the actual intervention outreach.

5.5. Summary

The microinsurance industry continues to evolve through product innovation and process improvement. Rigorous evaluations of existing and innovative insurance schemes play an important role in guiding the industry to develop schemes that are more effective and sustainable. The typical evaluation method that compares the well-being of participants before and after enrolment in the insurance scheme, has a number of methodological problems in accurately measuring the programme impact. Randomised controlled trials are an effective tool

to measure the exact impact of micro-insurance schemes and inform the industry about potential product and process innovation. RCT designs are flexible and adjustable to specific operational and budgetary constraints as well as ethical concerns.

Whilst the RCT methodology is increasingly employed in evaluations of development projects around the world, there have been few RCTs on micro-insurance schemes. The replication of RCTs in various settings is critical to accumulate knowledge of what works and what doesn't. Whilst individual evaluations provide useful information for the microinsurance providers whose products are being evaluated, the replication of evaluations would benefit the microinsurance industry as a whole. By testing the impact of insurance products in different contexts, we will gain better understanding of effective product designs and distribution mechanisms in various settings, or *why* and *when* a particular insurance product works. This accumulation of knowledge could effectively accelerate the innovation and expansion of microinsurance.

References

- Anagol, S., S. Cole, and S. Sarkar. 2013.** Understanding the advice of commissions-motivated agents: Evidence from the Indian life insurance. Harvard Business School Finance Working Paper No. 12-055, Harvard Univ.
- Bauchet, J. and J. Morduch. 2010.** An introduction to impact evaluations with randomized designs. A Research Framing Note for the Financial Access Initiative, New York Univ. http://www.financialaccess.org/sites/default/files/blog/RFN12_Introduction_to_Impact_Evaluations_with_Randomized_Designs.pdf
- Burns, C. and A. Dalal. 2010.** Explaining insurance: Implementing consumer education in CARE-India's Insure Lives & Livelihoods Program. A research report for the Financial Access Initiative, New York Univ.
- Cole, S., X. Giné, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery. 2013.** Barriers to household risk management: Evidence from India. *American Economic Journal: Applied Economics*, American Economic Association Vol. 5(1): 104-135.
- Duflo, E., R. Glennerster, and M. Kremer. 2008.** Using randomization in development economics research: A toolkit. In: *Handbook of Development*. Vol.4, ed. T. P. Schultz and J. Strauss, 3895-3962. Amsterdam and New York: North-Holland.
- Giné, X. and D. Yang. 2009.** Insurance, credit, and technology adoption: Field experimental evidence from Malawi. *Journal of Development Economics* 89(2009): 1-11.
- Goldberg, N. and D. Karlan. 2007.** Impact evaluation for microfinance: Review of methodological issues. Doing Impact Evaluation Series, No. 7. Washington, D.C.: World Bank.
- Karlan, D., R. D. Osei, I. Osei-Akoto, and C. Udry. 2012.** Agricultural decisions after relaxing credit and risk constraints. National Bureau of Economic Research (NBER) Working Paper No. 18463. Cambridge, MA. <http://www.nber.org/papers/w18463>
- Kremer, M. and R. Glennerster. 2012.** Improving health in developing countries: Evidence from randomized evaluations. In *Handbook of Health Economics*. Vol. 2, ed. M. V. Pauly, T. G. McGuire, and P. P. Barros, 201-315. Oxford: North-Holland.
- Schultz, E., M. Metcalfe, and B. Gray. 2013.** The impact of health insurance education on enrollment of microfinance institution clients in the Ghana National Health Insurance Scheme, northern region of Ghana. Research Paper No. 33. International Labour Office, Microinsurance Innovation Facility, Geneva.
- Siegel, P., J. Alwang, and R. S. Canagarajah. 2001.** Viewing microinsurance as a social risk management instrument. Social Protection Discussion Paper Series No. 0116, World Bank, Washington, D.C.
- Thornton, R., L. E. Hatt, E. M. Field, M. Islam, F. Solis Diaz, and M. A. Gonzalez. 2010.** Social security health insurance for the informal sector in Nicaragua: A randomized evaluation. *Health Economics* 19(Suppl. 1): 181-206.

Non-experimental methodologies for quantitative analysis

*Markus Frölich, Andreas Landmann,
Markus Olapade, and Robert Poppe*



6.1. Introduction

The ultimate objective of quantitative analysis is to establish causality. Researchers want to know the causal influence of a factor—the effect that can be attributed to this factor and to the factor only. Done correctly, quantitative analysis allows both quantifying the magnitude of this causal effect and computing statistical precision of estimation (confidence intervals). The goal of an impact evaluation is to measure the causal effect of a policy reform or intervention on a set of well-defined outcome variables.



Understanding reasons for participation in insurance helps to avoid bias in evaluation of results

Knowing causal relationships is useful for making predictions about the consequences of changing policies or circumstances; they answer the question of what would happen in alternative (counterfactual) worlds. As an example, one could try to identify the causal effect of introducing a community-based health insurance on health status or on out-of-pocket spending for health of the insured in a specific district of a developing country.

6.2. Selection bias and comparison issues

The fundamental problem of impact evaluation is the impossibility of observing an individual in two states at a moment in time; each individual is either in the programme under

consideration or not, but not both. The impact of a development programme can only be identified by comparing realised outcomes of those who did receive and of those who did not receive an intervention. Thus, data on non-participating individuals needs to be collected as well. The issue of selection bias is of central concern in this context: selection bias may arise when treated and non-treated individuals are different with respect to observed and unobserved characteristics. One reason could be for example, a project manager who deliberately chooses some individuals to be eligible for the programme but not others. Another important source for selection bias is *self-selection*, i.e., when individuals themselves choose to be treated or not.

Thus, any impact evaluation should be based on a detailed understanding about why some individuals or communities participated, whilst others did not participate in the programme; otherwise, results are likely to be biased.

Box 1: Potential outcomes and objects of interest

To make ideas more precise, let Y_i^1 denote the outcome if individual or community i is exposed to development intervention D . Before programme start, each individual (or community) has two hypothetical outcomes: a potential outcome Y_i^1 if individual i participates in the programme, and a potential outcome Y_i^0 if individual i does not participate in the programme. The causal effect of the intervention is defined as the difference between Y_i^1 and Y_i^0 , i.e., the effect of participation in the programme relative to what would have happened had individual i not participated in the programme. The individual effect of the intervention is usually averaged over the population of interest, defined as the average treatment effect $E[Y^1 - Y^0]$. It can be interpreted as the average treatment effect for a person randomly drawn from the population or, alternatively, as the expected change in the average outcome if the individual status indicator variable of development intervention D were changed from 0 to 1 for every individual (provided that no general equilibrium effects occur), where individual i either receives ($D=1$) or does not receive ($D=0$) the treatment. In a policy evaluation context of particular interest is the average treatment effect on the treated (ATT) defined as $E[Y^1 - Y^0 | D=1]$. It may be more informative to know how the programme affected those who chose to participate in it than how it affected those who could have participated but decided not to.

One way to avoid selection bias is to randomly assign individuals to a treatment and a control group. We refer to randomised trials as methods that randomly assign individuals who are equally eligible and willing to participate into distinct groups; they are generally considered the most robust of all evaluation methodologies and sometimes referred to as the *gold standard*

(Angrist 2004). Given appropriate sample sizes, the two groups will have approximately the same characteristics and differ only in terms of the treatment status. They will be approximately equal with respect to variables like race, sex, and age, and also for difficult to measure variables, such as lifestyle-related risks, quality of social networks, and health awareness.

Box 2: How randomisation eliminates selection bias

Randomisation allows a simple interpretation of results: the impact is measured by the difference in means between treatment and control group. Why is this so? With randomisation, the average treatment effect on the treated (ATT) can be written as

$$\text{ATT} = E[Y^1 - Y^0 | D=1] = E[Y^1 | D=1] - E[Y^0 | D=1] = E[Y^1 | D=1] - E[Y^0 | D=0].$$

Thus, randomisation allows replacing the expected (unobserved) counterfactual outcome, $E[Y^0 | D=1]$, with the expected observed outcome of the non-participants, $E[Y^0 | D=0]$. Essentially, $E[Y^0 | D=0]$ is used to mimic the counterfactual. Because of randomised assignment, it holds that $E[Y^0 | D=1] = E[Y^0 | D=0]$, i.e., the expected non-programme participation outcome is the same whether an individual actually participates or does not participate in the programme. This last equality usually does not hold in non-experimental studies. Individuals have, for example, a better non-programme participation outcome if there is positive selection into the treatment ("they would have done better anyway"). This would lead to upward biased results.

In practice, there are several problems with randomisation. Firstly, it may be unethical to deny access to benefits for a subgroup of individuals.¹ Secondly, it may be politically unfeasible to randomly deny access to a potentially beneficial intervention. Thirdly, there may not be any individuals who are unaffected if the scope of the programme is nationwide. Fourthly, problems arise when, after randomised assignment, individuals cross over to the treatment group. For example, people might travel to another municipality to buy insurance after having learned that a microfinance institution offers a new life insurance scheme there. Fifthly, individuals assigned to the treatment

group may not take up treatment, or individuals assigned to the control group may seek similar treatment through alternative channels.

Sometimes, randomised trials are impractical. However, impact evaluations are most valuable when we use data to answer specific causal questions as if in a randomised controlled trial. In absence of an experiment, we may look for a natural or quasi-experiment that mimics a randomised trial in that there is a group affected by the programme and some control group that is not affected. If it is credible to argue that the groups do not differ systematically, such a natural or quasi-experiment can be used for evaluation instead of a randomised experiment.

1 For a discussion see Burtless and Orr (1986).

Unfortunately, it is often hard to justify that programmes that were not ex-ante planned as randomised experiments do in fact fulfill this criteria. This is why it is preferable to think about evaluation and the appropriate design before the programme is implemented.

In non-experimental studies, researchers often try to approximate a randomised experiment by using statistical methods. We will discuss several non-experimental methods in the following paragraphs. The issue of selection bias is of central concern here. Rather complex statistical methods are required in order to deal with selection bias when using non-experimental data. The methods differ in the way they correct for differences in (observed and unobserved) characteristics between the treatment and the control group and by their underlying assumptions.²

Two non-experimental methods—differences in means and before-after estimation (also called reflexive comparisons)—usually do not give a satisfactory solution to the selection issue when using non-experimental data. In general, this is because changes in the outcomes cannot be attributed to the programme. The former method, differences in means, is based on cross-sectional data using the outcome of the non-participants to impute

*Statistical methods
can help reduce
selection bias in
non-experimental
data*

the counterfactual outcome for the participants. The underlying assumption is that individual characteristics, on average, do not play a role for the difference between the treated and the non-treated, which is a strong assumption. The latter method, the before-after estimator, is based on (at least) two cross-sections of data—one cross-section before programme start and one cross-section after the programme. It uses the participants' pre-intervention outcome to impute the counterfactual outcome for the participants. The drawback of this method is that it is impossible to separate programme effects from general effects that occurred during the same period.

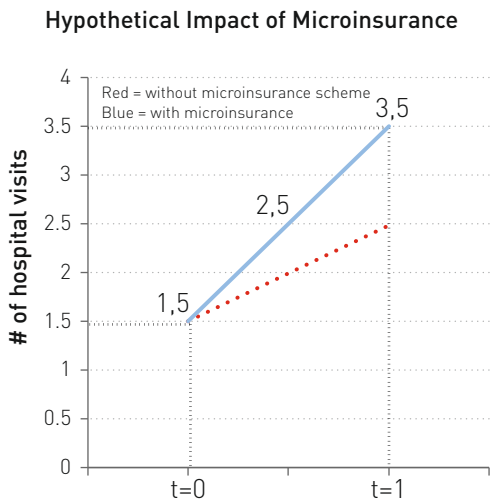
Throughout this chapter, a hypothetical example will display the different evaluation methods. By use of a microinsurance example for inpatient and outpatient hospital visits, we will explain the main concepts to determine the effect of this insurance on our outcome variable:

² For a more general and accessible introduction to impact evaluation see further Leeuw and Vaessen (2009).

number of hospital visits. This example can then be generalised to any other insurance and outcome of interest.

Figure 1 displays the impact health insurance has on the number of hospital visits. At time $t=0$, the microinsurance for inpatient and outpatient services in a nearby hospital is introduced in our village. At that point of time the villagers visit the hospital 1.5 times a year on average. At time $t=1$, the number of hospital visits has increased to 3.5 visits per year (blue line). Without the introduction of microinsurance, the frequency would have increased to 2.5 times only (dotted red line). As we will explain, the outcome 2.5 constitutes the counterfactual outcome. Thus, the (true) impact of the microinsurance scheme is 1 hospital visit per year.

Figure 1: Real impact of microinsurance on hospital visits



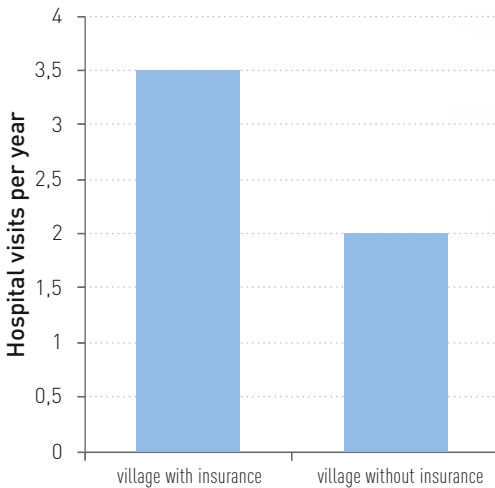
The cross-sectional comparison by *difference in means* is our starting point for the impact evaluation. Hereby, we compare two villages at $t=1$, where it so happens that in the treatment village microinsurance is available. If the inhabitants from this village buy a policy, they can make claims to the insurer after a hospital visit. Inhabitants from the control village do not have this opportunity and need to pay the total hospital bill from their own money. In a cross-sectional comparison we directly compare the hospital visits of the treated and untreated groups.

Figure 2 displays the naïve estimator of the treatment effect using a cross-sectional comparison. The treatment village has 3.5 hospital visits a year per person on average, where village 2 has only two visits per year. The naïve estimator is simply the difference between those two outcomes, which is 1.5, a biased estimate of the treatment effect.

It is obvious that the effect might be partially attributable to the availability of a health insurance, but it cannot be ruled out that other reasons have also affected this result.

1. Were the two villages different from the outset? Different in observables like wealth, education, or different in unobservables such as *trust in hospital staff*?

Figure 2: Cross-sectional comparison



Neither before-after nor comparison of means analyses between two villages provide reliable impact results

- Has one of the villages had other influences that might affect the number of hospital visits? The road leading to the hospital might have been blocked for village 2, hindering people from visiting the hospital and leading to an overestimation of our treatment effect.

Myriad scenarios can be constructed to answer these two questions. According to our example from figure 1, however, an actual comparable control village should have 2.5 hospital visits per year in $t=1$. This shows that using cross-sectional comparison, the impact evaluator cannot be sure whether the effect of microinsurance on hospital visits results only from the availability of microinsurance or also from other confounding factors.

Another naïve estimator would be the *before-after comparison*. Here, we need to conduct a baseline survey amongst the population of the treatment village (shortly) before the microinsurance is made available. We would ask explicitly how often a week the inhabitants go to the hospital.³ Then the insurance is made available and after a certain period of time the same survey questions are gathered again from the village. Of course, the time between the introduction of the microinsurance and the follow-up survey needs to be long enough for certain incidences of sickness and claims to occur. As depicted in figure 3, the estimated impact of our hypothetical insurance is the difference between the number of hospital

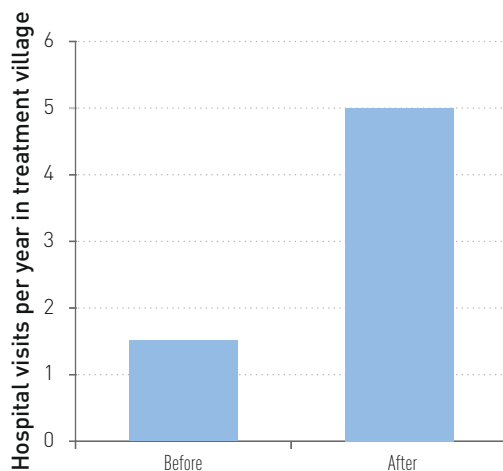
3 Of course, it might be more reliable to gather this information directly from the hospital: "How many inhabitants from village 1 visit the hospital a week?"

visits before and after introduction, which is two hospital visits per year. The interpretation would be that introducing microinsurance that covers the cost of inpatient and outpatient care increases the number of hospital visits by two visits per year.

However, this estimator relies on the important assumption that, in between our two surveys, no other factors have occurred that might cause a change in hospital visits. This means that, for the before-and-after estimator to produce reliable results, our researcher must be sure no other confounding effects have occurred between the two surveys. In fact, figure 1 shows that without microinsurance there still would have been an upwards trend in hospital visits and that, therefore, our before-after comparison delivers unreliable results. Reasons for an upwards trend in hospital visits independent of the treatment could be increases in prosperity, decreases in transportation costs, and many other scenarios.

In sections 6.3 through 6.6 of this chapter, the following four non-experimental approaches will be explained:

Figure 3: Before-and-after comparison



1. Instrumental variables
2. Regression discontinuity design (RDD)
3. Propensity score matching (PSM)
4. Difference-in-differences (DID)

Amongst them, the first two approaches, if applicable, usually give the most convincing results. We will also stress the importance of internal and external validity in each case. Internal validity is the extent to which the results are credible for the population under consideration. External validity is the extent to which this sub-population is representative for the whole population (of interest). Some

Internal validity is the extent to which the results are credible for the population under consideration

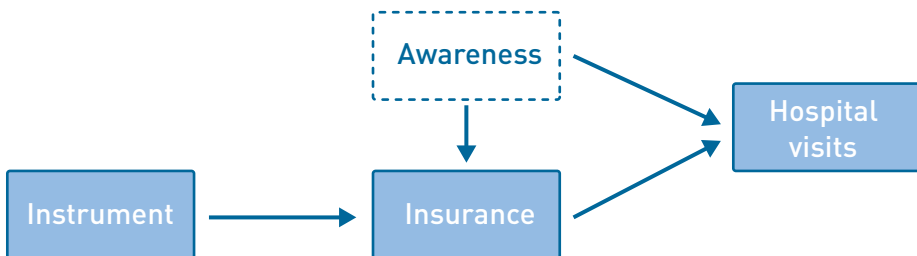
methods give results with high internal validity, but low external validity, and vice versa. We conclude with a discussion of where non-experimental methods should be applied.

6.3. Instrumental variable

Selection bias occurs when an omitted variable has an effect on the outcome variable of interest and the treatment. It is also called *selection on unobservables*, whereby treatment selection is affected by a variable that the researcher cannot observe in the data. For example, individuals with insurance could have had a higher (unobserved) awareness for health issues from the outset. Consequently, they would show different health behaviour, than those without insurance. Figure 4 illustrates this simple case with arrows indicating directions of influence and dashed lines indicating unobserved variables (health

awareness). The instrument affects insurance take-up without being itself affected by different levels of awareness about insurance. In the absence of a good instrument, one could not tell apart the insurance's effect and the awareness' effect on hospital visits. *Instrumental variable* methods solve this problem of omitted control variables. An instrumental variable is a variable which has an effect on whether an individual takes up or does not take up treatment and at the same time is permitted to affect the outcome variable of interest via the treatment variable only. This is called the *exclusion restriction*. In other words, individuals with different values of the instrument differ in their treatment status. But otherwise, these individuals are comparable. Often, the exclusion restriction will be only valid conditionally, that means when controlling for individual characteristics.

Figure 4: Setup with instrumental variable



Box 3: How the instrumental variables method solves the problem of unobservables

More formally, instrumental variable Z affects treatment status D , but there is no direct relationship between Z and the outcome variable Y . Hence Z is allowed to affect Y only indirectly via D . Suppose we have

$Y_i = \alpha + \beta D_i + \gamma X_i + \varepsilon_i$ where Y_i is the outcome variable, D_i is the treatment indicator variable, X_i is an unobserved variable that is correlated with D_i , and ε_i is a random error term.

If we now estimate (since X_i is unobserved),

$$Y_i = \alpha + \beta D_i + \varepsilon_i$$

D_i will be correlated with the residual ε_i because, in effect, we have

$$Y_i = \alpha + \beta D_i + (\gamma X_i + \varepsilon_i).$$

X_i and D_i are correlated, so D_i and $(\gamma X_i + \varepsilon_i)$ are correlated as well and, thus, D_i is endogenous.

The instrument can now be used to get an unbiased estimate of the effect of the endogenous variable. Researchers use a method that is called *two-stage least squares*: in the first stage, the instrument(s) Z is used to give estimated values of the endogenous treatment variable D for every individual (or community):⁴

$$\hat{D}_i = \alpha + \beta Z_i + \varepsilon_i$$

Then, in the second stage, this new variable is plugged in the equation of interest:

$$Y_i = \alpha + \beta \hat{D}_i + \varepsilon_i.$$

The coefficient β gives now an unbiased estimate of the treatment effect.

Using an instrument for the evaluation of microinsurance has not been done often and, in general instruments are hard to find. We therefore use a hypothetical example to illustrate our point in the context of this guide. Suppose the government sets up a health insurance programme for the poor. Everybody who is interested has to register

Instrumental variables
are hard to find but
can support high internal
validity of evaluation

⁴ Note that using predicted values as additional regressors in the way presented here only works in linear models.

and purchase the product at the local insurance administration centre of the neighbourhood or municipality. Now imagine two households that are very close, but on different sides of the border between two neighbourhoods. The distance to the administrative centre might differ considerably, but otherwise the two neighbours should be very similar. For such pairs, distance to the administrative centre could be used as a predictor of insurance take-up that is otherwise unrelated to individual characteristics—in other words, a good instrument. Here, the instrument is correlated with insurance take-up but not with awareness. Thus, instead of comparing the treated to the untreated, we compare those with high values and low values of the instrument. This example is analogous to the famous distance-to-school instrument used by Card (1995) for schooling. Typically, an instrument requires including additional X variables, e.g., quality of the neighbourhood, degree of urbanisation, family background, etc.

We may also generate instrumental variables ourselves by randomly assigning incentives or encouragements to individuals (random encouragement design). This approach looks very much like a proper randomised experiment, except that we have imperfect control over the beneficiaries. An encouragement or incentive is given to the individuals in the treatment group,

whilst the individuals in the control group do not receive such an encouragement or incentive (or receive a different one). It is up to the individuals whether they sign up for the actual treatment. For example, imagine that the price of insurance is varied randomly across communities, creating a random incentive to buy insurance for the population facing a lower price. The instrumental variable that is generated here helps resolve the problem of selection bias and allows consistent estimation of the effect that insurance take-up has on health and other outcome measures. Similarly, we may vary the effort related to take-up by, for example, varying service hours, density of offices in a community, etc., from the insurer's side. If areas or individuals cannot be exclusively chosen for a programme at random, we may at least give them varying incentives to do so.

If there is an instrument that fulfills the exclusion restriction as explained above, internal validity is high. However, external validity depends on another quality of the instrument. If the instrument predicts treatment status accurately, external validity is also likely to be high. Otherwise the instrumental variable results cannot be generalised to the whole population. The reason is that only those who are induced to take up treatment by the instrument can be used for the estimation of the treatment effect.

6.4. Regression discontinuity design

Although not as rigorous as random assignment, the regression discontinuity design (RDD) approach may give more convincing results than propensity score matching (PSM) and difference-in-differences (DID) methods (see below). The idea of RDD is to exploit some cut-off point that is important for selection into treatment and compare people near this cutoff. Thus, this approach implicitly compares treated subjects to a control group that is very similar.

In an RDD approach, people above and below a cut-off point are compared

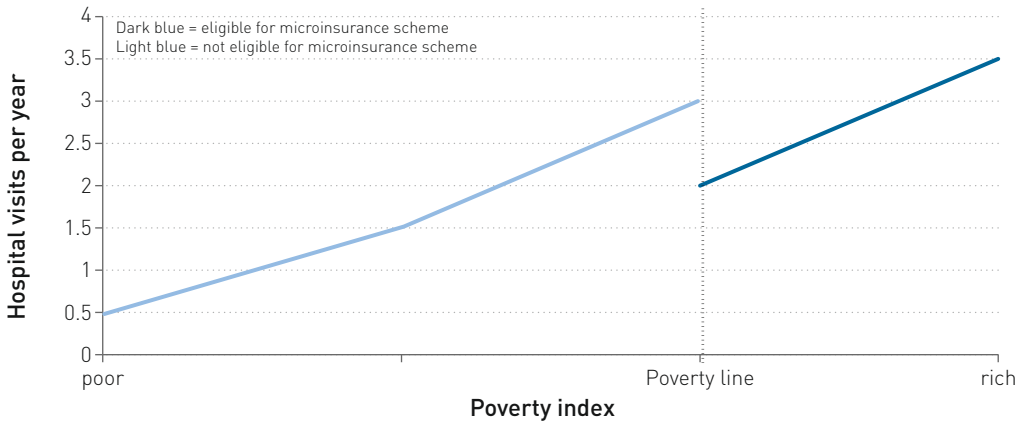
A standard application is when enrolment into treatment is limited and selection of participants is conducted according to an objective rule. Frequently, such form of targeting is done on the basis of a poverty index: individuals above the threshold receive the treatment, whilst it is withheld from those individuals below the threshold. When comparing individuals very close to this threshold, their characteristics barely differ, except with respect to their treatment status. It is basically random as to whether an individual is below or above the cutoff given that the individual is close to the cut-off value. However, for this approach to be valid, individuals must be unable to manipulate their value of the index such that they would become eligible

for the treatment.⁵ To test for the plausibility of this assumption, we can use statistical methods. Although discontinuities in evaluation studies are often unplanned, they may also be integrated ex-ante.

To give an illustrative example, imagine the government wants to introduce microinsurance especially for the poor. The village administration is responsible for the distribution of the insurance and relies on a poverty index to determine the eligible households. Only households that are considered poor under this index are eligible and can buy the insurance. Such discontinuities do not necessarily have to be

⁵ Even if they have some influence, the approach is feasible as long as they are unable to manipulate their assignment precisely. The solution in this case is the so-called fuzzy regression discontinuity design.

Figure 5: Regression discontinuity design



planned as part of the intervention (even though it is certainly beneficial to have it planned beforehand). Instead, the evaluator could detect and exploit any rule used in practice to determine participation in the programme.

Figure 5 shows the number of hospital visits by ranking in the poverty index sometime after the programme started in our hypothetical village.

We see that the number of hospital visits increases and that there is a jump exactly at the poverty line. This jump is a result of the microinsurance programme and the restriction that only poor people have access to this programme. Any household that is above the poverty line has no access to the insurance product. RDD assumes that households just above the poverty line are, in fact, similar to those slightly below the index in all relevant aspects.

Therefore, we can use the households that are eligible for the insurance and very close to the threshold as our treatment group, whilst those slightly above the threshold serve as control group.

The benefit of RDD is that it does not need actual randomisation. However, the interpretation of the estimated impact is limited to the population that is close to the threshold. As a result, external validity of this approach is rather limited. Further, it usually requires a large sample for estimation.

6.5. Propensity score matching (PSM)

The basic idea of propensity score matching is to match at least one non-participant to every participant with identical or highly similar values of observed characteristics X . The difference in outcome, Y , between these

two individuals is then averaged over the whole population. A practical simplification is to match non-participants

to participants on the basis of the propensity score, which is defined as the probability of treatment.

Box 4: How matching eliminates selection bias

Using matching, the average treatment effect on the treated (ATT) can be defined as $ATT = E[Y^1 | D=1, X] - E[Y^0 | D=1, X] = E[Y^1 | D=1, X] - E[Y^0 | D=0, X]$. As in the randomised trial, the expected counterfactual outcome, $E[Y^0 | D=1, X]$ can be replaced by the expected observed outcome of the non-participants, $E[Y^0 | D=0, X]$, but only conditional on a set of observable covariates, X . If treated and non-treated differ in terms of observable characteristics X only and not in terms of unobservables (the so-called selection on observables assumption), it holds that $E[Y^0 | D=1, X] = E[Y^0 | D=0, X]$. Otherwise, selection bias will remain an issue.

Instead of matching on X , it suffices to match on the propensity score $p(X)$, i.e., the probability of treatment defined as $p(X) = Pr(D=1|X)$ (Rosenbaum and Rubin 1983). We can therefore also write

$$ATT = E[Y^1 | D=1, p(X)] - E[Y^0 | D=1, p(X)] = E[Y^1 | D=1, p(X)] - E[Y^0 | D=0, p(X)].$$

It is important to include all variables in X that affect the outcome and selection into the programme at the same time. Another requirement is that the X variables need to be unaffected by the treatment, i.e., they should be measured before treatment starts. PSM requires both a thorough understanding of the selection process and a large data basis. Qualitative interviews with local project managers and participants may be helpful to determine which variables to collect in order to ensure that all important variables are included in X .

Panel data, if available, would allow testing for the plausibility of the underlying assumptions by conducting a pseudo treatment test. The idea is to pretend that the participants received the treatment before the start of the intervention and then to estimate the impact. Because the intervention had not been in place yet, the estimated effect should be zero. If the estimation leads to a different finding, then this should be taken as evidence that participants and non-participants differ with respect to unobserved characteristics. If we are willing to assume

*PSM gives rather low internal validity,
but external validity can be high*

that these differences are time-invariant, then we can use a DID matching approach. If, however, we suspect that these differences change over time, then we need more or better X variables or a better understanding of the selection process.

Propensity score matching gives rather low internal validity due to its reliance on the selection on observables assumption. In other words, the results might be biased if there are variables that are correlated with insurance take-up and the outcome of interest (such as hospital visits), but cannot be observed in the data. External validity can be high, except in the case that we cannot find sufficiently comparable untreated individuals to be matched with every treated individual (the so-called common support requirement). These treated individuals would then need to be excluded from the analysis, which would reduce external validity.

6.6. Difference-in-differences (DID)

Relying on the assumption that selection is on observables only can be

difficult to justify. Often we need a method that can also take care of confounding variables that are unobserved. However, as already mentioned, good instruments are hard to find. Therefore, we would like have other tools to deal with unobservables. The DID estimator uses data with a time or cohort dimension to control for unobserved but time-invariant variables. It relies on comparing participants and non-participants before and after the treatment. The minimum requirement is to have data on an outcome variable, Y , and treatment status, D , before and after the intervention. (It can be carried out with or without panel data and with or without controlling for characteristics, X .) In its simplest form, we take the difference in Y for the participants before and after the treatment and subtract the difference in Y for the non-participants before and after the treatment. As a result, time-invariant differences in characteristics between participants and non-participants are eliminated, allowing us to identify the treatment effect. Consequently, this approach accounts for unobserved heterogeneity as long as it is time-invariant.

Box 5: How the DID estimator accounts for time-invariant unobservables

To make ideas more precise, suppose we have

$$Y_i = \alpha + \beta D_i + c + \varepsilon_i,$$

where c is a time-invariant variable.

With observations at time $t=0$ and $t=1$ we can take the first difference

$$Y_{i,1} - Y_{i,0} = \alpha + \beta(D_{i,1} - D_{i,0}) + (c - c) + \varepsilon_{i,1}$$

Importantly, the time-invariant characteristic c drops out. As discussed in box 3, using just one cross-section of data will lead to a biased treatment effect β if c is correlated with ε_i . By using the DID approach we get rid of the problematic unobservable.

Panel data are helpful but not strictly required. Having cross-sectional data before and after the treatment may suffice. For instance, if villages participate in the intervention entirely, whilst other villages do not participate, it will suffice to conduct representative surveys in the villages before and after the intervention, i.e., interviewing the same individuals in the villages is not required. Thus, this method allows us to avoid problems with attrition commonly found in panel surveys.

The simple DID approach eliminates time-invariant heterogeneity. However, it fails to account for systematic differences in time trends between participants and non-participants. Therefore, we should include additional control

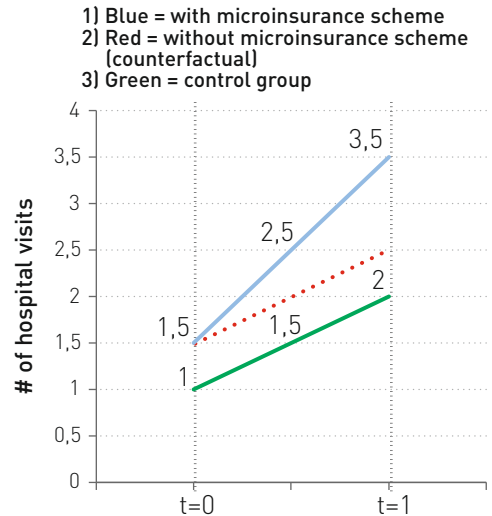
variables, X , if we can argue that time trends are the same at least for treated and non-treated with the same X . This can for example be done with PSM. Another extension is to use additional differences of unaffected comparison

Additional control variables should be included to account for systematic differences in time trends

groups. For instance, imagine an insurance product applies only to individuals below the age of 40. We can then compare the time trend of individuals above the age of 40 in the treatment villages with the time trend of those above 40 in the control villages. This difference in time trends can be used to eliminate differences in time trends of those under 40. A further possibility is to use data for more than one point in time before the treatment is introduced. This would also allow eliminating differences in time trends. Having more than one survey after the treatment implementation additionally allows the estimation of time-varying and long-run treatment effects.

In order to apply the DID estimator to our hypothetical example, we need data at two points in time for two different villages: one village where insurance is available and another one where it is not. The treatment effect is then calculated as the difference in hospital visits between the two villages after the introduction of our insurance in village 1 ($t=1$) and the difference between the two villages before insurance was introduced ($t=0$). Two scenarios show how this method relies on the assumption that time-trends of the two villages are equal. In figure 6, the two villages have parallel time-trends. This means that the counterfactual of the village with insurance (red dashed line) changes over time in the same way as

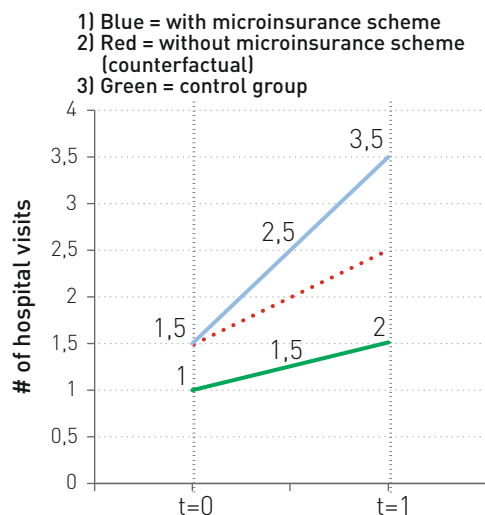
Figure 6: Difference-in-differences with parallel time-trends



the control village (the green line). In this scenario the treatment effect is $(3.5-2)-(1.5-1)=1$. We receive a reliable result for our treatment effect.

However, the second scenario shows that if the time-trends are different the DID estimator does not produce reliable results: $(3.5-1.5)-(1.5-1)=1.5$. Here, our estimated treatment effect overestimates the true treatment effect by 0.5 hospital visits because the treatment and control villages do not have parallel time-trends. Reasons for the difference in time-trends can be, for example, macroeconomic effects that affect treatment and control villages differently or any other confounding factors that influence the number of hospital visits in one village but not in the other.

Figure 7: Difference-in-differences without parallel time trends



Internal validity of the DID approach hinges on the assumption that participants' and non-participants' outcome variables under consideration have the same time trend. As explained, there are tests to check for the plausibility of this assumption. In addition to the common trend assumption, we require that there are no spill-over effects from the participants to the non-participants. If these assumptions are fulfilled, internal validity is high. External validity is high, as long as the sample our data is based on is representative for the population of interest.

6.7. Fields of application of non-experimental methods

Non-experimental quantitative impact evaluation can be applied to many areas.

However, the methods described belong to the field of microeconometrics, which are suited to evaluate interventions on the micro- or meso-level. A central element is that there exist different units—individuals, firms, hospitals, water works, villages, local administrations, neighbourhoods, districts, etc.—some of which were exposed by the treatment, whilst others were not. In order to evaluate projects on the macro level, such as budget support for balancing the national budget, other econometric methods are more suitable.

Convincing evaluations based on non-experimental methods require a detailed understanding of the selection mechanism and comprehensive and representative data on the treatment and comparison groups. Moreover, we usually need more than 1000 observations in order to obtain sufficiently precise estimates of the impact. Non-experimental methods are not suitable as a monitoring instrument for projects in the phase of introduction and should only be applied after resolving initial problems. Apart from initial obstacles, there are often larger modifications of the originally planned intervention, making a precise definition of the treatment more difficult.

Quantitative methodologies can be used to evaluate the impact of an intervention compared to a situation without. It is more informative, however, to

evaluate the impact of an intervention relative to other interventions or, alternatively, to evaluate different variants of an intervention, keeping context and data collection procedure constant. This, for example, would allow to look at the impact of different incentives or cost sharing arrangements for subsidised insurance.

Non-experimental methods generally give less convincing results than experimental methods. Moreover, if the confidence intervals turn out to be very wide, we should not interpret these non-significant results as evidence for the absence of an impact. This interpretation is only valid if the confidence intervals are very narrow. The correct interpretation would be

that the sample size was too small to draw reliable conclusions.

References

- Angrist, J. D. 2004.** American education research changes track. *Oxford Review of Economic Policy* 20(2):198-212.
- Burtless, G. and L. L. Orr. 1986.** Are classical experiments needed for manpower policy? *Journal of Human Resources* 4:606-639.
- Card, D. 1995.** Earnings, schooling, and ability revisited. *Research in Labor Economics* 14:23-48.
- Leeuw, F. and J. Vaessen. 2009.** Impact evaluations and development: NONIE guidance on impact evaluation. Washington, D.C.: The Network of Networks on Impact Evaluation (NONIE). http://siteresources.worldbank.org/EXTOED/Resources/nonie_guidance.pdf
- Rosenbaum, P. R. and D. B. Rubin. 1983.** The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41-55.

Qualitative designs

Tara Sinha, Christina May, and Jeremy Holland



7.1. Why use qualitative methods for impact assessment

The terms *qualitative* and *quantitative* refer to the type of data generated in the research process. Quantitative research produces data in the form of numbers whilst qualitative research tends to produce data that are stated in prose or textual forms. Whilst quantitative methods measure the impact of microinsurance in terms of quantitative indicators, qualitative methods can be used to understand the processes that explain this impact or, in the absence of quantitative impact studies, to explore what kind of impact can be expected. For example, quantitative methods may be used to measure how much impact microinsurance has, e.g., on the financial situation of the insured households, whilst qualitative methods ask *why and how* microinsurance impacts the financial situation and what the underlying processes are. Qualitative

methods are used to look in depth at impacts; the data and analysis generated is interpreted in context. Contextual methods are applied to a specific locality, case or social setting, and sacrifice breadth of population coverage and statistical generalisability in order to explore or understand issues in depth (Booth et al. 1998).

Particularly for a relatively new field like microinsurance research, qualitative studies are needed to explore potential impacts and processes. Results may then be the basis for succeeding quantitative studies. Without knowing what kind of impact microinsurance may have, it is difficult to measure it using a research design based on theoretical assumptions only. Moreover, qualitative research is able to use social analytical frameworks to interpret observed patterns and trends—including analysis of socially differentiated outcomes. Without these analytical insights into the complex *missing middle* between interventions and impacts, researchers and policy analysts tend to make *interpretive leaps* of analysis based on what is measured (Chambers 1995).

Due to this iterative relationship, qualitative and quantitative methods are particularly effective when used in combination. However, when considering ways to combine quantitative and qualitative methods and data, it is

Qualitative research to explain or explore impact dimensions

important to be aware of their comparative advantages and to recognise that “strong fences make good neighbours” (Appleton and Booth 2005).

Within this chapter, we firstly summarise the differences of qualitative and quantitative impact assessments and highlight their specific strengths. We will then discuss quality criteria for qualitative impact assessments before we conclude on possible research designs and data collection methods.

7.2. The nature of qualitative impact assessment

Though there is increasing dialogue between quantitative- and qualitative-oriented research methods and the application of mixed methods is becoming more and more popular, a differentiation between both approaches along the research process helps in grasping the distinctive nature of qualitative impact assessment.

Table 1: Differentiation of quantitative- and qualitative-oriented approaches to social science

Research process	Quantitative-oriented impact assessment	Qualitative-oriented impact assessment
Epistemological background and forms of knowledge	• Postpositivist: deductive procedures, probabilistic law	• Interpretivist: understanding subjective and contextual knowledge
Research problem and research question	• What and how much impact for whom? • Hypothesis-testing	• How and why does impact occur/ not occur: mechanisms and theories of change • Discovery-oriented: What unintended and unexpected impact occurs?
Sampling	• Random • Large n	• Purposive and stratified random • Small n
Data collection	• (Quasi)experimental setting • Tight prescription of research design • Large-scale, standardised surveys	• Naturalistic inquiry • High flexibility of research design • In-depth interviews (focus groups, key informants), observations and participatory methods • Involvement of researcher in data collection
Data analysis	• Calculation and statistical generalisation • Assess impact along predefined and externally imposed indicators • Tables and numbers	• Interpretation and particularisation / analytical generalisation / exploration • Understand and represent the target group's perception of meaning, existence and significance of impact • Rich and thick description

Epistemological background: Quantitative and qualitative approaches to impact assessment are based on different philosophies on how the world can be perceived and described. Quantitative approaches to impact assessments are based on the so-called postpositivist school of thought, which pursues objectivity and relies on deductive procedures and probabilistic laws to understand our world. Constructivist and interpretative thinking builds the background for qualitative research, supporting the view that social phenomena can only be understood when the perception of the world by human beings is taken into account. Thus, subjective knowledge and specific contexts are of importance (Della Porta and Keating 2008). For the remainder of the chapter, we will use only the terms quantitative and qualitative to differentiate between approaches.

Research problem and research question: Whilst quantitative-oriented impact assessments are interested in *what* and *how much* impact has occurred for whom, qualitative-oriented impact assessments focus on the *how* and *why* impact has occurred or not occurred. For example, qualitative methods would be suitable for the following questions on the impact of microinsurance:

- How does the insurance influence the financial/economic/social behaviour of the insured?
- How do the insured perceive the value of insurance? Why do some perceive value and others do not?
- How does the perception of impact influence the insured's enrolment decision?

In some cases, one is not sure whether all possible impacts of microinsurance have been accounted for. In these scenarios, explorative qualitative studies can help in identifying impact on areas which had not been considered before. Qualitative methods can also be useful for explaining missing, unexpected, or unintended impacts reported in quantitative surveys, e.g., to explore the underlying mechanisms when a quantitative impact assessment shows that a health microinsurance scheme does not lead to a reduction of or to a rise in out-of-pocket payments. For more examples of qualitative method applications, see examples 1-11 in section 7.7.

Change is not always a linear process and, thus, mechanisms and causal chains behind impact require close examination, for which qualitative methods are especially suited (Roche 1999; Kabeer 2003, 113; Faust 2010). Qualitative inquiry and impact assessments are discovery-oriented. They do not limit possible outcomes of their research beforehand and react

in a flexible manner to new discoveries, adapting the research process if necessary, and are thus better able to account for unexpected findings and impacts (Patton 2002, 39,44; Creswell 2009, 176; Kabeer 2003, 113).

Qualitative studies are better suited to capture unexpected results

Box 1: Examples of *how*-questions in qualitative impact assessments

In a study on a health insurance programme in India, McGuinness (2011) did not only intend to study whether the programme had an impact on household's financial protection, but also how the programme provided this protection. Additionally, the community-managed reimbursement process of the scheme and its influence was assessed. McGuinness chose to apply a qualitative approach, combining household case studies with claims data.

Hietalahti and Linden (2006) conducted a study "to gain a better understanding of how microcredit projects impact on rural women's livelihood structures, and how they can strengthen women's welfare", applying qualitative methods which were "designed to encourage respondents to describe their experience in their own words" (Hietalahti and Linden 2006, 202-204).

In his study on the impact of a life microinsurance product in Indonesia, Hintz (2010) explains how, due to the field and study situation, he moved to an explorative-qualitative approach and how he finally found that the impact of the product under study was not linear and mostly unintended.

There can be both theoretical and practical reasons for choosing a qualitative impact assessment design. In some cases, qualitative impact assessments are preferred over quantitative impact assessments because of limited resources and difficulties in fulfilling formal requirements (Hulme 2000; Copestake et al. 2005). Qualitative impact assessments can be less costly than quantitative, experimental, or

quasi-experimental research designs since they do not involve large-scale surveys and do not necessarily make use of a comparison group to investigate causality.

Sampling: The sample size of qualitative studies is small; its specific size depends on the study question and the study purpose. As Patton puts it, "The validity, meaningfulness, and insights

generated from qualitative inquiry have more to do with the information richness of the cases selected and the observational/analytical capabilities of the researcher than with sample size" (Patton 2002, 245). The sample in a qualitative study may be chosen purposefully or randomly. Whilst sampling purposefully, one should be aware of what cases are needed for the question under study, i.e., whether the case should be typical for the study population or display a certain characteristic. Typical cases can be defined by a preceding quantitative survey (e.g., individuals with more or less the mean characteristics of the underlying population). Less typical cases, which are insightful for a certain aspect under study (e.g., individuals with a particular chronic disease or disability, or households that have accumulated assets or diversified livelihoods in the face of prevailing trends), can be identified from survey data, from direct observation, or by snowball sampling through referral from group discussions and interviews, a sampling technique where initial respondents are asked for recommendations on who else to interview. The expectation is that persons belonging to a particular category already know each other and are thus able to provide references. For many research situations, however, stratified random sampling is preferable. For this, the population is divided into subpopulations by differentiating by characteristics like e.g.

occupation, income or religion. From each of the subpopulations, a certain number of elements are then randomly selected. This reduces selection bias (because of random sampling) whilst at the same time allows in-depth qualitative research to capture the variability of context (because of stratification).

Data collection: Qualitative impact assessments are naturalistic inquiries, data is collected "in the field" without creating a lab situation and manipulating the phenomenon under study (Creswell 2009, 175; Creswell 2007, 37). They can however make use of "natural experiments" where they are able to observe and document a change in the real world and its implications, for example when a new programme is implemented and the possibility opens up to accompany this process (Patton 2002, 42). Whilst quantitative-oriented approaches make use of large-scale, standardised surveys in order to generate representative findings, qualitative-oriented approaches conduct semi-standardised, or open interviews, and/or observations to gain an in-depth understanding of the phenomenon under study.

Data collection for a qualitative study differs significantly from that in a quantitative study. Since the data is gathered through in-depth interviews or focus group discussions, the tools for data collection, such as guides for the

interviews or focus group discussions, are open-ended and relatively unstructured. The investigator needs skills that enable him/her to probe along each of the topics listed in the guide, to follow up on comments or remarks made by the respondent, and to bring a conversation back on track if it is going off course. For this, the investigator needs the requisite skills, which are usually much more than those required of an investigator filling out a structured questionnaire for a survey.

Data analysis: Quantitative impact assessments aim at statistical generalisation, i.e., they intend to create results which can be generalised to whole populations. There are different opinions about the ability to generalise from qualitative inquiry: some researchers argue that qualitative research does not even aim to generalise and that it only means to explain the particular case in its specific context. Others want to extrapolate findings for cases with similar contexts or to generalise findings to broader theories—i.e., aim at analytical generalisation. Qualitative inquiry does not analyse its data statistically—though sometimes numeric measures can be used—but derives patterns and themes from the data and thus interprets it (Patton 2002, 460 et seq.). The core of analysis is to understand mechanisms of impact, and create and verify theories of change instead

*Qualitative data collection
requires skillful
facilitation in the field*

of testing hypotheses. For example, a hypothesis on the impact of health microinsurance would state that being insured with the health microinsurance scheme increases the utilisation of health-care services covered by the insurance package. The underlying theory of change is that the financial barriers to accessing health care are reduced through prepayment into the insurance scheme and, thus, insured people do access health care more often. Quantitative methods would create evidence for or against the hypothesised impact, whilst qualitative methods would inquire whether the mechanisms of the theory of change are as expected or not, and why. Whilst doing this, the qualitative researcher tries to understand the meaning and significance given to the phenomenon under study, by the study participants themselves, instead of applying his/her own point of view of what is important and what is not (Creswell 2009, 175–176; Creswell 2007, 147). For example, as Oakley et al. (1998) argue, the target group of a certain intervention might value not only long-term impact but also short-term outcomes of a project, which would not be accounted for in quantitative impact assessments.

Box 2: Comparison of quantitative and qualitative approaches to impact assessment

In a study on the impact of a microcredit programme in Peru, Copestake et al. (2005) compared both quantitative and qualitative approaches to impact assessments and their respective advantages and disadvantages, concluding with a call for combining both approaches whenever possible.

Both approaches [the econometric/quantitative and the interpretative/qualitative approach] have strengths and weaknesses. The econometric approach was based on a larger and more statistically reliable sample. It also offers precise estimates of key impact variables, most importantly household income. However, the approach remains open to selection bias, arising from unobservable differences between client and non-client samples. It was also limited in its inclusion of variables, and in the extent to which differences in impact between different sorts of clients can be measured. Reliability depended upon good survey management and skilled econometric analysis. The interpretative approach was based on a smaller sample, albeit one that was randomly selected from a baseline survey, and cross-analysed against its findings. The range of potential variables covered was more open, and a greater range of differences in impact were highlighted, albeit less precisely. The reliability of impact attribution hinged on the specialist skills of the qualitative researchers, in both interviewing and analysis of the data. In the absence of budgetary constraints then, there is a case for arguing that the two approaches are complementary... However, this is a luxury that most microfinance institutions cannot afford [Copestake et al. 2005, 719].

7.3. Quality criteria in qualitative impact assessment: validity, reliability and generalisability

Because qualitative impact assessments are flexible in their design and there are no universal processes to be followed, ensuring quality by responding to certain criteria is important. There are different sets of quality criteria, partly competing, which derive from different philosophical backgrounds of qualitative research (Patton 2002). Traditional scientific criteria

transfer quality standards from quantitative to qualitative research, as does, for example, Yin (2003). He describes four criteria for quality of research designs from all spheres of research and explains how these criteria can be fulfilled in case study research, but his recommendations are applicable to other forms of qualitative research as well (Yin 2003)¹, [see table 2].

¹ Other scholars argue that these traditional scientific criteria are not applicable to qualitative research since qualitative inquiry comes from a constructivist/interpretative philosophical background and not the traditional postpositivist thinking. For a more extensive discussion see Patton (2002, 542), and Marshall and Rossman (2011, 39 et seq.).

Table 2: Quality criteria in qualitative impact assessment²

Construct validity (correct operationalisation)	Internal validity (causal relationship)	External validity (generalisation to domain)	Reliability (operations can be repeated with same results)
The phenomenon under study has to be thoroughly defined and indicators for operationalisation relevant and justified.	The causal relationship (y is caused by x) inferred from the data has to be thoroughly described and justified.	Qualitative research/ case studies aim at analytical generalisation, i.e., generalising results to a broader theory. To achieve this generalisation, a replication of findings for different cases is necessary and helpful.	Close documentation of the research process is necessary to achieve reliability, each step of the formulation of the research question, the selection of cases, the data collection, and data analysis has to be documented so that others can repeat the procedure and come to the same conclusions.

Internal and external validity are probably the two most contested quality criteria for qualitative impact assessment and will thus be discussed more in detail.

Internal validity and reliability

According to Radermacher, von Arman-sparg, and Chen (2012, 342) non-experimental impact assessment designs have only limited internal validity due to the lack of a control or comparison group to verify the causal relationship between observed impacts and the specific intervention. However, there are qualitative impact approaches involving comparison groups for attributing causality. Roche (1999, 79 et seq.) presents alternatives for control groups to overcome the problem of attribution if these are not available, such as including

non-project respondents, using secondary data and other key informants, and ruling out other explanations than the assumed causal effect. Patton (2002, 479) argues that potential causal linkages should be addressed in qualitative evaluation research as long as it is made clear that these are only speculation and hypotheses. Considering rival explanations is recommended to increase internal validity of qualitative research (Patton 2002, 553; Creswell 2009, 152; Yin 2003, 32).

Another important quality criterion of qualitative impact assessment and qualitative research in general is the triangulation of findings by:

- applying different methods (e.g., verifying interviews by direct observation: do people really do what they say they do?);

2 Compare Yin (2003, 34 et seq.).

- collecting data from more than one source (e.g., insurance policyholders and insurance staff, male and female policyholders);
- involving more than one researcher in data collection and analysis; or,
- analysing data from different theoretical perspectives.

Triangulation can strengthen research findings

The goal of triangulation is to strengthen research findings. When results obtained through one method are mirrored in the results obtained from a different method, the research findings get confirmed. It thus offers a more accurate picture of empirical reality (Patton 2002, 555 et seq.; Roche 1999, 86 et seq.; Creswell 2009, 191).

Contradictions found during this process need to be addressed. They might reveal bias or specific interests and thus might not only strengthen the validity of the findings but provide new insights into the phenomenon as well (Roche 1999, 92). Presenting the findings to the study participants and discussing conclusions with them is not only a good method to increase the validity of the research (Creswell 2009, 191), but also increases the target group's involvement in the study and gives it a voice in the research process

that is intended to be for its own good. The recommendation to the researcher is to be transparent about all steps of the research process when reporting findings and provide any personal or professional information that might have influenced data collection, analysis, or interpretation. The researcher should not underestimate the effect of his/her presence on those under study, as well as on his/herself, and consider these effects and potential own biases when analysing his/her data (Patton 2002, 566 et seq.; Creswell 2009, 192). Possible follow-up actions of the research and potential consequences for the participants (e.g., introduction of a new insurance product or adjustments in the existing product) should be reported to the participants.

Box 3: Triangulation and documentation strategies applied in a qualitative impact assessment study

In their baseline study for an impact assessment of health microinsurance in Pakistan, McGuinness and Mandel (2010) adopted a variety of triangulation and documentation strategies to ensure the validity and reliability of their results.

The baseline Outcomes Assessment employed several methods to enhance the reliability and validity of the results. The research:

- included three separate studies covering all relevant aspects of the problem (the financial landscape, the health-care landscape and consumer perspectives);
- employed multiple data collection methods including interviews (66 total), focus group discussions (32 with a total of 243 participants) and a review of secondary data;
- cross-referenced research questions across the discussion and interview guides to allow for triangulation;
- employed multiple data sources reflecting diverse perspectives and experiences including key informants in various Network agencies, the financial services industry, health-care providers, community organisations, and community members; and
- used different investigators with specialized expertise for each of the three studies.

In addition to this, research procedures included

- using a documented research protocol including the data collection tools; establishing a chain of evidence through preliminary key research questions that are linked through the documented research protocol to the findings and their respective data sources; and
- developing of a research database.

These measures, which included triangulation of data, methods and investigators, establishing a chain of evidence, and documenting the research protocol and all data collected, enhance the study's construct validity and the reliability of its results (McGuinness and Mandel 2010, 26).

External validity and ability to generalise

The ability to generalise qualitative research findings is subject to an ongoing debate that is taking place against the background on the value of generalisation and particularisation in itself, which is seen as a trade-off between the breadth of a study and its depth (Patton 2002, 581). Most qualitative researchers accept the limited generalisability of their findings to whole populations, i.e., the lack of statistical generalisation, arguing that they do not strive for generalisation at all, but see the value in the context-specificity of qualitative research and its ability to investigate a case in-depth (Cresswell 2009, 193). Others aim for a generalisation to broader theories, i.e.,

Qualitative research findings lack statistical generalisation, but can provide an in-depth understanding of a case

to an analytical generalisation in the wording of Yin (2003), or—as a middle course—an extrapolation of findings to cases under similar, but never identical conditions (Cronbach and Associates according to Patton (2002, 584)).

Box 4: Limitations to generalisability for qualitative impact studies

Hietalahti and Linden (2006) are very careful about generalising from their data on the impact of a microfinance programme on livelihoods in South Africa, arguing that their findings would only be transferable to similar socioeconomic settings and even this only in a limited way.

Although many important economic and social impacts of microcredits have been clarified in this study, it is still unable to provide a final answer to the question of the total effect of microcredits on local livelihoods, even in this single case observed in Tzaneen area.

The results are still related only to a limited area, and the data have been gathered within only a limited period of time. It is, however, able to summarise cautiously the difficulties and opportunities that are directed towards microfinance in southern

Africa, in areas that share a similar socioeconomic setting. In this context it can be argued that, when properly designed, microfinance institutions such as The Small Enterprise Foundation have an important contribution to make to women's empowerment (Hietalahti and Linden 2006, 210).

Blanchard-Horan (2006, 140) also highlights the importance of the context of her study on the impact of a Ugandan health microinsurance on treatment-seeking behaviour for malaria, and stresses that "[n]either the intent nor the design of the study involves generalizing these data to a larger population. Nonetheless, for organizations interested in empowering women and improving their health care access this study hopes to raise their awareness to the potential value of health microinsurance."

7.4. Qualitative impact assessment research designs

The research design identifies the timing and frequency of observations to be undertaken in an impact assessment. Basically, we can differentiate between cross-sectional and longitudinal research designs for qualitative impact assessment. Additionally, we will discuss the case study approach.

Cross-sectional designs

Cross-sectional qualitative impact assessment designs collect data only

at one point in time. Depending on whether the intention of the impact assessment is to predict or identify the impact of a given intervention, the data is collected ex-ante or ex-post of the intervention. In the case of ex-ante assessment, it focuses on the present situation and potential changes due to a future intervention; in the case of ex-post assessments, it sets out to identify the present situation and past changes due to an intervention. The involvement of a control or comparison group is possible to compare findings.

Box 5: Example for retrospective, cross-sectional qualitative impact assessment

An example for a cross-sectional, retrospective qualitative design—though in combination with a preceding quantitative survey—is a study undertaken by Copestake et al. (2005) to assess the impact of a microcredit programme (Promuc) in Peru. Here, a representative quantitative baseline study was followed one year later by qualitative in-depth interviews with a smaller sample of clients, focusing on changes experienced during the previous year and reasons for these changes.

...the interviews were designed to elicit from respondents open-ended narrative explanations of changes they had experienced during the past year, and reasons for those changes, from which an experienced analyst could make a reliable assessment of impact attributable to their interaction with Promuc. An explicit part of the research was to pilot a qualitative in-depth interview protocol (referred to as the QUIP) for addressing the attribution problem using this interpretative, rather than a positivist, approach [Copestake et al. 2005, 717].

Cross-sectional research designs are easy to implement and very cost-efficient since they do not require a long-term commitment of participants and data needs to be collected only once. At the same time, they are prone to recall bias (e.g., when respondents need to assess their financial situation from two years before, but either do not remember correctly or their perception is influenced by their present financial situation, they give inaccurate information) and no direct comparison between data from before and after the intervention is possible. Involving comparison groups is beneficial, but might also create problems: e.g., a lack of commitment, since the group does not benefit from the intervention under study, drop out of the comparison group when a similar intervention is conducted in their area,

or spillover effects become apparent (Roche 1999, 79).

Longitudinal designs

To understand the impact of an intervention, it is preferable to conduct longitudinal studies that involve repeated observations of the same group. This is true for quantitative studies trying to quantify the scope and scale of impact, as well as for qualitative studies addressing theories of change, causal chains, and unintended and unexpected impacts. Ideally, the first data collection, often referred to as baseline study, should take place before the start of the intervention to enable a comparison of the situation before and after the intervention. Again, a comparison group can be involved to allow for a double difference comparison.

*Longitudinal studies
are preferred for
understanding the impact of an
intervention*

Box 6: Example of longitudinal qualitative impact assessment design

Hintz (2010) applied a longitudinal qualitative impact assessment design to assess the impact of a life microinsurance product in Indonesia.

Payung Keluarga is a clear development intervention. It was conceived to ameliorate the assumed post-mortem financial crisis of low-asset families, and to prevent a drop in their already low-asset base. Through qualitative-exploratory field research from 2006 until 2008 I investigated if this developmental intention was realized. Research components consisted of a baseline-end line comparison with insured customers, beneficiary interviews, and ancilliary [sic] research components (Hintz 2010, iii).

To prepare the ground for the end line intervention, the guideline for baseline intervention should be structured with end line questions already in mind. A clear theory of change in the design process helps in keeping all intervention rounds aligned to each other. Identical questions posed during different rounds facilitate comparison.

Such longitudinal designs involving more than one round of data collection and, in some cases, even a comparison group, are substantially more expensive and more difficult to conduct than cross-sectional designs. Participants need to be approached more than once and it might be challenging to maintain the comparison group long-term. Dropouts from both the study groups are possible and the researcher needs to consider beforehand how to deal with these.

Case study designs

Case studies are in-depth investigations of selected units (e.g., individuals, households, groups, insurance schemes). Different than other qualitative inquiry, case study research “involves the study of an issue explored through one or more cases within a bounded system” (Creswell 2007, 73), i.e., the case is seen as one example of a specific phenomenon. In other qualitative designs, information is gathered from different sources and then combined and synthesised to answer specific research questions. In the case study approach, each case is regarded as an entity and interpreted and analysed as such, though—of course—cross-case analysis and comparison is also done (Yin 2003).

For impact assessments, case studies are especially useful to explain

presumed causal links and explore what array of impacts an intervention has brought about.³ They can be cross-sectional or longitudinal, i.e., cases can either be studied at one

point of time or over a period of time and can either comprise only one or several cases (Gerring 2007, 39). They rely on a combination of different data sources, such as in-depth interviews, observations, and document analysis, and do not necessarily exclusively use qualitative data (Yin 2003, 85).

³ Yin (2003, 15) gives case studies a distinctive place in evaluation research.

Box 7: Example for case study approach in qualitative impact assessment

McGuinness (2011) chose a case study approach to assess the impact of a health micro-insurance programme in India and interviewed insured as well as uninsured households to compare how these coped with serious cases of malaria.

The key research questions addressed include:

- Does Uplift health microinsurance protect households financially? If so, how?
- Are the out-of-pocket (OOP) costs of healthcare lower for Insured households than for Uninsured households?
- How does the unique community-managed reimbursement process at Uplift influence the financial protection effect of the insurance?

These key questions are addressed in two ways. The first is through a case study which compares the specific experiences of 15 Insured households and 10 Uninsured households to assess whether Uplift-Insured households are financially protected when faced with a serious case of malaria. This case study yielded interesting insights into the experiences and coping strategies of low-income households faced with a serious, widespread disease, but the sample size was too small to make its findings generalizable to the HMF [Health Mutual Fund] population as a whole. In addition to the case study, (...) [we] carried out an analysis of Uplift's claims and financial data (McGuinness 2011, 12).

Table 3: Different qualitative research designs

Research design	Data collection	Advantages and disadvantages	Specifically useful for...	Further issues to consider
Cross-sectional	At one point in time, ex-ante or ex-post the intervention under study	<i>Advantages</i> • Easy to implement and cost-efficient • Fast results <i>Disadvantages</i> • Prone to recall bias (in case of ex-post intervention) • No direct comparison of data from before and after intervention possible	• Exploring unintended and unexpected impacts • Understanding impact observations from other studies or data	• Number and type of participants • Type of data collection tools to use
Longitudinal	At least at two points in time, ideally before and after intervention under study	<i>Advantages</i> • Direct comparison of data from before and after intervention, less recall bias <i>Disadvantages</i> • More expensive and difficult to conduct • Long-term commitment of study participants required	• Understanding impacts evolving over time, such as attitude and behavioural changes	• Number and type of participants • Type of data collection tools to use • Frequency of data collection rounds • Time between data collection rounds • How to motivate participants to commit to the study • How to deal with dropouts
Case study	At one or more points in time (cross-sectional or longitudinal)	<i>See advantages or disadvantages of cross-sectional and longitudinal designs (above)</i>	• Understanding the holistic impact of an intervention	• Number of cases to involve • Type of cases to involve: typical or untypical • Entity to treat as a case, such as individual, household, groups • <i>Specifically for cross-sectional or longitudinal case studies (above)</i>

7.5. Tools for data collection in qualitative impact assessment

In this section, we describe various methods of data collection used in qualitative research. These include interviews, group discussions, observation, storytelling, and participatory methods. In qualitative research, data collection methods are semi-structured and allow the researcher to explore issues in an open-ended manner. Whilst the issues explored are generally around a set of domains of enquiry central to the research question, the exploration is not limited to predetermined themes or topics. These may expand in response to findings in the field.

Key informant interviews

Key informant interviews (KIs) are qualitative, in-depth interviews conducted with individuals who are selected on the basis of their superior knowledge about the subject being investigated. This may be by virtue of their position in an organisation, or because they are a good representative user, or because they are an expert on the issue. The purpose of these interviews is to collect information from people who have a good understanding of the research subject.

were held with experts in the field of microinsurance and field officers from the two non-governmental organisations (NGOs) whose programmes were being studied. As part of his study on understanding the contribution of microinsurance to financial protection, Hintz (2010) conducted KIs with 24 experts from various fields: academics, religion, insurance, microfinance, and politics. These interviews were meant to gather background information on social, cultural, and religious perceptions and behaviours observed amongst customers and beneficiaries and to shed light on the relation of microinsurance to general social changes occurring in Indonesia.

According to Wright and Copestake (2004), more experienced clients can be used as key informants for discussing wider or indirect impact. In order to build trust and minimise response bias,

Key informant interviews collect information from subject experts

In Young's 2006 exploratory study of the impact of microinsurance, KIs

respondents in the interviews need to understand where the researchers are from, what the data is being collected for, and how it will be used. This information should be given to the respondents at the start of the interview.

The KII guide is semi-structured, and the interviewer should probe deeper for each of the themes in the guide. An example of a probe is, “Can you tell me something more about that?” If the interviewer is not clear about the respondent’s answer, he/she can repeat what she has understood so that the respondent can confirm his/her understanding or correct it. Probes can be both planned and spontaneous. It is very important to find out local terminology and use that in the interview to get good responses.

A trade-off exists between questions that act as prompts and open-ended questions. In the former, the researcher introduces new elements or invites responses to a specific topic. Such questions risk not allowing the informant to answer using their own categories, but rather imposing categories based on the prejudices of the researcher and thus tainting the nature of the response. But without some structure, on the other hand, interviews may become unacceptably time consuming or irrelevant (Wright and Copestake 2004, 361).

There is also a trade-off between degree of structure in interview and costs and expertise required for data collection, analysis, and interpretation. When an interview schedule is more structured, the responses are more likely to be comparable than when each interview is allowed to flow relatively freely. A less structured schedule requires a more skilled interviewer to ensure that all the key issues are covered and all the leads thrown up by the respondent in the course of the interview are followed through adequately.

Another important tip mentioned by Wright and Copestake (2004) is to be attentive at the end of interviews. Informants often reveal the most useful information when the formal interview is declared over. Each key informant interview is usually conducted in one or two sittings. During data collection, it is best to capture the responses of the interviewee/participants verbatim without trying to interpret at this stage. If possible, it is advisable to use an audio to record the interview if the respondent permits. Verbatim recording also allows the researcher to pull out quotations—in the words of the respondent—which is an effective technique to illustrate and explain points made in the data analysis and adds to the strength of the researcher’s interpretation.

Some Dos and Don'ts for KILs⁴

Dos

- Familiarise yourself with your questions as much as possible in advance
- Listen and pick up on cues
- Ask follow-up questions based on what you are hearing; use the participants' words where possible
- Encourage participants to share anecdotes and specific experience—avoid generalities
- Allow silences and accept pauses as natural; break the silence only if the respondent seems stuck
- Feel free to laugh and appreciate humour
- Monitor body language
- Catch discrepancies and try and seek clarification

Don'ts

- Ask leading questions (where the question contains hints or leads to answers)
- Make judgmental comments
- Interrupt the respondent or try to control the conversation

Focus group discussions

A focus group discussion (FGD) is a "... carefully guided discussion to obtain

information on a limited number of key questions (3 - 4) through the synergy of social interaction and discourse among participants" (Barnes and Sebstad 2000, 26).

FGDs are appropriate under a variety of circumstances. These include

- when you want to understand motivations and perceptions of clients;
- when you want to stimulate reflection and discussion on client satisfaction or group dynamics;
- when you need information to interpret quantitative data; and,
- when you need information quickly to address an issue, or participatory principles are a priority for the researching organisation.

An FGD is moderated by an individual who asks open-ended questions in a manner that generates discussion amongst the participants. The moderator has to be attentive to pick up leads and cues from the participants' responses so that the discussion yields meaningful depth. The moderator also needs to be fully attentive to group dynamics and to ensure that the discussion does not get hijacked by the agenda of some participants.

The size of the group should be large enough to generate diversity of opinion but small enough to give everyone the opportunity to participate in the

⁴ This section on *Dos and Don'ts* has been adapted from Padgett (2008).

Moderators of FGDs should be fully attentive to group dynamics

discussion. Ideally, FGDs should have six to eight participants, though FGDs are held with as few as three and as many as 15 participants.

Successful FGDs depend on equality and trust amongst participants. If the participants are in hierarchical relationships of say age, gender, or social class, the discussion will be dominated by the socially “superior” participants. It is therefore important to strike a balance between the homogeneity and heterogeneity of the group. Homogeneity, in terms of social class, will ensure that hierarchical relations do not prevent some people from being inhibited. At the same time, the group should comprise persons with diverse views and experiences to generate a useful discussion of the research issue.

In terms of procedure, it is useful to inform the FGD participants in advance about any remuneration they will receive in terms of transportation costs, a lunch, or compensation for their time for participating in the FGD.

An FGD is primarily a group discussion. However, either before the discussion or at the end, it is important to collect some basic demographic information about each of the participants. This is helpful

in terms of getting a sense of their socio-economic background and also useful in case there is a need to contact the participants at a later date. Respondents can fill out a simple form for this. If the FGD participants have limited literacy skills, the forms can be filled out for them.

During the FGD it is important to ensure that all participants are engaged in the group discussion and no one is particularly dominant or silent. This requires the moderator to be watchful of what is happening and of having the skill to make corrections without offending any of the participants.

Sometimes FGDs can be the primary source of research data. An FGD is appropriate when the information being sought is not personal or confidential. For example, one or two persons begin sharing their experience about having or not having insurance and it gets the others also to begin thinking how they are affected by having or not having insurance. Thus, one purpose of an FGD is to stimulate the thinking of participants through group discussion and make them come up with views and feelings that they may not have made explicit even to themselves prior to such a focused discussion on the topic.

Box 8: Example of a qualitative study using focus group discussions

Young (2006) set out to understand the impact of microinsurance and the level of financial protection it provided to members. Since it was an exploratory study, FGDs were the main data source. The two NGOs whose programmes provided the context of the study were FINCA Uganda and Save for Health Uganda (SHU). Amongst the FGDs conducted in the FINCA areas, there was a mix of current members, ex-members, and non-members of the insurance scheme. In the SHU groups, each of the FGDs had persons from discrete categories.

The selection of FGD participants needs to be determined by the researcher based on the research question and the study context. One common problem that can occur in an FGD is when the discussion becomes a series of one-on-one conversations between the moderator and each of

the participants. Instead of facilitating a discussion amongst the participants, the moderator takes on the role of an interviewer and seeks views and experience from each of the participants in turn. This falls short of the objective of the FGD, which is to generate a discussion within the group.

Box 9: Example for a qualitative study using serial individual interviews in a group

In a recent study on the impact of microinsurance, Hintz (2010) consciously used a method of serial individual interviews in a group, explicitly mentioning his reason for using this strategy.

I prefer to speak of group questionings rather than focus group discussions because different from the classical focus group concept, there was hardly any free or loosely moderated discussion amongst the research participants on a given topic [...]. Instead the sessions consisted mostly of a systematic question and answering process between me and the participants (Hintz 2010, 142).

He goes on to explain that he used this method "to achieve significant respondent numbers in a time-efficient way" (Hintz 2010, 143).

Storytelling

Storytelling is an analytical narrative tool that has been used successfully in evaluation over a range of organisational and program contexts. Dialogical, narrative techniques—written, acted, and videoed—are powerful tools to explain change, identify emergent, unpredicted changes, and test organisational assumptions about impact. They also provide space for local voices to be heard in the project evaluations.

The Most Significant Change (MSC) method, for example, is one such dialogical, story-based learning technique (Davies and Dart 2005). Essentially, the process involves the collection of significant change stories emanating from the field level and the systematic selection of the most significant of these stories by panels of designated stakeholders or staff. The designated staff and stakeholders are initially involved by searching for project impact. Once changes have been captured, selected groups of people sit down together, read the stories aloud, and have regular and, often, in-depth discussions about the value of these reported changes and which change they think is most significant. This is especially important in large programmes where there may be multiple levels at which significant change stories are pooled and then selected.



When the technique is implemented successfully, whole teams of people begin to focus their attention on programme impact. As part of a mixed-method approach to learning, MSC can generate hypotheses about changes that took place, which can subsequently be tested using other methods.

Observation (participant and direct)

Another method of choice for qualitative field studies, particularly applied in social anthropology, is (participatory) observations (Malinowski 1944).⁵ Observation as a data collection method is different from KIs and FGDs in that the latter two methods explicitly elicit responses from the interviewee or the FGD participant around some key topics in a short span of time. Both the

⁵ The term *participatory observation* has mainly been coined by Malinowski who was one of the first researchers to use observations as method for anthropological field studies. Compare Malinowski (1944) for more details on the fundamentals to participatory observations.

researcher and the research subject engage in an activity—conversation in this case—to generate data. Observation, in contrast, does not require the research subject to do anything other than continue with his/her normal routines. The researcher gathers data by observing the research subjects in their natural settings. Observation yields data that may not be reported in interviews and FGDs where respondents may not disclose or recall fully.

Observation may be participant or direct, i.e., non-participant. “When one’s concern is the experience of people, the way they think, feel and act, the most truthful, reliable, complete and simple way of getting that information is to share their experience (through participant observation)” (Douglas 1976, 112). Participant observation involves social interaction between the researcher and informants in the milieu of the latter. The objective is for the researcher to study first-hand the day-to-day experiences and behaviour of subjects (Waddington 2004). This method is usually time consuming and needs immersion in the local community to gain their acceptance.

In direct observation, unlike participant observation, the researcher may simply position his/herself as an observer of a situation without taking on a specific role in the research milieu.

Typically, a researcher using observation follows an inductive strategy for generating categories and uses his/her initial observations to formulate a hypotheses. Research ethics require that the observer make full disclosure about the decision to be an observer.

As a research method, observation has risk of reactivity—that is, changes that occur in the behaviour of the observed due to the researcher’s presence. A related issue is the degree of the researcher’s involvement from total participation at one end, to simply observing at the other end. Whilst a high degree of involvement of the researcher may yield rich insights, it runs the risk of researcher fatigue and the researcher becoming so involved that sight of the goal is lost—also known as “going native”. There are hardly any studies in the field of micro-insurance that have used observation as a strategy for data collection.

Box 10: Example for a qualitative study using observation as a supplementary technique

In his recent study described above, Hintz (2010) uses observation as a supplementary technique. He was actually a functionary in the insurance programme he was studying

and was thus ideally positioned to be in the role of a participant observer, since he had a legitimate position in the programme. However, he prefers to call himself an *observing participant*. This is perhaps because his research came after he had already started working in that context; he did not enter the context as a researcher. Hintz (2010) talks of personal observations as something that enriched the other data collection methods such as personal interviews and group questionings.

Participatory data collection methods

In order to avoid or minimise asymmetric situations between the researchers and the participants, other participatory data collection methods can also be used.

Box 11: Community-based oral testimony approach of the Goldin Institute

One example for a participatory data collection approach related to the method of story-telling described above is the community-based oral testimony. This strategy, developed by the Goldin Institute, a non-profit organisation from Chicago, involves people from the community under study in the data collection process (Goldin Institute, n.d.). This approach was used in a study on the impact of a microcredit programme in Bangladesh:

We adopted a strategy known as “oral testimony” which relies on extended semi-structured interviews to let participants tell their own stories in their own words, share their opinions and experiences and convey their own understandings of how development and poverty has transformed the history of their lives and villages. We wanted to take this approach a step further. Often, oral testimony research is coloured by power-dynamics between “researchers” and “subjects”. Within these dynamics, answers to questions are often pre-determined by what each party expects to hear from the other. We decided to address this by inviting microcredit recipients in Arampur, a village in rural Northern Bangladesh, to interview each other about their own experiences with loans. We hoped that the content of these interviews would be shaped by mutual dialogue, rather than by top down agendas and expectations about what we, as researchers, wanted to hear. In order to do this, we trained a group of villagers in basic, qualitative research techniques and invited them to interview their peers and neighbours. The result was open-ended, conversation-style interviews, recorded using digital-audio recorders, in which the interviewees participated in directing the discussion by framing conversations through stories, life experiences, and their own personal histories with microcredit lending organizations. Using this approach we heard what people had to say about microcredit on their own terms (Goldin Institute, n.d.).

A participatory method that could be easily adopted for impact assessments of microinsurance is the Peer Ethnographic Evaluation and Research (PEER) (Hawkins and Price 2000). Originally developed for monitoring the impact of sexual and reproductive health programmes, it is a powerful tool for understanding local interpretations of change and attributing changes to external interventions. PEER involves training members of a target or beneficiary community to conduct research, centred on conversational interviews, within their own social network or peer group under the supervision of an experienced researcher.

In Jamaica, for example, PEER was sequenced with participatory scorecards in urban and rural communities to monitor and evaluate the impact that social policy had on relations between youths and police officers. Whilst the PEER method maintains clearly-defined roles of interviewers and interviewees, participants in

the community-based oral testimony method (described in box 11) interview each other mutually, dissolving the distinction between researchers and subjects. In both methods, interviewers are no longer researchers from outside, but consist of members of the peer group of the participants.

Social mapping

Mapping of social difference and social change by local people became widely known and facilitated from the early 1990s. Chambers (2008, 133) observes a “phenomenal spread” in participatory mapping, through traditional methods as well as through innovations with spatial information technologies.

Although—at least to our knowledge—not yet used for evaluations in the field of microinsurance, mapping has been used effectively during the past two decades for the kinds of social analysis that underpins much qualitative research.

Box 12: Examples for social mapping

Targeted Inputs Programme (TIP) in Malawi

A participatory evaluation of the Targeted Inputs Programme (TIP) in Malawi, for example, aggregated numerical data from group-based mapping of food insecure households as a standardised subcomponent of a flexible participatory process (Barahona and Levy 2003).

Ubudehe in Rwanda

Social mapping has been taken to scale in Rwanda as part of the *Ubudehe* community group activity tradition. Some 15,000 communities across the country have produced and analysed social maps as the basis for household poverty categorisation and analysis, community planning and resource allocation (Joseph 2008).

MAP Kibera in Kenya

In Kenya, MAP Kibera is a social mapping project in Nairobi's largest slum.⁶ One MAP Kibera team, consisting of 13 young people from the community who were trained in open source mapping techniques, created a map around security and vulnerability. This map contained unsafe spaces (i.e., places where drugs and alcohol are consumed), safe spaces (i.e., girl groups, community centres, areas with enhanced lighting), resources (i.e., gender-based violence clinics) and more. This data was loaded to an online free and open source map, then shared with the community in order to prompt community reflection and action and empower community members to engage with local authorities on issues such as street naming and street lighting.

7.6. Conclusion

Qualitative approaches to impact assessment are especially well suited when processes of impact should be explored, identified, and understood in depth. They have the advantage of being open to unexpected findings, susceptible to perceptions of the study group, and flexible in their research design. Nevertheless, qualitative impact designs have been pushed into the background by the increasing popularity of experimental designs (i.e., randomised controlled trials (RCT)).⁷ However, as shown, qualitative impact studies have a value of their

own: besides exploring what *would be* valuable to be measured, qualitative methods can help to understand what *is* measured in quantitative and experimental approaches and open the black box to assess whether the assumed theories of change hold. Furthermore, they can give insight about impact of microinsurance in cases where quantitative methods—like RCTs—are difficult to implement or might not be sufficient (e.g., for cases where the insured event has very low probability of occurring). Both quantitative and qualitative approaches to impact assessment have their strengths and weaknesses and are appropriate for specific research interests. Thus, whenever possible, both approaches should be combined to obtain a full picture of the impact of a microinsurance programme.

⁶ For more information see <http://mapkibera.org>.

⁷ Compare Prowse (2007) and Patton (2008, 440 et seq.) as examples of critical voices on the overemphasis of RCTs and the lack of involvement of qualitative study parts in RCTs.

7.7. Examples of qualitative impact assessments in microinsurance and other fields of microfinance

Example 1

Young, P. (2006): “Microinsurance—Exploring Ways to Assess its Impact.”

Purpose of study

- To explore changes in household financial behaviour and risk coping strategies as a result of having insurance
- To identify and refine indicators to assess microinsurance impact

Issues examined

- How do behaviours, knowledge, skills, and attitude differ between the insured and uninsured?
- How does microinsurance improve the household’s ability to smooth consumption and income?
- How does microinsurance change the ability of households to withstand economic shocks?
- How does microinsurance protect asset base and what is the effect of this?
- Does microinsurance have impact at the level of the enterprise?
- The study also attempts to identify indicators to measure impact.

Context

The study wanted to explore the research questions in life and health insurance. They selected research sites where FINCA Uganda was operating in urban areas and Save for Health Uganda in rural areas.

Research design

Cross-sectional, retrospective

Methods used

Expert interviews with 24 experts

Example 2

Hintz, M. (2010): "Micro-Impact: Deconstructing the complex impact process of a simple microinsurance product in Indonesia."

Purpose of study

To study the social impact of an obligatory credit life microinsurance product

Issues examined

- What happens when microinsurance enters a community and why does this happen?
- Does credit life insurance ameliorate postmortem financial crisis amongst low-asset families and prevent further reduction of their asset base?
- Researcher has a holistic and social perspective, seeking to understand the dynamics of impact processes rather than exactly measure them, and takes a strongly localised approach.

Context

A credit life insurance scheme for low-asset families in Indonesia

Research design

Longitudinal (baseline/end line)

Methods used

Qualitative-explorative

- Baseline/end line comparison of insured members
- Beneficiary interviews
- Group questioning (174 persons in 17 groups for baseline; 139 persons in 19 groups for end line)
- Heavy reliance on triangulation
- Researcher as observing participant

Example 3

Copestake, J., H. Gosnell, D. Musona, G. Masumbu, and W. Mlotshwa (2001): "Impact monitoring and assessment of microfinance services provided by CETZAM on the Zambian Copperbelt: 1991–2001."

Source: Copestake et al. (2002)

Purpose of study

To understand the causation of impact of microfinance

Issues examined

- Changes in intra-household relations
- Changes at individual, business, household, and community level

Context

Impact assessment in a microfinance organisation on the Copperbelt in Zambia

Research design

Longitudinal (combination of closed questionnaire for baseline and qualitative interviews for end line)

Methods used

- In-depth interviews
- Group discussions

Reliability

- Attribution of impact was based primarily on coherence of respondents' own accounts of causal chains linking their membership of credit groups to issues examined.
- Plausibility of respondents' accounts was checked with direct observation and discussions with loan officers and other group members.

Example 4

Wright, K. (2001): "Women's participation in microcredit schemes: evidence from Cajamarca and Lima, Peru."

Source: Copestake et al. (2002).

Purpose of study

To understand the social processes affecting microcredit outcomes

Issues examined

- Microenterprise activities and motivation for taking out credit
- Changes in the family and microenterprise
- Decision making and male and female roles in the household
- Perceptions of and aspirations for the future

Research design

Cross-sectional, retrospective

Context

Microfinance programme in low income settings in urban and rural Peru

Methods used

Detailed interviews in the form of guided conversations

Example 5

McGuinness, E. (2011): "A Fine Balance: A Case Study of the Client Value of Health Micro-insurance — Uplift I.A."

Purpose of study

To investigate whether the Uplift health microinsurance program created financial value for the members.

Issue examined

- Does Uplift health microinsurance protect households financially?
- If so, how?
- Are the out-of-pocket payments of insured households lower than those of uninsured households?
- How does the unique community-managed reimbursement process at Uplift influence the financial protection effect of the insurance?

Context

Health microinsurance programme in India

Research design

Household case studies, cross-sectional and retrospective

Methods used

- In-depth interviews
- Short questionnaires
- Documentary analysis (claims data)

Example 6

McGuinness, E. and J. Mandel (2010): "Assessment of Health Microinsurance Outcomes in the Northern Areas, Pakistan - Baseline Report."

Purpose of study

To assess whether health microinsurance programme reduces vulnerability of households to risks and how.

Context

Baseline report for impact assessment of health microinsurance programme in Pakistan

Research design

Cross-sectional (baseline for later research)

Methods used

- Focus group discussions
- Semi-structured interviews
- Structured interviews
- Key informant interviews
- Secondary data and monitoring reports

Example 7

Copestake, J. et al. (2005): "Monitoring the Diversity of the Poverty Outreach and Impact of Microfinance: A Comparison of Methods Using Data from Peru."

Purpose of study

To assess poverty outreach and impact of a microcredit programme in Peru (comparison of quantitative and qualitative impact assessment methods).

Context

Microcredit programme in Peru

Research design

- Qualitative: cross-sectional (retrospective changes in last year)
- Quantitative: longitudinal with control group

Methods used

- Qualitative: in-depth interviews
- Quantitative: household survey

Example 8

Hietalahti, J. and M. Linden (2006): "Socio-economic Impacts of Microfinance and Repayment Performance: a Case Study of the Small Enterprise Foundation, South Africa."

Purpose of study

To gain a better understanding of how microcredit projects impact on rural women's livelihood structures, and how they can strengthen women's welfare.

Context

Microcredit programme in South Africa

Research design

Case study of microcredit programme

Methods used

Semi-structured interviews

Example 9

Solomon, Y. et al. (2002): "Outreach, Impact, and Sustainability of Informal Banking: A Case Study of the Ouelessebougou-Utah Alliance Microenterprise Program in Mali."

Purpose of study

To determine if this project is truly reaching people traditionally excluded from formal financial institutions, namely women and the poor.

Context

Microcredit programme in Mali

Research design

Case study of microcredit programme

Methods Used

- Individual interviews
- Focus group discussions
- Informal discussions and documents

Example 10

Goldin Institute (n.d.): "Community-Based Oral Testimony: A Different Approach to Knowledge."

Purpose of study

Learn about people's experiences with microcredit loans.

Context

Microcredit programme in Bangladesh

Research design

Cross-sectional

Methods used

In-depth, semi-structured interviews conducted by local members of the community

Example 11

Blanchard-Horan, C. (2006): "Malaria Treatment-Seeking Behavior in Kisiizi, Uganda: Health Microinsurance among Engozi Groups."

Purpose of study

To examine how health microinsurance influences malaria treatment-seeking behaviour in Uganda.

Context

Health microinsurance programme in Uganda

Research design

Cross sectional household case studies ("case-comparison study")

Methods used

In-depth, semi-structured interviews with open and closed-ended questions

References

- Appleton, S. and D. Booth. 2005.** Strong fences make good neighbours: Survey and participatory appraisal methods in poverty assessment and poverty reduction strategy monitoring. In: *Methods in Development Research: Combining Qualitative and Quantitative Approaches*, ed. J. Holland and J. Campbell, 119-135. London: ITDG.
- Barahona, C. and S. Levy. 2003.** How to generate statistics and influence policy using participatory methods in research. IDS Working Paper no. 212, Institute of Development Studies, Brighton, England.
- Barnes, C. and J. Sebstad. 2000.** Guidelines for microfinance impact assessments. Discussion paper from the proceedings of the CGAP 3 Virtual Meeting, October 18-19, 1999. Assessing the Impact of Microenterprise Services (AIMS). Washington, D.C.
- Blanchard-Horan, C. 2006.** Malaria Treatment-Seeking Behavior in Kisiizi, Uganda: *Health Microinsurance among Engozi Groups*. Ph.D., Walden University. ABI/INFORM Complete; ProQuest Dissertations & Theses A&I (304935281). <http://search.proquest.com/docview/304935281?accountid=10218>.
- Booth, D., J. Holland, J. Hentschel, P. Lanjouw, and A. Herbert. 1998.** Participation and combined methods in African poverty evaluation: Renewing the agenda. Social Development Department, Department for International Development (DFID), London. *Issues*. February.
- Chambers, R. 1995.** Poverty and livelihoods: Whose reality counts? *Environment and Urbanization* 7(1): 173-204.
- . 2008. *Revolutions in Development Inquiry*. London: Earthscan.
- Copestake, J., P. M. Dawson, J. P. Fanning, A. McKay, and K. Wright-Revollo. 2005.** Monitoring the diversity of the poverty outreach and impact of microfinance: A comparison of methods using data from Peru. *Development Policy Review* 23(6): 703-723.
- Copestake, J., S. Johnson, and K. Wright. 2002.** Impact assessment of microfinance: Towards a new protocol for collection and analysis of qualitative data. Working Paper 23746, Improving the Impact of Microfinance on Poverty: Action Research Program. University of Sussex, Imp-Act. <http://purl.umn.edu/23746>
- Creswell, J. W. 2007.** *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Thousand Oaks, CA: SAGE Publications.
- . 2009. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: SAGE Publications.
- Davies, R. and J. Dart. 2005.** The 'Most Significant Change' (MSC) technique: A guide to its use. *Monitoring and Evaluation NEWS*. <http://mande.co.uk/special-issues/most-significant-change-msc/>
- Della Porta, D. and M. Keating. 2008.** How many approaches in the social sciences? An epistemological introduction. In: *Approaches and methodologies in the social sciences: A pluralist perspective*, ed. D. Della Porta and M. Keating, 19-39. Cambridge and New York: Cambridge University Press.
- Douglas, J. D. 1976.** *Investigative social research: Individual and team field research*. Beverly Hills, CA: SAGE Publishing.
- Faust, J. 2010.** Wirkungsevaluierung in der Entwicklungszusammenarbeit. *Politik und Zeitgeschichte* 10:41-46.
- Gerring, J. 2007.** *Case study research: Principles and practices*. New York: Cambridge University Press.
- Goldin Institute. n.d.** Community-based oral testimony: A different approach to knowledge. <https://www.goldininstitute.org/images/CBOT.pdf>
- Hawkins, K. and N. Price. 2000.** *A peer ethnographic tool: For social appraisal and monitoring of sexual and reproductive health programmes*. Swansea Centre for Development Studies: University of Wales.

- Hietalahti, J. and M. Linden. 2006.** Socio-economic impacts of microfinance and repayment performance: A case study of the small enterprise foundation, South Africa. *Progress in Development Studies* 6(3): 201-210.
- Hintz, M. 2010.** *Micro-impact: Deconstructing the complex impact process of a simple microinsurance product in Indonesia*. Saarbrücken: Südwestdeutscher Verlag für Hochschulschriften.
- Hulme, D. 2000.** Impact assessment methodologies for microfinance: Theory, experience and better practice. *World Development* 28(1): 79-98.
- Joseph, S. 2008.** Creating spaces for citizen participation in self governance, poverty analysis, local problem solving and sector/district planning. Unpublished report.
- Kabeer, N. 2003.** Assessing the 'wider' social impacts of microfinance services: Concepts, methods, findings. *IDS Bulletin* 34(4): 106-114.
- Malinowski, B. 1944.** *A scientific theory of culture and others essays*. Chapel Hill: The University of North Carolina Press.
- Marshall, C. and G. B. Rossman. 2011.** *Designing Qualitative Research*. Los Angeles, CA: SAGE Publications.
- McGuinness, E. 2011.** A fine balance: A case study of the client value of health microinsurance — Uplift I.A. Washington, D.C.: Microfinance Opportunities. http://www.fsassessment.umd.edu/publications/pdfs/MFO_A_Fine_Balance_India_July2011.pdf
- McGuinness, E. and J. Mandel with H. Korda and A. Tayyab. 2010.** Assessment of health microinsurance outcomes in the northern areas, Pakistan—Baseline report. Financial Services Assessment Project, Univ. of Maryland, College Park. http://clients.caiters.com/wordpress/wp-content/themes/mfo/docs/Health_Assessment_Microinsurance_Outcomes_Pakistan.pdf
- Oakley, P., B. Pratt, and A. Clayton. 1998.** Outcomes and impact: evaluating change in social development. INTRAC NGO Management & Policy Series 6, Oxford.
- Padgett, D. K. 2008.** *Qualitative methods in social work research*. Los Angeles, CA: SAGE Publications.
- Patton, M. Q. 2002.** *Qualitative research & evaluation methods*. London and Thousand Oaks, CA: SAGE Publications.
- . 2008. *Utilization-focused evaluation*. Thousand Oaks, CA: SAGE Publications.
- Prowse, M. 2007.** Aid effectiveness: The role of qualitative research in impact evaluation. Background Note, Overseas Development Institute (ODI), U.K. <http://dspace.cigilibrary.org/jspui/handle/123456789/25170>
- Radermacher, R., K. von Armansperg, and T. Chen. 2012.** Impact analyses of micro health insurance. In: *Handbook of Micro Health Insurance in Africa*, ed. H. J. Rösner, G. Leppert, P. Degens, and L. M. Ouedraogo, 335-357. Wien: LIT Verlag.
- Roche, C. 1999.** *Impact assessment for development agencies: Learning to value change*. Oxford: Oxfam.
- Solomon, Y., B. Ballif-Spanvill, C. Ward, and A. Fuhriman. 2002.** Outreach, impact, and sustainability of informal banking: A case study of the Quelessebougou-Utah Alliance Microenterprise Program in Mali. *Journal of Developing Societies* 18(4): 290-314.
- Waddington, D. 2004.** Participant Observation. In: *Essential guide to qualitative methods in organizational research*, ed. C. Cassell and G. Symon, 154-164. London and Thousand Oaks, CA: SAGE Publications.
- Wright, K. and J. Copestake. 2004.** Impact assessment of microfinance using qualitative data: communicating between social scientists and practitioners using the QUIP. *Journal of International Development* 16(3): 355-367.
- Yin, R. K. 2003.** *Case study research: Design and methods*. 3rd edition. Thousand Oaks, London, New Delhi: SAGE Publications.

Young, P. 2006. Microinsurance – Exploring ways to assess its impact. From the Assessing the Impact of Innovation Grants in Financial Services project. Microfinance Opportunities, Univ. of Maryland, College Park. http://microfinanceopportunities.org/docs/Microinsurance_Ways_To_Assess_Impact_Young06.pdf

The case of mixed methods for impact evaluation in microinsurance

Howard White



8.1. Introduction

The preceding chapters discussed different approaches for assessing the impact of microinsurance, ranging from experimental and quasi-experimental (i.e., quantitative) approaches to qualitative approaches. However, these approaches should not be considered competing methodologies but rather complementary. **A combination of different methods, in particular a combination of quantitative methods together with qualitative ones, provides insights strengthening the policy relevance of impact assessments.**

Qualitative studies are employed as exploratory studies, preceding the quantitative work in order to explore topics to assess with quantitative analysis. However, there are many other ways of mixing methods, as shown by the following examples. Increased income is a commonly intended effect of crop microinsurance for insured farmers, since theory suggests that allowing farmers to take riskier but higher return investment decisions will result in higher income. This effect can be measured with quantitative methods, such as a randomised controlled trial (RCT). But what if a quantitative study applied to a microinsurance scheme finds that there is no difference in income between treatment and control group? A possible explanation for this finding could be that there simply is no causal relationship

Combining qualitative and quantitative methods can strengthen policy relevance of studies

between microinsurance and income. However, there exist various other potential and plausible explanations: for example, it could have been that no seeds were available for purchase other than the ones farmers used before. In this case, lack of options would be the reason for stagnant income rather than a missing causal relationship between microinsurance and income under the right conditions. Hence, a succeeding study would unpack the causal chain for the farmer's investment choices and explain the lack of increase in income. And, so, the insurance scheme can be adapted accordingly, and has also generated learning effects for other schemes.

However, qualitative methods alone can lead to unsatisfying results. Take, for example, a study with focus group discussions assessing the impact of a health microinsurance scheme. The participants might report that—despite being insured and health-care costs being at least partially covered by insurance—they still cannot seek sufficient treatment and that expenditures

on health remain very high. In this case, a quantitative study would be useful to measure the spending (total and out-of-pocket) on health care both for a treatment and a control group.

Both the above examples make the case for mixed methods: combining quantitative with qualitative analysis in order to assess the impact of microinsurance schemes.

8.2. What are mixed methods?

There are two parts to the definition of mixed methods: 1) methods, i.e., range of data collection processes, and 2) mixed, i.e., the combination of these diverse data into a single analysis. For most research teams, achieving a genuine mix of methods is a difficult challenge.

As Creswell and Plano Clark (2007) suggest in their definition of mixed methods research, it is not only the mix of methods, but also the underlying philosophical assumptions that determine this kind of research:

Mixed methods research is a research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative data

in a single study or series of studies. Its central premise is that the use of quantitative and qualitative approaches in combination provides a better understanding of research problems than either approach alone (Creswell and Plano Clark 2007, 5).

Whilst the examples above suggest combinations of quantitative and qualitative approaches and the term mixed methods is usually understood as this, the distinction between quantitative and qualitative can become difficult to maintain once engaged in actual field work. Participatory methods can generate numerical data while responses to quantitative surveys such as on subjective perceptions or views on priority projects may enter the qualitative domain. A more useful definition of the term mixed methods is analysis using data generated from different data collection processes (see table 1).¹

Mix data from different data collection processes

¹ See Bamberger et al. (2010, 3) for a discussion around this point.

Table 1: Data from different kinds of data collection instruments

Data collection instrument	Possible uses
Project documents	Describing intervention and elaborating programme theory
Expert knowledge (academic papers or interviews with experts)	Elaboration of programme theory
Academic literature on intervention type and region of intervention	Forming evaluation questions
Principal investigators exposure to field (with some structured components)	Contextualisation for study, understanding of causal linkages
Beneficiary focus groups	Identifying priority and possible unintended outcomes
Structured surveys	Statistical analysis of the counterfactual Statistically representative presentation of the factual

However, just using both qualitative and quantitative data collection instruments is not sufficient to be considered a mixed methods approach. In the case of impact evaluation, a further useful distinction is between factual and counterfactual analysis. A theory-based impact evaluation will mix (combine) both factual and counterfactual analysis. The counterfactual analysis is the analysis of attribution, that is, measuring the impact by comparing treatment and comparison groups, using experimental or quasi-experimental designs. In contrast, factual analysis does not rely on a comparison group, but is simply describing what happened in the treatment area. Such factual analysis is an important part of a theory-based impact evaluation. The factual analysis may be either quantitative, such as a targeting analysis of who benefits from the scheme, or qualitative, such

as focus group discussions of alternative risk coping mechanisms to understand the demand for microinsurance. Table 2 shows these categories. Many impact evaluations focus only on the categories described in the top right cell, which limits their ability to give policy relevant conclusions.

Theory-based
impact evaluation
combines factual
and counterfactual
analysis

Table 2: Categories of analysis

	Factual	Counterfactual
Quantitative	Analysis of impact questions in treatment group	Experimental or quasi-experimental analysis using comparison groups
Qualitative	Understanding of working of the causal chain	Constructing of small <i>n</i> counterfactuals

8.3. What are the potential advantages of mixed methods?

Mixed methods strengthen impact evaluation designs in the following important ways:

- Studies should be driven by issues (questions), not methods. Having a range of methods at hand means that all questions can be addressed, not just those amenable to particular methods. Some evaluation questions need quantitative data, and some need qualitative data.
- Quantitative analysis can be strengthened in a number of ways by using qualitative methods, such as data on context to inform survey design for quantitative data collection and casting light on the interpretation of quantitative results (see examples in section 8.5).
- Vice versa, qualitative analysis can be strengthened in a number of ways by using quantitative methods. This is true for both factual and counterfactual analysis, supporting qualitative results by quantitative

measures. For example, unpacking the causal chain (as in theory-based impact evaluation) by qualitative analysis often requires answering a number of quantitative evaluation questions as well, or at least having numbers on hand helps to do this.

- Purely quantitative impact evaluations are sometimes criticised as having strong internal validity but weak external validity (e.g., Leeuw and Vaessen 2009; Cartwright 2007). Mixing methods provides more context for the intervention and so a better understanding as to which settings the results may be generalised.

8.4. How mixed methods may be used

Mixed methods can be used in the following three ways, as described by Carvalho and White (1997):

1. Integrating methodologies. Combining quantitative and qualitative work in part of the evaluation, this can be both concomitant and sequential, and for the latter, both

kinds of studies can succeed or precede each other. As in the example given above, a qualitative study may succeed a quantitative one in order to explain the results from the quantitative part and to unpack the causal chain, in particular if unexpected results occur. However, a qualitative study may also precede a quantitative one in order to explore potential impacts and outcomes. Very often, microinsurance schemes are different in practice than planned on paper. Conducting quantitative and qualitative studies concomitant to each other can help with understanding the context and how the intervention works in practice at field level. Thus, they can strengthen each other's results. Take, for example, a health microinsurance scheme with insurance for outpatient treatment at particular local medical practitioner or in a particular clinic. Quantitative methods may show no increase in health care utilisation for the treatment group and, indeed, higher total costs of health care. However, qualitative methods, such as focus group discussions, might reveal dissatisfaction and lack of trust with the insurance's practitioner. Hence, the insured go to other practitioners, without making use of the insurance. In combination, these data reveal valuable insight into microinsurance in practice.

2. Confirming, refuting, enriching, and explaining the findings of one approach with those of the other. This is possible when applying the concept of triangulation, by which different methods support the same conclusions (confirming), or do so whilst adding more understanding (enriching). Triangulation is an important and often used concept in mixed methods research. Such triangulation makes the study findings more convincing. However, there may be cases when data conflict with one another (refuting), in which case further work is needed to resolve or understand this conflict. Take, for example, a health microinsurance programme. Asking about different illnesses by using a quantitative, structured survey may show that health has increased within the treatment group. However, at the same time, results of focus group discussions of insured or of key informant interviews with physicians or hospital staff reveal that health has worsened. On the one hand, qualitative samples might easily be biased in their coverage, with a more formal sample survey giving a more representative view. On the other hand, quantitative surveys might neglect important issues of the causal chains. Hence, complementary studies are needed to explain this apparent contradiction in findings.²

2 For more on this please see Pluye et al. 2009.

3. Merging the findings of the two approaches into one set of policy recommendations. Whilst qualitative work can identify possible problems in an intervention, representative structured surveys can better capture how widespread these problems are. But, the importance of storytelling in conveying policy messages is well known, with qualitative data being a good source of stories on which to base proposed policies. Hence, it can be useful to draw on both approaches to make policy recommendations.

8.5. Examples of mixed methods in microinsurance

If understood and used in the way described above, mixed methods can be a useful approach providing insights into many issues related to the impact of microinsurance. The following examples show this usefulness:

- Morsink (2012) applied mixed methods research in an empirical study to investigate and test the question of why low-income households from rural Filipino communities demand natural disaster rehousing microinsurance and whether microinsurance has an effect on poverty reduction for this socio-economic group. For this analysis, focus group discussions (FGDs) to assess village characteristics were

conducted with participants of different attributes such as gender, age, and social status. These FGDs preceded quantitative household surveys, whilst focus group discussions with insured and uninsured were conducted alongside the surveys. In addition, administrative data was used to solve issues with causality. The results show, amongst other things, that microinsurance can reduce the negative influence that a shock has on future economic growth.

- Although focusing on qualitative methods, Hintz (2010) complemented his explorative study of a credit life microinsurance pilot in Indonesia by quantitative methods. Qualitative interviews with Muslim beneficiaries were complemented with additional baseline and end line surveys with customers, as well as with additional research with members of other (functional) groups, like credit group leaders or loan officers. The study reveals that the intended developmental impact was very low—labelled “micro-impact”—but shows the complex interplay of the insurance product with the sociocultural context, leading to extensive social impact. Amongst others, crowding-out effects regarding traditional forms of family assistance were identified, as well as a possible inflation of funeral costs. However, there

was also evidence that financial literacy, as well as peace of mind of the insured, was increased by microinsurance, although, for some insured, this feeling of security decreased on account of religion and superstition.

- The “Client Math” approach, designed within the framework of the Microinsurance Learning and Knowledge (MILK) project by the MicroInsurance Centre, is based on a mixed methods approach, aimed at assessing the value of microinsurance for the poor.³ Based on a carefully developed theory of change, focus group discussions are used for refining the hypothesis and for deepening the understanding of the functioning of the product and the according value chain. This qualitative study part is then the basis for developing a sampling strategy (taking into account specific groups, e.g., by age or gender) and subsequent quantitative interviews. One distinctive feature of the Client Math approach is that participants for the treatment group are chosen after experiencing a financial shock, rather than waiting for a shock to happen. This has certain advantages for assessing the value of the insurance scheme, in particular for schemes covering low-frequency risks. However, because of this sampling approach, and the

fact that the number of participants both for treatment and control group is relatively low (usually including up to 30 participants for each), the approach does not allow for statistical evidence. Hence the approach cannot determine effects in a causal sense, but can rather be understood as complementary to statistical and, especially, experimental methods. By providing numbers, comparing them for claimants and non-insured, and taking into account contextual information, the Client Math approach can provide valuable insight into potential gains that microinsurance can have for the poor. Therefore, it can help to deepen the understanding of the value that microinsurance schemes may have. This does, in particular, hold for low-frequency risks for which impact of microinsurance is difficult to achieve with statistical methods. This approach has, for example, been used for gaining insight into the value of property microinsurance in coastal Columbia or catastrophe microinsurance in Haiti (Magnoni and Poulton 2013; Magnoni and Budzyna 2013).

8.6. Meeting the challenges to mixing methods in mixed methods studies

Mixed methods are frequently lauded in the social sciences, and the fathers of social science, such as Durkheim

3 Compare Magnoni, McCord, and Zimmerman 2012 for this and the following section.

and Marx, readily combined quantitative and qualitative data. But as different disciplines have specialised, the use of methods has grown apart. Economics—and in the United States, political science and sociology—have taken the path of mathematical modeling and advanced statistical analysis. This divergence of paths creates problems for agencies wishing to commission rigorous impact studies of their programmes. Those with the necessary skills in counterfactual analysis do not generally have skills in qualitative analysis or mixed methods. Moreover, motivation to adopt mixed methods is not facilitated by the need to publish in high-ranked journals in their discipline, which mostly requires technical sophistry rather than practical insights.

Across both Europe—especially the UK—and South Asia, there is a strong tradition of development studies, which has always strived to achieve interdisciplinary research. But the reality has been one of multidisciplinary research, with parallel studies. Achieving a true mixed methods impact evaluation

design is thus a serious challenge. Here are some pointers toward achieving that outcome.

- Establish very clear vocabulary on the evaluation questions, requiring an inception report which identifies the data to be used for answering those questions.
- Include a range of skills and experience in the evaluation team. Even more important is ensuring that the lead investigator is engaged in all aspects of the study and has the capacity to do so. Having other team members who can also bridge studies will help.
- Work with the whole team on the development of the causal chain and identification of underlying assumptions, requiring them to indicate what evidence they can bring to bear on which questions.
- Include a mix of skills in the external review panel. Ideally, the experts would have high professional standing combined with practical experience in the policy/programme domain.

An evaluation team with a diverse skills set and a team leader familiar with different methods are needed for a true mixed method study

8.7. Conclusions

This chapter has discussed how different methods can be mixed for impact assessments in microinsurance and how the application of mixed methods research can contribute to better insight and higher validity. Application of different methods—quantitative and qualitative, factual and counterfactual—in parallel and sequentially can be valuable. Often, qualitative studies are used in an exploratory way, being employed as preparation for quantitative, counterfactual studies (i.e., experimental or quasi-experimental ones). The range of possible and valuable combinations of qualitative and quantitative methods is much wider. This range should be fully assessed in order to deepen the insight into potential impacts and outcomes of microinsurance to the welfare of the poor. By doing so, not only can insight be gained ex-post, but learning processes can be initiated in order to improve other schemes. Moreover, by combining different methods, insights regarding the generalisation of successful schemes can be made, which is of particular importance in practice when conducting a pilot project, transferring schemes to another context, or when scaling-up. Therefore, applying mixed methods for impact assessments in microinsurance is crucial in order to gain insight into how microinsurance schemes can affect the lives of the poor.

References

- Bamberger, M., V. Rao, and M. Woolcock. 2010.** Using mixed methods in monitoring and evaluation: Experiences from international development. Policy Research Working Paper 5245, Poverty and Inequality Team, Development Research Group, World Bank, Washington, D.C. http://www-wds.worldbank.org/servlet/WDS-ContentServer/WDSP/IB/2010/03/23/000158349_20100323100628/Rendered/PDF/WPS5245.pdf
- Cartwright, N. 2007.** Are RCTs the gold standard? *Biosocieties* 2:11–20.
- Carvalho, S. and H. White. 1997.** Combining the quantitative and qualitative approaches to poverty measurement and analysis: The practice and the potential. World Bank Technical Paper no. 366, Washington, D.C. http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2000/02/24/000009265_3971031092544/Rendered/PDF/multi_page.pdf
- Creswell, J. W. and V. L. Plano Clark. 2007.** *Designing and conducting, mixed methods research*. Thousand Oaks, CA: SAGE Publications.
- Hintz, M. 2010.** *Micro-impact: Deconstructing the complex impact process of a simple microinsurance product in Indonesia*. Saarbrücken: Südwestdeutscher Verlag für Hochschulschriften.
- Leeuw, F. and J. Vaessen. 2009.** Impact evaluations and development: NONIE guidance on impact evaluation. Washington, D.C.: The Network of Networks on Impact Evaluation (NONIE). http://siteresources.worldbank.org/EXT/OED/Resources/nonie_guidance.pdf
- Magnoni, B. and L. Budzyna with J. Bauchet. 2013.** Doing the math. *Catastrophe Insurance in Haiti*. MILK Brief #15. MicroInsurance Centre. http://www.microinsurancecentre.org/milk-project/milk-docs/doc_details/933-milk-brief-15-qdoing-the-mathq-catastrophe-insurance-in-haiti.html

Magnoni, B., M. J. McCord, and E. Zimmerman. 2012. What is "Client Math"? MILK Brief #9, MicroInsurance Centre. http://www.microinsurancecentre.org/milk-project/milk-docs/doc_details/842-milk-brief-9-what-is-qclient-mathq.html

Magnoni, B. and D. Poulton. 2013. Doing the math. Property microinsurance in coastal Colombia. MILK Brief #18, MicroInsurance Centre. http://www.microinsurancecentre.org/milk-project/milk-docs/doc_details/949-milk-brief-18-qdoing-the-mathq-property-microinsurance-in-coastal-colombia-.html

Morsink, K. 2012. Weathering the storm: The demand for and impact of microinsurance. PhD Thesis, University of Twente (The Netherlands).

Pluye, P., R. Grad, A. Levine, and B. Nicolau. 2009. Understanding divergence of quantitative and qualitative data (or results) in mixed methods studies. *International Journal of Multiple Research Approaches* 3(1):58-72.

Further reading

Garbarino, S. and J. Holland. 2009. Quantitative and qualitative methods in impact evaluation and measuring results. Governance and Social Development Resource Centre Issues Paper. Department for International Development (DFID), U.K. <http://www.gsdrc.org/docs/open/EIRS4.pdf>

O'Cathain, A., E. Murphy, and J. Nicholl. 2008. The quality of mixed methods studies in health services research. *Journal of Health Services Research & Policy* 13(2):92-98.

Pluye, P. 2012. Les méthodes mixtes pour l'évaluation des programs. In: *Approches et pratiques en évaluation de programme*, ed. V. Ridde and C. Dagenais. Seconde édition sous presse. Montréal: Presses de l'Université de Montréal.

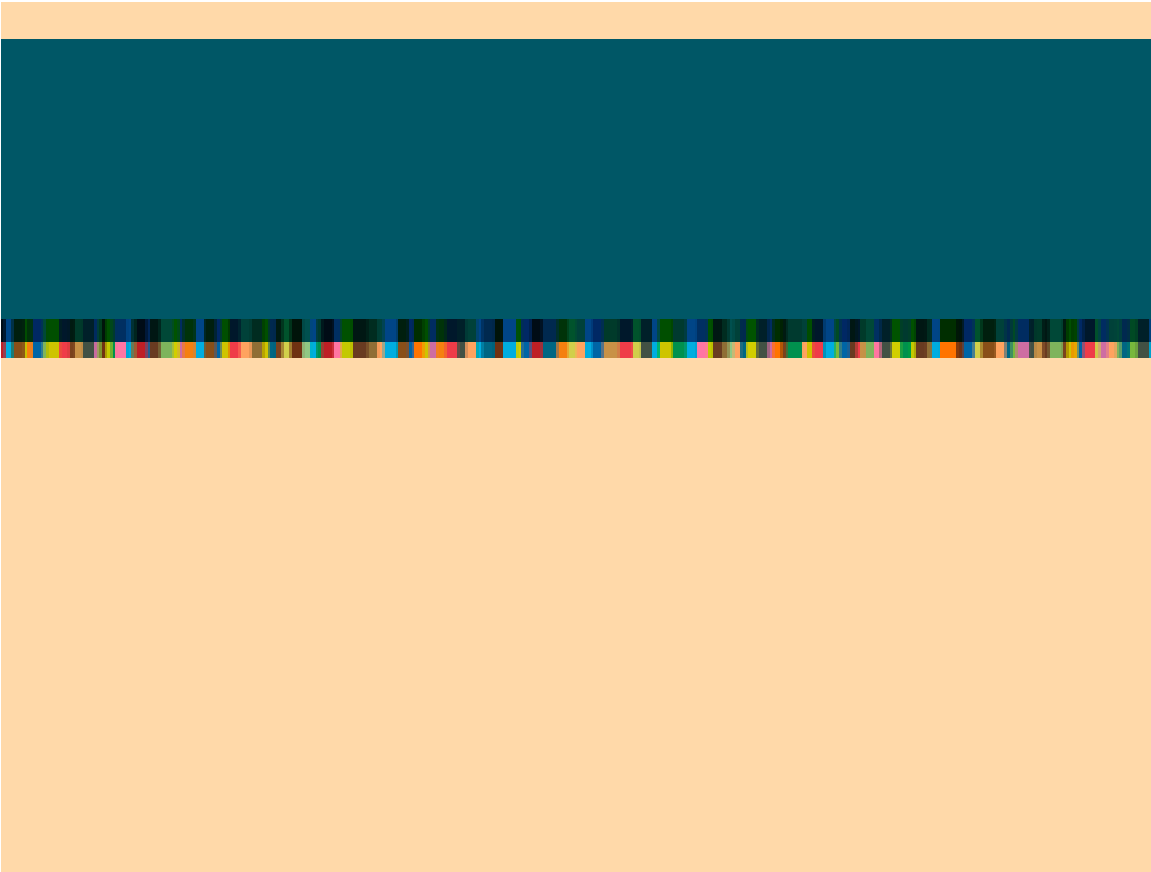
Ridde, V., S. Kouanda, M. Yameogo, K. Kadio, and A. Bado. 2012. Why do women pay more than they should? A mixed methods study of the implementation gap in a policy to subsidize the costs of deliveries in Burkina Faso. *Evaluation and Program Planning* 36(1):145-152.

Tashakkori, A. and C. Teddlie. 2010. *SAGE handbook of mixed methods in social & behavioral research*. 2nd ed. Los Angeles, CA: SAGE Publications.

White, H. 2002. Combining quantitative and qualitative approaches in poverty analysis. *World Development* 30(3) (March):511-522.

———. 2008. Of probits and participation: The use of mixed methods in impact evaluation. *IDS Bulletin* 39(1):98-109.

———. 2011. Achieving high quality impact evaluation design through mixed methods: The case of infrastructure. *Journal of Development Effectiveness* 3(1):131-144.





Planning and Conducting Impact Assessments in Microinsurance

Defining indicators for impact assessments

Conor Doyle and Pradeep Panda



9.1. Introduction

In recent years, extensive efforts have been made to improve and standardise the methods by which the impacts of microinsurance programmes are measured. In particular, the increased use of the scientifically rigorous randomised controlled trial (RCT) technique in programme evaluations has led to a welcome increase in the accuracy of recorded impacts on many development interventions. This has raised the methodological bar against which evidence is judged. However, relatively little time has been spent considering what should be measured to assess impact. Indicators may be accurately measured under a rigorous methodology—but if a poor metric of the intended outcome is used, or important impacts are not examined, the evaluation may generate misleading evidence and malformed policy.

This chapter provides guidance to evaluators on how to design a set of indicators that will effectively and holistically measure the impact of their microinsurance intervention. Section 1 provides some background on how to determine the events to monitor and the types of indicators to use on a microinsurance evaluation. Section 2 provides guidance on formulating high-quality indicators, and provides a list of the characteristics to which effective indicators should conform.

Section 3 presents an overview and typology of the forms of indicators that have been commonly used to date in microinsurance evaluations.

9.2. Choosing research areas and methods

A performance indicator is any metric used to measure or assess the effect of an intervention or some aspect of an intervention. Researchers look at performance indicators to answer key evaluation questions. An insurance scheme, like any intervention, generates a huge number of effects: processes run, opinions changed, behaviours affected. A vast amount of data could potentially be captured. However, limits of time and cost dictate that an evaluation can capture only a sliver of this information. The key problem faced by the evaluator is choosing a set of indicators that illuminate the most important elements of the intervention and will form a cohesive story of the whole from a limited number of pieces. This is like taking

WHAT should
be measured?
---> choosing indicators

3 levels of indicators:
process,
outcome,
impact

pictures of a family holiday: the holiday might be two weeks long, but you can only take snapshots of a few key moments, and hope that you choose these in such a way that they give a good outline of the course of events. Evaluators have the additional handicap of having to decide much of what they are going to measure in advance; so, it is really like trying to decide on the photos you will take before your holiday begins!

In order to make educated decisions in advance as to what areas and types of information are likely to be worth examining, the evaluator must be aware of the range and formats of indicators available. The following discussion outlines some key qualities of indicators that evaluators can use to determine what to measure and how.

Processes, outcomes, and impacts

A first set of distinctions between indicators refers to the level at which an indicator is measured. Indicators can usually be placed into one of three levels:

- **Process indicators** are metrics that examine the extent to which

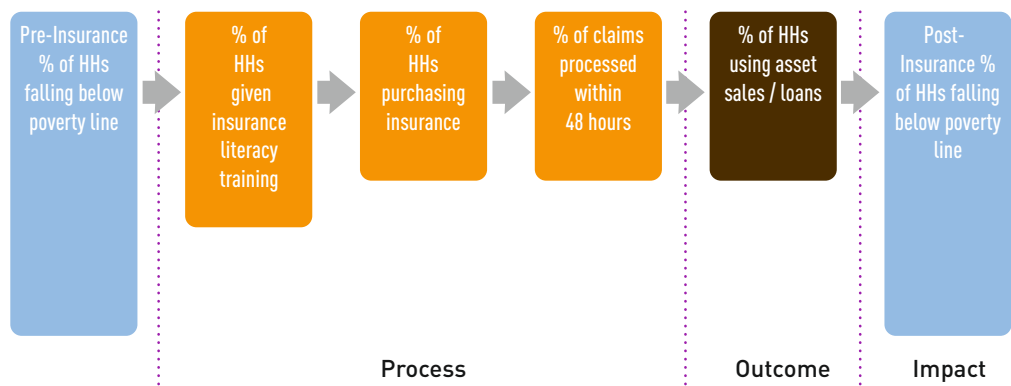
a scheme has run in the manner expected or the success of a scheme in reaching operational targets. They examine the inputs made into the operation of the scheme, i.e., internal aspects of scheme administration that can be considered to be largely under the control of the staff working on and managing the insurance scheme. In the insurance context, process indicators might include: the proportion of a local population who have been contacted by an insurance education campaign, the proportion of insurance clients to whom claim forms have been pre-distributed, or the proportion of claims processed via correct formal channels.

- **Outcome indicators** examine the extent to which the operation of the scheme has been turned into a tangible and direct short- or mid-term result. They often examine external changes in the behaviour or situation of those covered by the scheme that are a direct result of the scheme being in operation. In a health insurance context, the average reduction in out-of-pocket (OOP) expenditures on inpatient (IP) care is a commonly-used indicator in the outcome category.

- **Impact indicators** examine the extent to which the scheme in question has led to long-term effects. These are often external effects of the scheme that are indirectly facilitated, but not directly caused by it. To return to the health insurance example, an impact indicator might be the reduction in the number of individuals falling below the poverty line due to the costs of an inpatient care episode.

It should be apparent that these three levels of indicators are linked to each other: processes need to run in order to create an outcome, and outcomes need to happen in order to generate an impact. Figure 1 gives an example of the relationship between process, outcome, and impact indicators. Imagine evaluators are examining a health microinsurance scheme in a developing country that provides cover in the case of high-cost health events.

Figure 1: Example of a relationship between process, outcome, and impact indicators for households (HHs)



The ultimate development aim of the scheme is to reduce the incidence of poverty (percentage of households (HHs) falling below the poverty line) resulting from selling household assets or taking on high-interest loans in order to finance high-cost health events. A baseline survey is taken to measure the percentage of households that have fallen below the poverty line due to high-cost health

events in recent months. Following this, five key indicators are defined. Evaluators begin by focusing on the operation of the scheme itself and the extent to which it has run as planned. They define three key process indicators: firstly, the extent of an insurance literacy training programme administered to the local population; secondly, the proportion of the population convinced to take up insurance;

and, thirdly, the percentage of claims processed quickly enough that members do not need to resort to alternate sources of financing for inpatient care (defined here as 48 hours). The cumulative effect of these processes is expected to produce an outcome. In this scheme, it is hoped that a well-functioning insurance scheme will lead to a drop in the proportion of households that rely on methods such as asset sales or high-cost loans to finance health care, and this is explicitly defined as an indicator. Finally, the outcome can lead to an impact. In this case, it is hoped that the rate at which households are falling below the poverty line is reduced as a result of the insurance scheme, and this figure is also explicitly defined as an indicator.

When determining the set of indicators they will be monitoring, evaluators can improve the quality of their research by systematically selecting indicators from across the processes, outcome, and impact levels, following the *theory-based impact evaluation*, or TBIE, approach (Weiss 1997a; White 2009). In TBIE, an underlying theory of change is postulated prior to the intervention, outlining the causal steps by which a programme is expected to generate an impact and the assumptions underlying each causal step. Indicators are then selected systematically such that the assumptions made at each step in the theory of change are tested.

Essentially, the TBIE approach allows evaluators to audit the logic underlying their intervention. We can unpack the chain of events and examine not only what impact has been achieved, but also whether it has been achieved in the expected manner and, if not, why not. Applying TBIE methods also increases the ability of evaluators to suggest improvements in how an intervention is targeted or implemented and to examine the extent to which impacts might be replicated in different contexts. In addition, TBIE can improve to the ability of managers and funders to provide strategic oversight: failures that can be traced to a process level may be solvable via specific management or business process interventions, whilst failures at an outcome or impact level can indicate changes in targeting or scheme design are required. As Rogers (2009) notes, TBIE is an appropriate evaluation strategy for every type of intervention.

Use theory of change
for selecting
indicators



Unintended consequences, missing indicators, and confirmation bias

Any intervention leads to multiple effects on the environment in which it is undertaken. These can be both positive and negative. However, evaluations are generally undertaken on the basis of a theory that posits that a particular intervention will have a beneficial effect on a particular aspect of the participant's environment. Thus, one common trap into which evaluators fall when selecting indicators is to choose only those that examine the intended positive consequences of the scheme. This is an example of the well-known problem of confirmation bias, an inherent psychological tendency:

...when testing an existing belief, to search for evidence which could confirm that belief, rather than for evidence which could disconfirm it (Jones and Sugden 2001, 59).

The evaluation sets out to test a theory, and chooses a set of indicators which will help elaborate upon the extent to which the expected impacts are generated or not. Indicators that capture other effects are often excluded from the evaluation design. This problem of missing indicators can have severely deleterious consequences on the findings of an evaluation.

To demonstrate, let us take a hypothetical example of a health insurance scheme that generates an unintended adverse impact. Imagine a scheme has been set up by a primary health-care clinic in a developing country to cover its clients in the surrounding area. The scheme administrators hope that by using insurance as a prepayment mechanism, they will increase the rate of primary care usage amongst the population they are serving. The indicator they select to measure whether or not the scheme has met its targets is the total number of visits made to the clinic each month. In the months after the scheme is started, the total number of visits per month triples. By the indicator being measured, the scheme is adjudged to be a great success. However, a closer examination of the scheme reveals deep flaws. The clinic is in a remote rural area, and finds that it is not able to attract additional qualified staff to meet the increased demand for services. With increasing numbers of patients each week, the existing doctors have had to cut the average number of minutes they spend per consultation. As a result, there is a higher rate of misdiagnosis and a fall in the real quality of care provided. Moreover, as the rate of misdiagnosis rises, the number of unnecessary additional visits increases and this further drives up the visits-per-month indicator. In fact, the lower the quality of service becomes, the more successful the scheme is

adjudged to be! However, as neither visit times nor number of second visits are monitored indicators, these important negative effects go unrecorded, and a scheme with substantially harmful effects is recorded as a major success.

To avoid generating erroneous conclusions, evaluators should include some indicators that examine unintended consequences of the scheme. In the example, a careful evaluator could have easily included indicators of quality of care (e.g., average length of consultation, rate of re-diagnosis on second visits, or patient satisfaction with care received) that would have led him/her to identify the problems experienced. The key problem remains: having to determine in advance what is important to monitor. One possible approach is to use *negative program theory* (Weiss 1997b). This is essentially an extension of the TBIE approach, under which a causal framework is used to predict some of the broader consequences of an intervention, including potential adverse impacts, and indicators are inserted to capture these. However, this approach will not suffice to identify the truly unexpected or unpredictable effect. A second approach is to incorporate rounds of participatory qualitative research (Chambers 2009), both before and after an intervention, allowing programme participants to divulge problems they expect to face, or have faced. Unintended consequences

Include indicators
for unintended
consequences

identified pre-intervention can either be corrected or quantified via additional indicators, whilst those identified post-intervention can be made the subject of new subevaluations. Obviously, unintended adverse effects are more serious than unintended beneficial ones, and more care should be taken to search for them.

Quantitative, qualitative, and hybrid indicators

Further salient features to be considered when choosing indicators are the differences between quantitative, qualitative, and hybrid indicators:

- *Quantitative indicators* are those capturing objective realities, i.e., verifiable facts that are external to the views of those taking part in the research. These indicators can be expressed as numbers, directly or indirectly. For example, the number of individuals who have visited a hospital is a quantitative indicator: it is directly expressible as a number, and that number is independent

of the opinion of the researcher or hospital manager. Whether or not an individual has joined an insurance scheme is also a quantitative indicator: it is an objective fact, and can indirectly be rendered as a number (e.g., 1=joined /0=not joined).

- *Qualitative indicators* are those capturing the subjective conceptions of those taking part in the research regarding the world around them. The hospital manager's opinion as to why the number of individuals visiting his hospital has changed is a qualitative indicator: it depends solely on this individual's interpretation of the trends he/she has experienced. Purely qualitative data is generally best expressible in the form of textual statements and descriptions.
- *Hybrid indicators* fall between the other two classes. These are indicators in which people are asked to rank, scale, or codify their opinions and attitudes. Whilst the indicators gained are based on individual respondents' perceptions and are not subject to external confirmation or disconfirmation, they also code this information numerically into a number of pre-defined categories, allowing opinions to be compared and aggregated across individuals. For example, the hospital administrator might be asked to rank the importance of a list of reasons as to why the number of patients has increased, thus organising primarily qualitative

responses into a numeric (though not strictly "quantitative") framework.

The table below presents the relationship between quantitative, qualitative, and hybrid indicators. Quantitative and qualitative indicators examine fundamentally different dimensions of information. To the extent that this reflects potentially differing epistemological standpoints, academics have sometimes identified "paradigm wars" between adherents of the two schools of research (Johnson and Onwuegbuzie 2004) or even an implicit "incompatibility thesis" positing that these methods should not and cannot be mixed (Howe 1988). As a practical matter, this should not be considered to be the case. As quantitative and qualitative methods examine differing types of information, they can often be applied together to generate a significantly richer set of knowledge for interpreting the impact of an intervention. A common simplification used by evaluators is to say that quantitative research answers the question of "what happened?", whilst qualitative research answers the question of "why did it happen?" Choosing a set of indicators that utilises a mix of quantitative, qualitative, and hybrid indicators designed to provide mutually supportive insights can potentially deepen the explanatory power of an evaluation.¹

¹ This "mixed methods" approach to research is a developing stream within evaluation studies—see Pronyk et al. (2008) for an example in a developing country.

Table: Relationship between quantitative, qualitative, and hybrid indicators

Quantitative	Qualitative	Hybrid
Amount of money spent on IP services	Patient's feelings as to whether money was well spent	Patient's rating of quality of care along a 5-point scale
Number of people who have visited a hospital	Administrator's opinion as to why number of patients has changed	Ranking of a set list of reasons why number of patients might have changed
Number of individuals who have joined an insurance scheme	Member's statement as to why he/she joined the insurance scheme	Ranking of a set list of reasons as to why a scheme has been joined

Evaluators should avoid the fallacy of assuming that either quantitative or qualitative indicators are inherently more accurate than the other. Most indicators rely on the collection of information from a respondent and suffer from some common potential problems in the information reported to investigators. In any branch of research, questions that are incorrectly phrased (either on the survey sheet or orally by the investigator) can lead to incorrect responses being given. Alternately, questions may be phrased correctly, but understood incorrectly by respondents, again leading to inaccurate responses. Even when a question is fully understood, respondents may recall the required information incorrectly or with some bias. Finally, respondents may feel some incentive or some pressure to misreport their answer. **A combination of these factors may also apply.** Whilst quantitative data is generally exact, and qualitative data is usually rich; neither is necessarily accurate.

Combining indicators

The qualities of different types of indicators, and the ways in which different types of indicators can be systematically combined to help increase the usefulness of the information gathered, have been outlined to help guide the evaluator's choices. The categorisations presented are intended to be: 1) mutually exclusive within themselves, and 2) unrelated among themselves. Thus, an indicator must be one

*Quantitative,
qualitative, and
hybrid indicators:
it is helpful
to adopt a mix*

*Indicators must be specific:
what is to be measured, how, and when?*



of three types (process, outcome, or impact), and it must be one of three methods (quantitative, qualitative, or hybrid). However, any combination is possible: for example, a process indicator could be quantitative, qualitative, or hybrid.

9.3. Defining high-quality indicators

The next task is to design the specific indicators to be used. This section looks at the finer points of calibrating and specifying performance indicators. Having followed the advice of section 1, all indicators under consideration should already possess two key characteristics. Firstly, they will be important, in the sense that they address issues of key concern to the evaluation at hand. Secondly, they will be consistent, in that they have a clear theoretical link to the insurance intervention being applied. Specifying channels of impact in advance allows evaluators to avoid accusations of ad hoc data mining. What follows is a checklist of additional qualities of well-designed indicators: the more of these qualities your indicator has, the clearer and more useful the data yield on that indicator is.

Specific

In order to be usable, indicators must be defined in such a way that it is clear exactly what is to be measured, how, and when: that is, they must be made highly specific. Refer again to the key impact indicator used in figure 1, *percentage of households falling below the poverty line*. This may seem clear, but in order to become a usable indicator, several pieces of additional information need to be specified:

- *Time period*: The time period over which the value of the indicator is measured must be defined. In this example, the evaluators have a range of choices: they could examine the percentage of households falling below the poverty line over the course of one year, five years—even a decade. Definitions of time period should generally be set so they are long enough to contain a sufficient number of observations to make meaningful measurements, but short enough to not overburden the capacity of survey respondents to accurately respond. For example, a common rule of thumb in health microinsurance evaluations is to examine inpatient care over the

last year: inpatient care is so infrequent that one year of observations is required to provide a reasonable number of occurrences, but extending the time period any further can impair accurate recall of details.

- *Unit of measurement:* The units in which an indicator is to be measured must be defined. In this example, we must determine a unit against which we can determine whether a household is poor. A number of options are available, for example: gross income in local currency, the average calories consumed per day, the respondent's score on a synthetic index of housing conditions, and personal asset ownership.
- *Unit of aggregation:* When individual observations on an indicator are calculated as an aggregate figure over a group of people (a unit of aggregation), standard definitions of group inclusion and exclusion need to be determined to avoid any inaccuracy in data collection and indicator calculation. In this example, the unit of aggregation is the "household", so, a standard definition of household members is required. For example, evaluators must specify whether family members who are currently travelling are part of the household, and, whether a group of relatives living in two separate houses, side-by-side constitutes one household or two.

- *Other operational definitions:* Indicators may require further definitions of any specific terms used. In this example, the poverty line needs to be defined. Operational definitions should refer to commonly used standards to the extent possible. Importantly, the data source for an indicator should also be specified in advance.

An actionably specific formulation of this indicator is presented in Schneider and Hanson (2006, 17), who examine microinsurance schemes in Rwanda. This indicator can be rendered as:

Percentage of households in which per capita consumption expenditures in the last two weeks are above the government standard poverty line of RWF 4920 per adult per week, but drop below this line once OOP medical expenditures in the last two weeks are subtracted.

This indicator employs the following definitions:

- *Time period:* last 2 weeks
- *Unit of measurement:* the local currency, Rwandan Francs (RWF)
- *Unit of aggregation:* household, i.e., all those people who live together in a dwelling for more than 6 months per annum, and eat from the same kitchen
- *Poverty line:* Rwandan government defined standard of RWF 4920, per adult per month.

- *Consumption expenditure*: household self-report of value in RWF of expenditures over last 1 month, including both food and non-food items, captured via a standardised survey instrument
- *Out-of-pocket medical expenditure*: household self-report of value in RWF of expenditures in last month on inpatient care, outpatient diagnostic services, drugs, testing, transport and fee co-payments made to any medical provider.

The outcome measure is now fully specified: the exact methodology by which it will be measured is made explicit, as are all necessary definitions, and the source of the data that will be used.

Simple

Some indicators require less extensive or complex information to be used in their calculation than do others. **Simplicity in terms of information required to calculate an indicator has several benefits: it minimises the scope for errors at the data collection stage; it increases the transparency and communicability of measured outcomes; and, it may enhance the trust that external readers place in the accuracy**

of an evaluation's results. It may also save time in data collection.

To demonstrate, consider first a complex indicator, commonly used for evaluations of health microinsurance products covering inpatient care:

Incidence of annual inpatient hospital expenditure greater than 10% of annual income, after reimbursement

This indicator is highly specific, important within the context of the trial, and theoretically consistent with the effects of an inpatient health insurance product. However, calculating the value of the indicator for any one individual requires the input of a large amount of complex data, each piece of which is potentially subject to errors. Firstly, an estimate of each household's income must be generated. This is a difficult and time consuming task (Tschirley and Rose 2000), especially in a developing country environment in which individuals have multiple, unpredictable income streams. An exact estimate requires multiple detailed questions on each income stream of each household member. An error or misreport in the answer to any one of these questions will make the calculated value of the indicator inaccurate. Next, total hospital bills must be calculated. Again, this may require detailed questioning of individuals who have been hospitalised, subject to

Keep indicators simple!

potential recall errors, recall biases, and/or intentional misreporting. Moreover, the hospitalised individuals may not be available, in which case other people will report what they remember of the episode; this may introduce further errors. Finally, details of insurance reimbursement need to be collected from the insurance provider. The final indicator is then calculated by dividing one potentially inexact and/or misreported estimate (total HH income) by a second potentially inexact and/or misreported estimate (net hospital expenditures). Evaluators would be right to be wary of the accuracy of this indicator.

In contrast, an indicator used on an evaluation by Rao et al. (2009, 6), provides a lesson in simplicity:

Percentage ill respondents who sought outpatient department (OPD) treatment outside the home, last one month

Again, the indicator is specific, important, and consistent. However, the room for error in the calculation of this indicator is much lower. Calculating the value of the indicator requires asking only two basic questions: 1) have you been sick this month? and 2) did you visit any health-care provider to get help? The information asked for is simple to recall and easy to report. Even in cases where a survey respondent is

answering for someone else, we can be fairly certain that he will report correctly. As a result, both the evaluator and the external observer will be able to place a high level of confidence in the accuracy of measurement of this indicator.

Minimum coefficient of variance

This is a characteristic that relates specifically to quantitative indicators. When examining a quantitative indicator, evaluators generally want to know whether the average value of the indicator has changed over the course of the intervention. Intuitively, if observations on an indicator are very spread out around the average, then the indicator is likely affected by very many different forces and events. If we find a change in the mean of the indicator post-insurance, it is difficult to say with confidence that this change is due to the insurance, and not some other random event. Conversely, if an indicator is very tightly clustered around a certain value, insurance is introduced, and we then find that the indicator has moved a considerable amount, we can be fairly certain the insurance intervention and not any other factor caused this change.

Different quantitative indicators have differing intrinsic levels of variability. In order to maximise their ability to detect changes in an indicator, the evaluator should choose indicators that

Amongst similar indicators choose those with low variability

are less dispersed. That is, amongst a set of indicators that yield qualitatively equivalent information on the effect of an insurance package, the evaluator should choose the indicator with the lowest level of variability. This will increase his/her ability to detect changes in the average of the indicator caused by insurance.

As an example, when evaluating an insurance product that covers inpatient care in India, the average number of visits to hospital and the average amount spent on hospitalisation care might be considered equivalently good indicators of the amount of hospital care received. Formally, statisticians use a metric called *standard deviation* to measure how dispersed observations are around the average.

The more spread out the observations are around the average, the higher the standard deviation. Indian national statistics show that, in rural households, the average number of hospitalisations per year is 1.17, with a standard

deviation of 0.53, and that the average amount of money spent per household on hospitalisations is rupees (Rs) 3,775, with a standard deviation of Rs 9,457. It is impossible to compare the standard deviations directly, as they are in different units: it makes no logical sense to try to assess which is smaller: 0.53 visits or 9,457 rupees. To make the standard deviations comparable, we calculate the coefficient of variance (CV). This is simply the standard deviation divided by the average. The CV for number of hospitalisations is $0.53/1.17 = 0.46$. However, the CV for expenditures is $9457/3775 = 2.51$. It will be to the evaluator's advantage to choose number of hospitalisations as his key indicator: its smaller CV indicates that it is less inherently variable and that it will be easier to discern statistically significant changes in its value. At the indicator design stage, evaluators should use secondary data to determine which of the indicators under consideration have a relatively lower CV. More advanced studies may wish to take on formal power calculations to ensure that changes in their key indicators are likely to be detected.

Comparable indicators enable comparison amongst different schemes

Comparable

If an evaluator is interested in areas of scheme performance that have been investigated for other schemes, he/she should choose indicators that are comparable to those which have been used on these other schemes. The main advantage of choosing comparable indicators is that it enables the evaluator to answer a broader set of questions about their scheme: not merely how well it has worked, but how well it has worked in comparison to similar schemes undertaken elsewhere. A second benefit is that it increases the transparency and, potentially, the credibility of the evaluation. In the context of quantitative indicators, comparability means, at a minimum, using indicators calculated in the same way as those used on other schemes. Strictly, comparability further implies that the data used for calculating the value of an indicator is collected in the same way across studies, by using the same survey instruments.

- Extensiveness of service usage
- Intensiveness of service usage
- Volume of service usage
- Needs-based usage
- Financial protection
- Household socioeconomic status (SES)
- Equity

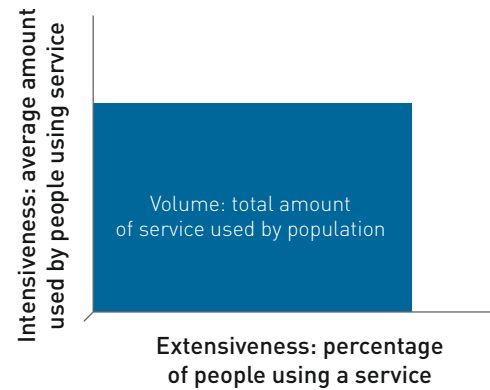
The first three classes are the most studied aspects of microinsurance schemes and are related as shown in figure 2. Extensiveness measures look at how much of the population uses a service. This is graphed along the horizontal axis. Intensiveness measures look at how much service is used by those availing of it. This is graphed along the vertical axis. Total volume indicators examine the total amount of service use by the population as a whole. Total usage is graphed by the orange box, and can be thought of as being the product of (extensiveness x intensiveness). The other four classes of indicators are not related to each other.

9.4. Commonly used indicators in microinsurance evaluations

This section outlines the classes of indicators most commonly used in microinsurance evaluations. Seven broad classes of indicators are discussed:

For each class of indicator, a description of the general form of the indicator is provided along with some examples of how it has been used in practice. A discussion of the interpretation and limits of each class of indicator is provided, along with details on the complexity of the data that needs to be collected for calculation and the problems generally

Figure 2: Relationship between extensiveness, intensiveness, volume of usage



found in data collection. The vast majority of microinsurance schemes that have been subject to an impact evaluation are health microinsurance (HMI) schemes. The review process for this chapter identified 36 papers examining the impact of health microinsurance schemes and only six papers examining non-health microinsurance schemes. The majority of examples discussed are drawn from health literature. However, many of the indicators discussed can be applied equally well to other types of insurance. As far as possible, descriptions of how indicators might be applied to non-health microinsurance schemes are provided.

Extensiveness of service usage

An indicator of extensiveness of service usage measures the proportion of individuals who have used a service.

It quantifies how widespread usage of services has been rather than how deep usage has been. In its purest form, an extensiveness indicator is calculated as:

$$\frac{\text{Number of people using service at least once}}{\text{Number of insured people}}$$

The result is expressed as a percentage. To appreciate what is being measured, consider a population in which the extent of outpatient (OP) service usage is being examined. Before insurance, 30% of people go to the doctor each month, each making only one visit. Imagine that insurance causes an increase in the extent of usage, with 90% of people now visiting the doctor once a month. The value of the indicator will triple, fully capturing the three-fold increase in extent of service usage. Now, alternately, imagine that after insurance is launched, the same 30% of people continue to visit the doctor each month, but increase the number of visits they make from one to three. Even though the number of visits triples, the extent of usage does not increase: the percentage of people using OP services at least once remains at 30%.

There would be no change in the value of the indicator, and measuring impact using only an extensiveness indicator would lead to the erroneous conclusion that insurance had not had any effect. This is not to say extensiveness indicators are poor measures; it is merely a warning that, like all of the indicators

discussed in this section, they capture only one aspect of impact, and evaluators should take care to understand the limits of the conclusions they can draw from a single measure.

Extensiveness indicators are the most studied form of indicator in microinsurance evaluations. They have two great advantages. The first is the relative simplicity of the data needed for their calculation. Data on these indicators is typically gathered via a household survey. Only very simple information needs to be gathered to calculate the value of the indicator. Survey respondents of any age or education level can typically remember whether or not they have been to the doctor this month, or had any surgeries this year. Also, it is likely that even if the survey respondent is not the person who used the service, they will be able to accurately state whether other people in their household had used a particular service or not. Estimates gained may be more reliable and contain less reporting errors than those requiring respondents to provide more detailed information.

The second major advantage of extensiveness indicators is versatility. Extensiveness indicators can be tailored to provide an array of highly informative measures by appropriately varying the time period, subpopulation, and definition of “service” used in calculations. They are thus of broad application.

Some examples of specific applications of extensiveness indicators from the HMI literature serve to illustrate the versatility of this class of indicator:

- *Percentage of people using OP/IP services at least once.* This is the single most commonly applied indicator in evaluations of microinsurance schemes, and the broadest possible measure of the extent of use of outpatient and inpatient services. The time period over which usage of OP service is counted can be varied, e.g., last 3 months (Polonsky et al. 2009), or last 12 months (Ranson 2001). For IP services, it is typically set at one year (Thornton et al. 2010; Diop et al. 2006). The following indicators can be thought of as restricted versions of this indicator.
- *Percentage of men/women aged 16-59 using OP service, last one year.* This indicator is used by Trujillo, Portillo, and Vernon (2005) to study the extent of service usage amongst non-elderly men and women. By defining and separately studying different subclasses of the insured people over which the indicator is measured, more detailed and specific results are generated than with a more general formulation of the indicator.
- *Percentage of people attending modern OP services for treatment of fever, last two weeks.* In this indicator, used by Franco et al. (2008), the definition

of service is restricted, providing highly detailed information on the extent of usage of a very specific service of interest.

- *Percentage of deliveries in a modern facility.* The most commonly used indicator of the extent of maternity-care service usage is the percentage of all deliveries that have taken place in a facility defined as modern (see Dror et al. 2005; Diop et al. 2006; Smith and Sulzbach 2008). Calculations are restricted to the subgroup pregnant women.
- *Percentage of pregnant women making any prenatal care (PNC) visits.* This is the second most commonly used indicator of extent of use of maternity care (see Diop, Sulzbach and Chankova 2006; Thornton et al. 2010). Franco et al. (2008) and Smith and Sulzbach (2008) define *service* more strictly as having made four or more PNC visits.

As the examples demonstrate, these indicators have been widely applied within the HMI literature. Service usage is a topic of primary concern on health insurance schemes, where insurance provides direct incentives for using health services when an adverse event occurs. With other forms of insurance, it may also be useful to examine extent of service usage where insurance provides incentives for using particular services in the recovery from an adverse event. For

example, in crop insurance, evaluators might wish to examine the percentage of farmers repurchasing seeds from a commercial dealer after a failed harvest. In livestock insurance, evaluators might look at the percentage of households purchasing at least one replacement animal after a livestock disease outbreak.

Intensiveness of service usage

Intensiveness indicators examine how deep or intensive usage of a service has been amongst the insured, rather than how widespread usage has been. Intensiveness indicators are calculated using the formula:

$$\frac{\text{Total number of usages of a service}}{\text{Number of people using a service}}$$

The outcome is a rate, expressed as number of usages per person. This is a pure intensiveness indicator: it examines solely how much of a service is used by those availing of it. Returning to the example of a population in which 30% of people make one visit to the doctor per month before insurance is introduced and three visits per month afterwards, this increase in intensiveness of care will be fully captured by this indicator, which will triple in value. However, a change in extensiveness only will have no effect on this indicator: if insurance instead caused usage to expand so that 90% of people went to the doctor each month,

but all continued to make only one visit, there would be no change in the value of the indicator. Any change in the total number of service usages caused by a change in the extent of service usage (measured in the top half of the fraction) is matched by an equal increase in the number of people using a service (measured in the bottom half of the fraction). These two changes automatically cancel out each other, leaving an indicator that only changes in response to changes in the intensity of service usage.

Intensiveness measures involve more complex reporting than do the extensiveness indicators discussed above. Survey respondents must remember, at a minimum, the exact number of times a service was used rather than simply the fact that the service was used. This is especially difficult to accurately report when the respondent was not the person using the service. An intensiveness indicator may thus be more prone to errors caused by inaccurate reporting than an extensiveness indicator.

A pure intensiveness indicator has been used only once in a HMI evaluation: Criel, Van der Stuyft, and Van Lerberghe (1999) examined the number of days that people, admitted as inpatients, spent in hospital. As it is measured only over those people who are actually admitted as inpatients, it

can be thought of as a pure indicator of intensiveness in which each day represents one service usage.

Despite the reservations above, intensiveness indicators are of only middling complexity in terms of data required, and provide huge scope to define time period, subpopulation, and definition of service, making them as highly versatile as extensiveness indicators. They are recommended as a valuable and underexploited form of indicator, which evaluators should consider incorporating into their research designs.

Volume of service usage

Indicators of the volume of service usage are based on a count of the total number of visits made to, or uses made of, a certain service. Therefore, all such indicators respond to changes in both the extent and the intensity of service usage: whether more people use a service, or the same people use a service more often, or both, the total count of how often the service is used will increase. Whilst primarily applied for counting visits to health-care providers on HMI interventions, the concept of counting visits can be extended to usage of non-health services: total number of visits to livestock or seed dealers might be useful measures of responses to livestock or crop insurance, for example. Usage rates can be applied to any microinsurance

evaluation in which the insurance may lead to a change in the intensity of use of some service.

In its most raw form, a volume measure may simply be a count of how much of a service has been used as a whole. Two examples found in the literature are:

- *Total admissions per health centre area.* Used by Criel, Van der Stuyft, and Van Lerberghe (1999), this indicator counts the total number of admissions in a number of different health centres.
- *Number of deliveries at health centre, last one year.* Utilised in Diop, Schneider, and Butera (2000) in an evaluation of microinsurance schemes in Rwanda, this indicator counts the total number of deliveries that had taken place in different health centres. The authors also counted the number of postnatal care visits made and the number of vaccinations administered in the past year.

Much more common than simple counts are volume indicators that express usage in per person terms, i.e., measures that divide the total number of visits by the total number of insured people. This type of measure is often termed *utilisation rate*. As discussed, the time period over which visits are counted, the definition of service, and any subpopulations of interest can be defined by the investigator to suit his/

her needs. Examples of these indicators in use include:

- *Number of OP/IP visits per person.* A very commonly used indicator, e.g., Dror et al. (2009) and Schneider and Hanson (2006) where evaluators simply count the number of visits made to OP or IP providers by insured people in a certain period of time (respectively, two years and two weeks), and divide this by number of insured people. This is a very broad indicator, using a broad definition of services and not subdividing the population under examination into any groups.
- *Number of surgeries per person.* There are two examples of how definitions can be tightened to give more detailed information on specific aspects of service usage from Aggarwal (2010) and Criel, Van der Stuyft, and Van Lerberghe (1999). Both looked at the rate of surgical procedures, or certain types of surgical procedures, amongst the insured population, thus restricting their definition of service to only a subset of all IP care.

Collecting the information needed to calculate volume measures of service usage is of medium difficulty: the same caveats apply as those described for intensity measures. Evaluators should again be aware of the limitations of this class of indicators. **Total usage**

Total usage indicators can provide a useful overview, but cannot shed light on the drivers of change in utilisation levels

indicators provide a potentially useful overview of the grand trends in service usage; by incorporating all information on extent and intensity of use, these indicators are able to provide a useful overview of how overall usage levels have evolved. However, the indicators are unsuited to detailed analysis of the drivers of changes in utilisation levels. With volume measures, a 10% increase in intensity of use is observationally equivalent to a 10% increase in the extent of use. Thus, evaluators will not be able to say which indicator has led to an observed change. Moreover, if changes in intensity and extent of use are in different directions, even large changes may have negligible effects on a volume indicator. A 90% increase in extent of usage, coupled with a 91% decrease in intensity of usage will lead to only a 1% change in a volume measure. In general, evaluators will need to examine all three classes of service usage indicators to build an integrated picture of how microinsurance has altered the usage pattern.

Needs-based usage

Rather than examining usage over the population as a whole, needs-based

usage indicators are examined only for those people reporting that they had a need to use the service in question. Either the extent or the volume of needs-based usage can be measured in this manner, using the formulae:

Extensiveness:

$$\frac{\text{Number of people using service at least once}}{\text{Number of people reporting need to use service}}$$

Volume:

$$\frac{\text{Total number of service usages}}{\text{Number of people reporting need to use service}}$$

By allowing the number below the line to vary over time, this class of indicator removes that part of variation in extensiveness of use, which is due to factors other than insurance. This is of particular importance for services that experience systematic variations in demand. Consider an HMI evaluation in South Asia that takes a baseline during dry season and an end line during monsoon. Monsoon season is associated with a seasonal uptick in vector borne disease, such as malaria, dengue, etc. With higher disease levels, imagine that the recorded number of people making an OP visit doubles. Standard extensiveness and volume indicators

will also double, and evaluators are at risk of erroneously assigning this effect to the insurance intervention. However, a needs-based usage indicator removes this error, by allowing the demand level to vary over evaluations. That is, whilst the recorded number of people recording OP visits doubles, the recorded number of people stating they had a need for OP treatment will also at least double. Thus, the value of the needs-based indicator will remain unaffected by the change in background conditions between baseline and end line, and erroneous conclusions will not be made.

This class of indicators is commonly used in both extensiveness measures (Rao et al. 2009; Gnawali et al. 2009), and volume measures (Gnawali et al. 2009; Schneider and Hanson 2006). It can be extended to study non-health insurance schemes, as discussed, for other classes of usage indicator. However, it involves one complication at the data collection stage: respondents are required to report that there was a need to avail of a service, but that this step was not taken. This can involve a difficult judgment call, which may be inconsistently made over different respondents.

Financial protection

The core purpose of insurance schemes is to reduce the expenditure a member

must incur to remedy some expensive adverse event. This is otherwise termed *financial protection* (FP). FP is the most intensively studied impact of microinsurance schemes after service usage patterns. The indicators used to study this area are readily extended to most forms of insurance: one might examine expenditures on funerals in the case of life/funeral insurance, expenditures on restocking a farm after some problem in the case of crop or livestock insurance, or expenditure on replacing broken or stolen items in case of home insurance. The examples presented are drawn from the health insurance context common on micro-insurance evaluations.

Three forms of FP indicators are commonly used. The most basic and commonly applied are indicators based on *out-of-pocket spending* (OOPS). In health literature, OOPS is defined as, “direct outlay of households including gratuities and in-kind payments made to health practitioners, suppliers of pharmaceuticals, therapeutic appliances, and other goods and services whose primary intent is to contribute to the restoration or enhancement of the health status of individuals” (Than Sein and Waheed 2003, 10-11). That is, OOPS is the total outlay that must be financed by a household’s own resources to cope with the impact of the risk that is insured against net of any expenditure financed by insurance.

A reduction in the average level of OOPS is often considered an important target of microinsurance schemes: those who are insured should find that when an adverse event occurs, they need to spend less money from their personal finances to cope with it.

Indicators based on OOPS are easily applied to most insurance schemes, but are far less malleable than the service-usage indicators described above. By definition, OOPS can only be calculated over those people who experience the adverse event they insured themselves against. Other than in very large evaluations, this group will usually not include enough people to perform subgroup analysis. OOPS should also be calculated as a total: it often makes little sense to look at subexpenditures within OOPS. Only the time frame of measurement can be easily varied. Common indicators based on OOPS include:

- *OOPS per visit to a health-care provider.* Average OOPS per visit to a health-care provider is the most commonly used FP indicator. It has been applied to outpatient care visits (Gumber and Kulkarni 2000), inpatient care visits (Jütting 2004), and deliveries (Smith and Sulzbach 2008).
- *OOPS per illness/maternity episode.* A less-used indicator is total OOPS over the course of a particular

episode of illness (see Gumber 2001). OOPS is calculated as the sum of all expenditures over all visits to all health-care providers to treat a particular ailment. The indicator is sometimes defined as covering both OP and IP care (Diop, Sulzbach, and Chankova 2006). Whilst this measure gives a more rounded view of the level of financial protection insurance offers per event, it also requires survey respondents to report more information, relating to a longer time period, increasing the scope for inaccuracies.

- *OOPS per episode, within a time frame.* One potential midway point between OOPS per visit and OOPS per episode is OOPS per episode over a limited time frame. Respondents may be asked to report all OOPS on an illness within the last one month (Rao et al. 2009) or over a longer period (Wagstaff et al. 2009). This approach balances the problem of recall issues in the *per episode* indicator against the limited insight of the *per visit* indicator.

One issue with FP measures based purely on OOPS is that they are absolute rather than relative. As an example, imagine that two people have had an outpatient visit for exactly the same ailment—a twisted ankle—and both have been prescribed the same treatment for this ailment: painkillers and an x-ray. One person has taken the

full course of treatment; the other has skipped the x-ray. It may be that the person who skipped the x-ray has done so because he/she is much poorer than the person who did not. Although he/she has spent less money, this money may have been more critical to basic purchases like food. As a consequence, the person with lower OOPS may now be both financially and physically worse off than the person with higher OOPS.

The second class of FP measures, catastrophic expenditures (CE) indicators, offers a partial solution to this problem. Expenditures on a given adverse event are considered catastrophic when the OOP payments incurred to deal with an adverse circumstance are of a level “at which a household is forced to sacrifice other basic needs, deplete productive assets, incur debt, or be impoverished” (O’Donnell et al. 2005). In the microinsurance literature, this is generally proxied as 10% of household income. Evaluators proceed to calculate the proportion of adverse events on which catastrophic expenditures are incurred, i.e., the proportion of insurance claims on which total OOPS amounted to more than 10% of HH income (Ranson 2002 and Devadasan et al. 2007). CE indicators thus provide some measure of the relative impact of expenditures on the overall welfare of those incurring them. However, this measure is partial only:

a household spending of 9.9% of its income on a hospitalisation is counted as non-catastrophic, whilst a household spending 10.1% is included. Moreover, all expenditure levels above the 10% threshold are considered equally catastrophic: no differentiation is made between a household spending 10% of its income on replanting failed crops and one spending, say, 60% of its income. A simple indicator that would provide a fully relative measure of the extent of the impact of OOPS on HH welfare is *OOPS per adverse event / HH income*. The only applied use of this indicator is in Ranson (2002), who examines costs per hospitalisation as a percentage of HH income.

A third class of FP indicators is total expenditures (TE) on all events that are insured against, including insurance premiums. That is, *OOPS for all insured events + insurance premium*. Assuming no change in background conditions, average TE would not be expected to decrease: a decrease in TE would indicate that the insurance scheme is paying out more than has been paid into it. However, the variance of TE may be expected to decrease, as insurance decreases the incidence of extreme high-cost events amongst the insured. This potentially useful indicator has not been used in any evaluation to date. TE type indicators are generally calculated on a per annum basis, i.e., the timeframe over which an insurance

policy is usually active. TE indicators are sometimes calculated in absolute terms (Gumber and Kulkarni 2000) and sometimes as a percentage of household income (Franco et al. 2008). They remain rarely used.

All forms of FP indicators discussed contain two inherent limitations: complexity of reporting and limited scope. Regarding complexity of reporting, all FP indicators require, at a minimum, that full data on OOPS be collected. Provider records often do not suffice, as OOPS covers all expenditures made by an individual over multiple providers. Individual insurance members/households must usually be asked via a household survey. Respondents are typically asked detailed questions on expenditures on a variety of items over an extended period. For CE and TE measures, they may additionally be asked for details of their household income, a further complex task requiring detailed questioning. The complex and difficult-to-remember data required to evaluate FP indicators makes them potentially more prone to reporting errors than most other classes of indicators.

Evaluators should also be aware of the limits of the interpretations that can be drawn from FP indicators, and specifically, that they examine monetary outcomes rather than welfare outcomes. Returning to the example outlined above, the man who has skipped the x-ray would be found to be more financially secure on the basis of OOPS per visit. A measure based on OOPS would, therefore, rank him as having had a better result, even though he has skipped recommended treatment. Evaluators would need to combine OOPS data with separate health-seeking behaviour data to determine this fact. In order to provide valid conclusions on the impact of an insurance scheme, measures based on OOPS need to be carefully evaluated in conjunction with other information on the income, status, and treatment needs of the individuals involved.

Household socioeconomic status

Theoretically, one of the key impacts offered by insurance is its potential to increase the socioeconomic status of the insured household. By mitigating the financial impact of adverse events,

Full data on OOPS is difficult to gain ---> financial protection indicators prone to reporting errors

insurance may directly allow households to avoid taking on loans, selling assets, or cutting back on investments when such events occur. This translates into potentially higher income levels for the insured. Moreover, by mitigating the impact of adverse events, insurance may liberate households to undertake riskier but more profitable, production patterns. For example, farmers with crop insurance may be more likely to move into monoculture cultivation of profitable cash crops than uninsured farmers. For these reasons, the impact of insurance on measures of household socioeconomic status and income levels are an important topic in microinsurance impact. Whilst this is a less studied area than service utilisation or financial protection, a number of indicators have been used:

- *Change in household income level, last one year.* Two studies examine the impact of health microinsurance on the income levels of participants (Aggarwal 2010 and Hamid, Roberts, and Mosley 2010). Both studies used detailed questioning of households to establish estimates of their gross income level, i.e., total income, ignoring money spent on inputs into the household's farm and/or microenterprises. Aggarwal takes one survey prior to insurance being introduced and one survey after, and thus examines the effect of insurance on *change* in

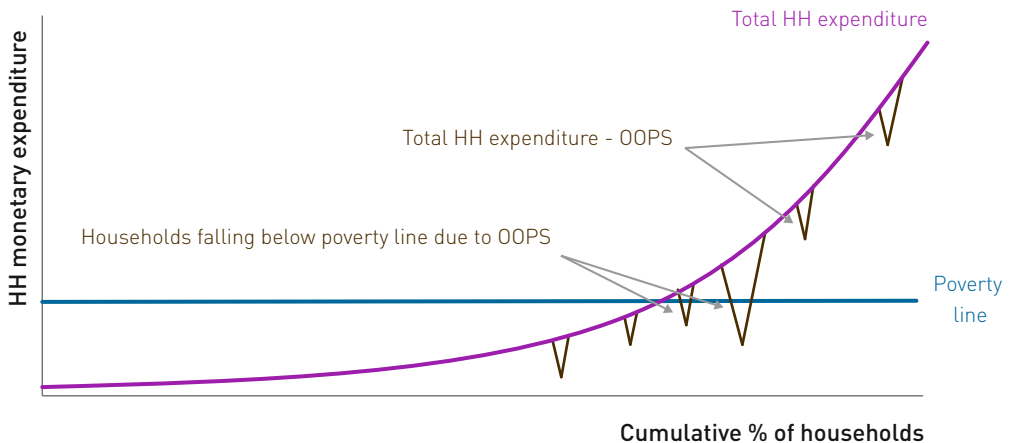
household income. Hamid, Roberts, and Mosley have only one observation, post-insurance, and thus look at the effect of microinsurance on the *level* of household income. One innovation in Hamid, Roberts, and Mosley is the use of an age-based weighting system to examine the impact of insurance on individual income levels, as opposed to household income levels.

- *Household asset levels.* Hamid, Roberts, and Mosley (2010) have further examined two measures of household asset levels: the present value of all non-land household assets and the present value of all non-land productive household assets. The rationale for examining non-land assets as an indicator of household socioeconomic status is that these represent the stock of assets available to finance expenditures in case of adverse events. The productive assets were those separately identified by the household as being used in one of the household's microenterprises, and thus represent the stock of income generating assets. Land was excluded due to the different prices across the various study sites, perhaps indicating underlying non-compatibility. No time period is applied: assets that are held by the household at time of interview are recorded.
- *Change in percentage of households falling below poverty line, last*

two weeks. Schneider and Hanson (2006) pioneer the use of this indicator in a microinsurance context. Detailed information is separately collected on household non-medical expenditure levels and household OOPS on health care, in the past two weeks. Total expenditure is calculated as the sum of these two components and taken as a proxy for HH income. All households are then ranked by total expenditure, from lowest to highest, and ordered along a horizontal axis

(see figure 3—this form of graph is sometimes called a Penn's Parade). Both their total expenditures and total non-medical expenditures are graphed against a vertical axis. A poverty line is drawn. The evaluators then measure the proportion of HHs above the line in terms of total expenditure, but below the line once medical OOPS are subtracted (in figure 3, it can be seen that two such cases exist). This percentage is compared across insured and non-insured groups.

Figure 3: Demonstration of Schneider and Hanson's (2006) percentage of households falling below poverty line



These approaches, in particular the second and third, follow techniques widely applied outside the microinsurance literature. Unfortunately, all existing indicators of HH socioeconomic status rely on the collection of complex data from respondents who

may suffer from recall issues. Consider how difficult it would be to accurately recall all of your expenditures in the last month, and you will have some idea of the potential that exists for errors to be introduced. Issues of deliberate misreporting affect these

types of indicators. For example, respondents may deliberately misstate income or health spending data in the hope of gaining some benefit from the party administering the survey, out of a lack of trust in the surveyor, or in fear that their answers will be made public within their community. Unfortunately, these indicators, whilst following current best practices, remain highly complex to calculate and prone to reporting error.

Equity of impact

The final class of indicators used to evaluate the impact of microinsurance is one that examines the equity of service usage, expenditure levels, or other variables across insured individuals of differing socioeconomic status. **Strictly, equity measures can be thought of not as a separate class of indicators, but as a different way of examining impact using the indicators already described.** The indicators most commonly evaluated for equity across insured people are the extent and rate of OP/IP visits. Evaluators have also, on occasion, examined how equitable

the distribution of deliveries is in modern facilities, probability of self-medication when ill, cost of consultations, and cost of drugs per illness episode. Three separate approaches have been previously applied. Each is technically rigorous.

- *Change in probability of care across income quartiles.* Yip, Wang, and Hsiao (2009) apply this technique. They begin by measuring the levels of a variety of health-care usage extensiveness indicators, including incidence of OP visits, incidence of visits to different OP providers, and incidence of self-medication. They measure before and after insurance is implemented. Measurements are taken across two groups: treatment and control. Next, they subtract the level of each variable pre-insurance from its level post-insurance, giving a net change in the probability of utilising each health-care service. They then match each individual in the treatment group with similar individuals in the control group. The groups are then ranked in ascending order of income and divided into

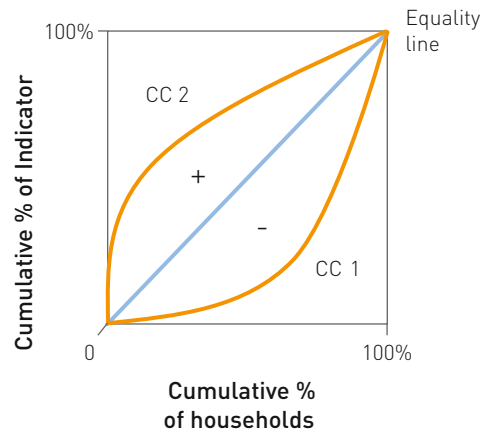
Equity measures are not a separate class of indicators, but a different way of examining existing indicators to make comparisons amongst subgroups

four cohorts. Finally, the difference in net change in the probability of utilising each service across treatment and control individuals is measured and averaged. Differences in the changes of utilisation patterns across income cohorts are then examined. This is a technically demanding approach, involving both double-difference estimation and propensity score matching.

- *Concentration-curve based indicators.* This approach is followed by Dror et al. (2005), who also examine equity across a number of health-care usage extensiveness variables (incidence of hospitalisation, incidence of OP consultation, and attendance of professional staff at deliveries). A household survey covering these indicators and separate estimations of household income is conducted. Households are then ranked according to their reported income (on an X-axis), and their cumulative percentage usage of the care variable in question plotted along a vertical axis (see figure 4). This plot is the concentration curve (CC).

If usage of care is entirely equitable (i.e., 10% of care is being used by the poorest 10% of the population, and so on), then the concentration curve lies exactly along the 45 degree line—or equality line—between these axes. If poorer people use proportionately

Figure 4: Concentration curves and needs-based care curves



less care, the concentration curve lies below the equality line (e.g., 5% of care is used by the poorest 10%, or CC1). If the opposite is the case, the curve lies above the equality line (e.g., 15% of care is used by the poorest 10%, or CC2). The concentration index is calculated as twice the area between the equality line and the concentration curve. It ranges from +1 to -1, with negative values indicating a distribution of usage favouring the richer, and positive values indicating the opposite. This approach is less technically demanding than that used in Yip, Wang, and Hsiao (2009). It can also be used to rank equity in expenditure levels, as in Dror et al. (2009), or many other variables.

- *Needs-based indicators.* Finally, Wagstaff and Van Doorslaer (1998) suggest employing an approach in which

the actual concentration curve for care is measured not against the equality line, but against a “needs adjusted” concentration curve. This is a statistically generated curve, which displays the amount of care each person in the sample would have received, had they received the same amount of care which others with the same needs received on average. The difference between this curve and the concentration curve then provides a measure of inequity in treatment levels. This approach is of specific application to health insurance evaluations. An example of its use in an HMI evaluation is found in Schneider and Hanson (2006).

9.5. Conclusions

The problem of choosing indicators has been presented here as one of trying to determine in advance which few pieces of information are most profitable to collect from a much larger set of potential choices. The qualities of different types of indicators and the ways in which different types of indicators can be systematically combined to help increase the usefulness of the information gathered have been outlined to help guide the evaluator’s choices. From a practical point of view, indicators most common in microinsurance impact evaluations together with their advantages and drawbacks have been discussed.

However, fully evaluating the meaning of indicators requires the evaluator apply them in combination with other types of information. Rather than looking at a single indicator and trying to draw conclusions from this, it is more useful to assess a combination of different indicators and to take them into a broader context, applying other types of information.

References

- Aggarwal, A. 2010.** Impact evaluation of India’s “Yeshasvini” community-based health insurance program. *Health Economics* 19:5-35.
- Chambers, R. 2009.** So that the poor count more: Using participatory methods for impact evaluation. *Journal of Development Effectiveness* 1(3):243-246.
- Criel, B., P. Van der Stuyft, and W. Van Lerberghe. 1999.** The Bwamanda hospital insurance scheme: Effective for whom? A study of its impact on hospital utilization patterns. *Social Science and Medicine* 48(7):897-911.
- Devadasan, N., B. Criel, W. Van Damme, K. Ranson, and P. Van der Stuyft. 2007.** Indian community health insurance schemes provide partial protection against catastrophic health expenditure. *BMC Health Services Research* 7(43):1-11.
- Diop, F., P. Schneider, and D. Butera. 2000.** Summary of results: Prepayment schemes in the Rwandan districts of Kabgayi, Byumba and Kabutare. *Technical Report no. 59*. Bethesda, Maryland: Partnerships for Health Reform.
- Diop, F., S. Sulzbach, and S. Chankova. 2006.** The Impact of mutual health organizations on social inclusion, access to health care and household income protection: Evidence from Ghana, Senegal and Mali. Bethesda, MD: The Partners for Health Reformplus Project, Abt Associates Inc.

- Dror, D., R. Koren, and D. M. Steinberg. 2005.** The impact of Filipino micro health insurance units on income-related equality of access to healthcare. *Health Policy* 77(3):304-317.
- Dror, D., R. Radermacher, S. B. Khadikar, P. Schout, F. X. Hay, A. Sing, and R. Koren. 2009:** Microinsurance innovations in low-cost health insurance. *Health Affairs* 28:1788-1798.
- Franco, L. M., F. Diop, C. R. Burgert, A. G. Kelley, M. Makinen, and C. H. T. Simpara. 2008.** Effects of mutual health organisations on use of priority health-care services in urban and rural Mali: A case-control study. *Bulletin of the World Health Organization* 86(11):830-838.
- Gnawali, D. P., S. Pokhrel, A. Sié, M. Sanon, M. De Allegri, A. Soares, H. Dong, and R. Sauerborn. 2009.** The effect of community-based health insurance on the utilization of modern health care services: evidence from Burkina Faso. *Health Policy* 90(2-3):214-22.
- Gumber, A. 2001.** Hedging the health of the poor: The case for community financing in India. Health, Nutrition and Population Discussion Paper. Washington, D.C.: World Bank.
- Gumber, A. and V. Kulkarni. 2000.** Health insurance for informal sector: Case study of Gujarat. *Economic and Political Weekly* 35(40):3607-3613.
- Hamid, S. A., J. Roberts, and P. Mosley. 2010.** Can micro health insurance reduce poverty? Evidence from Bangladesh. Sheffield Economic Research Paper Series 2010001. Sheffield, UK: University of Sheffield.
- Howe, K. R. 1988.** Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher* 10:10-16.
- Johnson, R. and A. Onwuegbuzie. 2004.** Mixed methods research: A research paradigm whose time has come. *Educational Researcher* 33(7): 14-26.
- Jones, M. and R. Sugden. 2001.** Positive confirmation bias in the acquisition of information. *Theory and Decision* 50(1):59-99.
- Jütting, J. P. 2004.** Do community-based health insurance schemes improve poor people's access to health care? Evidence from rural Senegal. *World Development* 32(2):273-288.
- O'Donnell, O., E. van Doorslaer, R. P. Rannan-Eliya, A. Somanathan, C. C. Garg, P. Hanvoravongcha, M. N. Huq, A. Karan, G. M. Leung, K. Tin, and C. Vasavid. 2005.** Explaining the incidence of catastrophic expenditures on health care: Comparative evidence from Asia. Working Paper 5. EQUITAP Project.
- Polonsky, J., D. Balabanova, B. McPake, T. Poletti, S. Vyas, O. Ghazaryan, and M. K. Yanni. 2009.** Equity in community health insurance schemes: Evidence and lessons from Armenia. *Health Policy and Planning* 24(3):209-216.
- Pronyk, P. M., T. Harpham, J. Busza, G. Phetla, L. A. Morison, J. R. Hargreaves, J. C. Kim, C. H. Watts, and J. D. Porter. 2008.** Can social capital be intentionally generated? A randomised trial from rural South Africa. *Social Science & Medicine* 67:1559-1570.
- Ranson, M. K. 2001.** The impact of SEWA's medical insurance fund on hospital utilization and expenditure: A household survey. Health, Nutrition and Population Discussion Paper no. 28895. Washington, D.C.: World Bank.
- Ranson, M. K. 2002.** Reduction of catastrophic health care expenditures by a community-based health insurance scheme in Gujarat, India: Current experiences and challenges. *Bulletin of the World Health Organization* 80(8):613-621.
- Rao, K. R., H. Waters, L. Steinhard, S. Alam, P. Hansen, and A. J. Naeem. 2009.** An experiment with community health funds in Afghanistan. *Health Policy and Planning* 24(4):301-311.
- Rogers, P. 2009.** Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation. *Journal of Development Effectiveness* 1(3):217-226.
- Schneider, P. and K. Hanson. 2006.** Horizontal equity in utilisation of care and fairness of health financing: A comparison of micro-health insurance and user fees in Rwanda. *Health Economics* 15(1):19-31.

Smith, K. V. and S. Sulzbach. 2008. Community-based health insurance and access to maternal health services: Evidence from three West African countries. *Social Science and Medicine* 66(12):2460-2473.

Than Sein, U. and A. Waheed. 2003. National health accounts: Policy brief on concepts and approaches. *Regional Health Forum WHO South-East Asia Region* 7(2):1-19.

Thornton, R., L. E. Hatt, E. M. Field, M. Islam, F. Solis Diaz, and M. A. Gonzalez. 2010. Social security health insurance for the informal sector in Nicaragua: A randomized evaluation. *Health Economics* 19 [Suppl. 1]:181-206.

Trujillo, A., J. Portillo, and J. Vernon. 2005. The impact of subsidized health insurance for the poor: Evaluating the Colombian experience using propensity score matching. *International Journal of Health Care Finance and Economics* 5(3):211-239.

Tschirley, D. and D. Rose. 2000. Developing cost effective methods for estimating household income and nutrient intake adequacy. USAID Policy Synthesis 54. Washington, D.C: USAID.

Wagstaff, A., M. Lindelow, G. Jun, X. Ling, and Q. Juncheng. 2009. Extending health insurance to the rural population: An impact evaluation of China's new cooperative medical scheme. *Journal of Health Economics* 28(1):1-19.

Wagstaff, A. and E. Van Doorslaer. 1998. Equity in Healthcare Finance and Delivery. *Handbook of Health Economics* 1(2):1803-1862.

Weiss, C. 1997a. *Evaluation: Methods for studying programs and policies.* 2nd ed. Upper Saddle River, New Jersey: Prentice Hall.

———. **1997b.** How can theory-based evaluation make greater headway? *Evaluation Review* 21(4):501-524.

White, H. 2009. Theory-based impact evaluation: Principles and practice. *Journal of Development Effectiveness* 1(3):271-284.

Yip, W., H. Wang, and W. Hsiao. 2009. The impact of rural mutual health care on access to care: Evaluation of a social experiment in rural China. *Health Economics* 18(2):S65-S82.

Core outcomes, impacts, and indicators for microinsurance

Katja Roth, Ralf Radermacher, and Pradeep Panda



10.1. Introduction

A lot of attention is focused on how to best measure impacts and outcomes in microinsurance (or other interventions). However, as important as deciding on *how* to measure impacts and outcomes, it is to decide on *what* exactly to measure. This chapter discusses the most useful indicators for an evaluation of microinsurance schemes, as identified through a Delphi study amongst experts.

The objective of this chapter is to enable practitioners, researchers, and others to correctly choose and define purposeful indicators.¹ Quantitative and hybrid indicators, specifically, are taken into account, whilst purposefully neglecting qualitative indicators since they differ substantially from one another. The chapter also touches upon creating research tools based on the indicators discussed in the first part of the chapter.

10.2. Identifying core impacts, outcomes, and indicators

In an attempt to define core impacts and outcomes attributed to microinsurance, almost 30 scholars and practitioners

from the microinsurance arena participated in a Delphi study in 2012. In the study *impacts* are defined as *long-term effects* of microinsurance. This follows the definitions used by the Development Assistance Committee (DAC) Working Party on Aid Evaluation from the Organisation for Economic Co-operation and Development (OECD) (2002).

In contrast, *outcomes* are defined as *short-term and medium-term effects*. These outcomes are frequently causally related to the long-term impacts, since, in many cases, they lead to the long-term impacts and can be considered as the step before them in the causal—and chronological—chain. **Impacts and outcomes can be positive or negative, i.e., desired or undesired.**

The Delphi study used a three-step approach. In the first step, the participating experts submitted suggestions for important outcomes and impacts. They then rated their suggestions

Impacts and outcomes can be positive or negative, i.e., desired or undesired

¹ This chapter is linked to appendix 1 “Guidelines for Indicators”, of this publication. The appendix provides detailed guidelines on how to measure each of the indicators described here. Appendix 1 is not included in the printed version of this book, but it is available online at www.microinsurancenet.org.

according to importance during the second round and further discussed and refined them in the third round.

In the ranking a five point scale was used. Impacts and outcomes which on average got at least the second highest ranking or for which at least half of the respondents attributed the highest ranking were considered “core” impact or outcome.

Using this ranking, a set of 12 core impacts and 17 core outcomes were identified. Out of the 17 core outcomes, eight apply to all types of

microinsurance, while nine specifically target health microinsurance. Core outcomes, particularly appropriate for other microinsurance types (for example, agricultural, life or funeral microinsurance) could not be identified. The 12 core impacts apply to all types of microinsurance and can be grouped into six thematic areas:

1. Financial protection
2. Living standards
3. Health
4. Education
5. Perceptions / psychological issues
6. Social life and community

Figure 1: Overview of core outcomes and impacts

Core outcomes		lead to	Core impacts
General	• Risk taking behaviour • Risk management strategies (ex-ante) • Risk management strategies (ex-post) / coping strategies in case of shock – sale of assets for managing expenses related to shock event – reliance on informal risk sharing networks • Variability of costs or profits with respect to the coverage of the insurance scheme • Total costs in case of shock • Total out-of-pocket spending (OOPS) in case of shock	→	Dimension 1: financial protection • Risk of poverty / financial protection / financial vulnerability Dimension 2: living standards • Economic situation of the household – level of assets Dimension 3: health • Nutrition – quality of nutritional intake – quantity of nutritional intake / extent of hunger • Physical health
	• Health-care services – quality of health-care providers – quantity of health-care providers • Receiving (appropriate) health care – Health-care utilisation (needs based) – delay in health-care seeking • Equity regarding health and health care – equity in health care regarding different socioeconomic groups – equity (needs-adjusted) in health care regarding different household members	→	Dimension 4: education • Educational attainment of children – child labour – duration of school attendance of children Dimension 5: perceptions / psychological issues • Peace of mind / perception of financial vulnerability Dimension 6: social life and community • Social capital

Also identified were potential indicators for all core outcomes and impacts. The measurement of the indicators depends on the particular setting and the particular situation of the intervention and the microinsurance programme. For this reason, two or more indicators are suggested for most of the effects.

10.3. Core indicators for microinsurance outcomes

As mentioned above, 17 core outcomes have been the result of the Delphi process, eight of which are general in nature whilst nine are specifically for health microinsurance (see figure 1). For other types of microinsurance, there are no specific core outcomes suggested.

10.3.1. General core outcomes and corresponding indicators

For the general core outcomes, issues of risk taking and risk management are dominant (see table 1). Changes in risk taking behaviour has been rated as one of the core outcomes microinsurance may have in general. The rationale behind this is that more income security (or smoothing of expenditures) through insurance may increase the willingness of the insured to make investments that are more risky, but potentially have a higher return, and hence can constitute a more efficient

allocation of resources. This does, in particular, apply to productive investments and corresponding loans. Due to the effects described, the insured may be willing to invest more and, furthermore, take more loans for investments. For this reason *productive investments as percentage of total income of the household* and *total amount of loans taken* are being suggested as indicators for risk taking behaviour (indicating a rise in risk taking behaviour by a higher number of loans). Following the same argument, the *total amount of savings* is suggested as an indicator, indicating a rise in risk taking behaviour by a lower amount of savings, since no money has to be put aside for financial consequences of shock events and hence savings are freed up for investments. However, another argument is that households are able to save more money and need fewer loans due to fewer financial shocks and higher income caused by higher productive investment. **Additional information is needed as both the amount of loans and the amount of savings are indicators, which, if they stand alone, are ambiguous and difficult to interpret. Context information is hence needed.**

Additional qualitative impact assessments studies are helpful for examining these effects in more detail. For example, qualitative interviews with household members examining the reasons

Some indicators can be ambiguous and understanding them requires more context information

they save money or for what they use loans can be useful for this. It also has to be noted that it depends on the type of insurance and on the particular risk that is taken whether increased risk taking behaviour is desirable or not with respect to an increase of welfare of household members. In the case of crop or weather insurance, increased risk taking behaviour (and decreased risk diversification) may lead to higher crop yields, a desirable outcome. Yet more risk taking in the area of health, may lead to an adverse health shock that is not desirable.²

Changes in ex-ante risk management strategies have been ranked as core outcome. Similar to risk taking behaviour, here the rationale is also that lesser alternative risk management strategies are necessary due to microinsurance. This refers to savings, as well as liquid assets, the number of income sources per household, and the number of saving networks per household (like Rotating Savings and Credit Associations (ROSCAs), Accumulated

Savings and Credit Associations (ASCAs), etc.). But the opposite may also be possible: a higher household income (or less expenditure) leads to more activities in this field.

Ex-post risk management strategies have been scored as core outcomes for microinsurance, together with the two subordinated outcomes *sale of assets for managing expenses related to shock event* and *reliance on informal risk sharing networks*.³ In this case, indicators are the amount of loan taken and the amount of savings used. It is important to note that in this case, loan amounts and savings *taken or used in case of shock* are suggested. This is different from risk taking behaviour and ex-ante risk management strategies where the *total amount* of savings or loans of the household is the indicator. For ex-post risk management strategies, the indicators refer to financial behaviour closely related to the shock event itself. The *difference in total household expenditures before and after shock events (not including paying back loans*

² These explanations are based on input given by Xavier Giné for the Delphi study.

³ Additional subordinated outcomes exist, but have not been scored as core and are not included.

taken for the shock event and paying back interest for those) can be used as an indicator, as well as *changes in food intake*, since, without microinsurance, households might be more required to reduce other expenditures than is the case with microinsurance. Moreover, without microinsurance, households might be required to take children out of school (in order to make them contribute to the household's income or to save money on school-related expenditures, like school uniforms, books, transport, or tuitions). This action can serve as indicator as well.

Regarding the subordinated outcome of sale of assets, both the *total value of sold assets in case of shock* as well as the *percentage of those assets which have been replaced or recovered six months after the shock event* can be useful as indicators. Whilst the former one shows the general need of the household to sell assets in case of shock events, the latter indicator takes into account that the payout of insurance claims may happen with time lags. Insured households may be forced—despite microinsurance—to first sell their assets in case of a shock event, but can later recover or replace them due to claim payouts. A second subordinated outcome of ex-post risk management strategies, *reliance on informal risk sharing networks*, has been ranked as core. This refers to financial contributions (with or without interest,

as loan or as gift) to be received from different sources like family, neighbours, and institutions (i.e., churches).

Three other core outcomes are: 1) the variability of costs or profits, 2) the total costs in case of shock, and, 3) the total out-of-pocket spending (OOPS) in case of shock. In particular, smoothing the variability of costs or profits is one of the central expected outcomes of microinsurance. For the particular case of health insurance, indicators for total out-of-pocket spending can refer to varying categories of treatment, like hospital stay, deliveries, self-treatment, ambulatory care from formal providers, or inpatient care.

10.3.2. Core outcomes and corresponding indicators specific for health microinsurance

The nine core outcomes identified for health microinsurance refer to health-care services in general, the receiving of health care, equity regarding health, and health-care utilisation (table 2). The first core outcome of health-care services includes two subordinated outcomes, which are quality and quantity of health-care providers. The rationale behind this is that due to health microinsurance, the demand for health-care services rises and may lead to an improvement in both quality and in quantity supply of health-care services. For quality of health-care

Table 1: General outcomes of microinsurance

General outcomes of microinsurance	Indicators
• Risk taking behaviour	• Productive investment as percentage of total income of the household • Total amount of loans taken • Total amount of savings
• Risk management strategies (ex-ante)	• Total amount of savings • Amount of liquid assets • Number of income sources per household • Number of memberships in ROSCAs, ASCAs, and other informal savings networks per household
• Risk management strategies (ex-post) / coping strategies in case of shock	• Amount of (formal) loans taken in case of shock • Amount of savings used in case of shock • Difference of total household expenditures before and after shock events (without paying back of loans and interest) • Food intake (quality and quantity, self-reported) • Percentage of children taken out of school due to shock event
- Sale of assets for managing expenses related to shock event	• Total value of assets sold in case of shock • Percentage of assets recovered/replaced after being sold in case of shock (i.e., some land/livestock may be repurchased after insurance payout, houses may be repaired) six months after shock event
- Reliance on informal risk sharing networks	• Total amount of borrowing for shock-related expenditures from informal networks • Total amount of borrowing for other expenditures from informal networks • Total amount of contributions received from family in case of a shock (as loan) • Total amount of contributions received from family in case of a shock (as gift) • Total amount of contributions received from informal or semiformal organisations, (e.g., local church association, employer) in case of a shock (as loan) • Total amount of loans (currently pending) given to family members and other community members
• Variability of costs or profits	• Variability of net costs or net profits (including insurance premiums and claim payouts)
• Total costs in case of shock	• Total costs in case of shock
• Total out-of-pocket spending (OOPS) in case of shock	• Net OOPS per shock event In case of health: • Net OOPS per full episode of illness • Net OOPS on varying categories of treatment: hospital stay, deliveries, self-treatment, ambulatory care from formal providers, inpatient care

services, *hospital mortality rates* and *scoring based on quality assessments* can be used as indicators. For quantity of health-care providers, the *number of modern health-care providers within*

a defined area/radius is suggested as an indicator.

For the outcome of receiving (appropriate) health care, two subordinated

outcomes are also ranked as core and refer to health-care utilisation and the delay in health-care seeking. For these outcomes, the total number of treatments (regarding various types of health-care services), as well as the number of illness episodes involuntarily self-treated, can be used as indicators. Furthermore, to assess health-care utilisation of children, utilisation rates of children regarding different treatments, or prevention measures can be used as indicators (for example, for immunisations). For delay in treatment, *the number of days the symptoms persisted before treatment was sought* can be a valuable indicator.

Of key concern is the third main outcome that has been identified as core for health microinsurance: *equity regarding health and health care*. This does not refer to total numbers or results over all subgroups, like the other core outcomes and impacts, but refers to *differences* between various subgroups.

Particularly, equity between different socioeconomic subgroups (such as extremely poor, poor and non-poor households, or different groups of professions) and between different household members (e.g., the household head/main bread winner, daughter-in-law, parents-in-law, boys, girls) are considered as core. This can basically be assessed by any of the indicators

Especially in health, subgroup analysis is important for understanding aspects related to equity

mentioned above for the outcome of receiving (appropriate) health care, split by the specific subgroups of interest and comparison of potential differences (and the differences amongst subgroups in the control group). In addition, indicators measuring the total costs of health-care expenditures (per illness episode or per time period, e.g., 6 months) can be used to assess equity regarding health care. Equity is important to assess since all other indicators give average numbers for all household members or socioeconomic groups. But it is important to examine whether, for instance, boys get on average six immunisations in their first year of life, whilst girls get only two, (i.e., leaving girls behind in terms of health), or whether both groups get a similar number of immunisations.

Table 2: Outcomes specific for health microinsurance

Outcomes specific for health microinsurance	Indicators
• Health-care services	
- Quality of health-care providers	• Hospital mortality rate • Scoring on quality assessments
- Quantity of health-care providers	• Number of modern health-care providers within a defined area/radius
• Receiving (appropriate) health care	
- Health-care utilisation (needs based)	• Total number of visits to outpatient services per household member within the last month • Total number of visits to inpatient services (with at least 24 hours of hospitalisation) per household member within the last month/year • Total number of illness episodes involuntarily self-treated per household member within the last month Regarding children: • Percentage of children (below age 5) seeking diarrhea treatment • Percentage of children (below age 5) sleeping under a mosquito net • Percentage of children (below age 5) getting vitamin A supplements • Number of immunisations for children below age 1 (per child)
- Delay in health-care seeking	• Number of days symptoms persisted before treatment was sought
• Equity regarding health and health care	• Use any of the indicator mentioned above for subgroups
- Equity in health care regarding different socioeconomic groups	• Use any of the indicator mentioned above for socioeconomic subgroups
- Equity in health care regarding different household members	• Use any of the indicator mentioned above for subgroups of household members

10.4. Core indicators for microinsurance impacts

Unlike outcomes, impacts are long-term changes, often taking place over five years or more after a programme has started. Pressure to prove positive

impact is strong from all sides. Private and public financial supporters, donor organisations, local partner organisations, and internal or political institutions are all looking for an immediate positive impact. Unfortunately, the impact needs time to unfold; also, assessing it takes time. If early results

*Impact needs time to unfold,
measuring it too early is not possible*

are needed due to external or internal forces, it might be better to measure indicators based on short-term outcomes or for social performance⁴. Using these indicators as proxies for impact is more advisable in this case than trying to measure impact and being disappointed by the final results. However, to apply outcome indicators and social performance indicators as proxies, a reliable theory of change is necessary.

Impact dimensions are likely interrelated: a positive impact on one aspect of a client's life can lead to a negative or positive impact on another

The experts of the Delphi study identified twelve impacts as core impacts. Unlike the outcomes described above, these core impacts and their indicators can be applied for all types of microinsurance. The dimensions and impacts discussed in this chapter are closely interrelated. For example, better health may lead to higher financial protection, and vice versa, and higher education may lead to better living standards, and vice versa. The impacts mentioned below, thus, cannot be understood as standalone effects, but are pieces of the same puzzle. This has implications when conducting impact assessments: if one is able to measure a positive impact on financial protection, but at the same time cannot identify positive impact on living standards, health, or education, more research might be

necessary to examine these relations.⁵

Moreover, impacts of microinsurance for different dimensions of life might have different directions: whilst the impact on health might be positive, impact on social life could be negative. The interdependencies of impacts have not yet been sufficiently understood in microinsurance, but, nevertheless, have to be taken into account when doing impact assessments.

10.4.1. Dimension 1: financial protection

Whilst most general core outcomes point to risk management and risk taking, only one impact related to financial protection was ranked as core, which is

4 Appui au Développement Autonome (ADA), Belgische Raiffeisenstichting (BRS), and the Microinsurance Network (2011) have provided a set of social performance indicators for microinsurance that can be used.

5 An example for this interrelation is described in Hintz (2010) regarding funeral microinsurance where the insurance payouts lead to more costly funerals, which, at the same time, lowers the amount of money left for the household to deal with the cut in income due to the loss of the breadwinner. The insurance payout does only partially affect the future welfare of the household.

Table 3: Core impacts and indicators on financial protection

Impacts on financial protection	Indicators
• Risk of poverty / financial protection / financial vulnerability	Percentage of households living below poverty line: • Percentage of households living on less than US\$1.25 (purchasing power parity) per person • Percentage of households living between US\$1.25 and US\$2 (purchasing power parity) per person

risk of poverty/financial protection/financial vulnerability (table 3). However, we are convinced that this impact is one of the most important that microinsurance can have, which mirrors the fact that this relationship has been subject to many studies.

The effect on financial protection is very closely related to the eight general outcomes we have discussed. They come prior to financial protection in the causal chain and are often a prerequisite for financial protection.

10.4.2. Dimension 2: living standards

Closely related to the dimension of financial protection described above is the dimension of living standards. Since microinsurance aims to target the poor, improvement of the economic situation of the household is a potential core impact of microinsurance.

There are different approaches on how to best assess the economic situation of a household. We have mentioned level of assets, level of income, and level of liabilities as potential

impacts identified by the participants of the Delphi process, but there are also expenditure related effects, like level of expenditures and change of expenditures patterns (i.e., a shift in relative allocation of expenditures regarding food, non-food, leisure, savings, and other items). However, only *level of assets* has been rated as a core potential impact. As indicators for this, the experts suggest *housing conditions, level of household assets, and consumer appliances*, as well as *savings and other working capital*.⁶ As seen in the table below, *housing conditions* and *level of household assets / consumer appliances* consist of groups of indicators rather than representing a single indicator. We suggest using scores for the single indicators that can then be summed up in a joint score. For some situations, a weighting of scores might be useful. Similarly, for the third indicator, *savings and other working capital*, we suggest adding up all sums of the different kinds of capital mentioned.

⁶ These suggestions refer mainly to input from Rebecca Thornton, who has used similar indicators for assessing a FINCA programme.

Table 4: Core impacts and indicators on living standards

Impacts on living standards	Indicators
Economic situation of the household	<i>See indicators below in subcategory</i>
o Level of assets	- Housing conditions: - Number of rooms in the dwelling - Main material of the floor of the dwelling - Main material of the exterior walls in the dwelling - Main material of the roof of the dwelling - Percentage of households with access to tapped water within 30-minute walk - Percentage of households with tapped water at their home - Percentage of households with a toilet in their home - Percentage of households with electricity in their home - Level of household assets / consumer appliances: - Number of pots and pans in dwelling - Existence of bicycle, motor bike, car (and number) - Existence of television, refrigerator, washing machine, air condition, Hi-fi) - Existence of mobile phone(s) - Savings and other working capital: amount of liquid savings (bank accounts, cash-on-hand, money in savings groups, gold, jewelry, etc.), illiquid savings (land ownership irrigated/non-irrigated), livestock, housing) and working capital (e.g., machinery, fish boats, stocks of goods))

10.4.3. Dimension 3: health

The third dimension for core impacts is health (table 5). It is important to note that the related effects are not restricted to health microinsurance only (as are the health-related outcomes we described previously). Since higher financial protection and better living standards may go hand in hand with better health, the latter should be considered as core potential effects for all types of microinsurance.

Firstly, nutrition is an issue. Both quality and quantity of nutrition (hunger) are concerns. For nutrition in general, we suggest using an indicator measuring the *total expenditures on food per person*

per week. The impact of nutrition has two subordinated impacts, which are *quality of nutritional intake* and *quantity of nutritional intake*. In case of shocks, some households might reduce their spending on diversity and quality of

All kinds of microinsurance can have an impact on health

Table 5: Core impacts and indicators on health

Impacts on health	Indicators
• Nutrition (in quantity and quality) - o Quality of nutritional intake	• Total expenditures on food per person per week • Frequency of eating vegetables or fruits (number per week) • Frequency of eating meat (if people eat meat at all) (number per week)
o Quantity of nutritional intake / extent of hunger	• Average number of meals eaten per day in last month • Number of days when food was insufficient for household in last month
• Health / morbidity	• Number of sick days in a household per person within last month (per subgroups, such as children, women, elders, etc.) • Number of days household members were unable to perform usual activities because of poor health per person within last month • Body mass index (for all) • Percentage of children with anemia • Percentage of women with anemia

food, while other households have to even reduce quantity. The picture is often not homogenous for all households included in the study and can depend on the socio-economic status of households. For quality of nutritional intake, the *frequency of eating vegetables or fruits (number per week)* and the *frequency of eating meat (if people eat meat at all) (also number per week)* are suggestions for indicators. These can be adapted to other eating or food habits of the region where the study takes place. For quantity of nutritional intake, the *average number of meals eaten per day in last month* and the *number of days when food was insufficient for household in last month* are possible indicators.

Secondly, besides nutrition, change in health itself (or morbidity) is a potential core effect of microinsurance that

builds upon the assumption that better risk management and lower variability in expenditures will lead to being able to afford better and more food (as we have described) and, furthermore, to a better utilisation of health-care services and, thus, to better health.⁷ The *number of sick days in a household per person within last month* or the *number of days household members were unable to perform usual activities because of poor health per person within last month* (which might be working or going to school) are suggested as indicators for this core effect, eventually per subgroups (such as children, women, main breadwinners, or elders). If one

⁷ Outcomes regarding health-care utilisation have been rated as core for health microinsurance only, but are potential outcomes for other types of microinsurance, too. The rationale is that by lowering the variability in expenditures, households do not require as much savings as before (for financial shocks), freeing financial resources for other purposes.

does have access to them, medical indicators like *body mass index* (per subgroups) or *percentage of children or women with anemia* would be useful indicators for this core impact as well. These data are difficult to assess on one's own, however, but are useful in cases where they have been assessed by others already (e.g., through a health camp or in hospital or school).

10.4.4. Dimension 4: education

The fourth dimension of impacts is education (table 6). Whilst education can refer to both the education of children and adults and can take place in school or somewhere else, the focus here is on school education of children.

Primary impact in the dimension of education is the educational attainment of children. Two indicators, *the percentage of children aged 6 - 16 attending school* and *days missed in school per child within last month* are recommended in order to measure this impact. These indicators need to be adapted with respect to the school age that is usual in the particular country (e.g., in some countries, school starts for children at age 5 and not age 6), and with respect to the particular time of the year the intervention is taking place. For example, if an intervention is taking place during or soon after key annual school holidays, *days missed in school* would not be a helpful indicator.

The impact of *educational attainment of children* has two subordinated impacts which are *child labour* and *duration of school attendance of children*. Child labour refers to “work that deprives children of their childhood, their potential and their dignity, and that is harmful to physical and mental development” (International Labour Organization (ILO) 2012). This includes work that is “depriving children of the opportunity to attend school..., obliging them to leave school prematurely... [or] requiring them to attempt to combine school attendance with excessively long and heavy work” (ILO 2012). Although the reduction of child labour is an objective in itself and cannot only be understood as a measure for better education of children, it is described in this chapter within the dimension of education and has been classified as a subcategory of educational attainment due to the very close relationship between education or school attainment and child labour, as shown in the quote above. (However, child labour could well have been established a dimension of its own.) Indicators like the *percentage of children below a certain age engaged in income generation, housework, farming, or other economic activity as their main occupation* (depending on the particular country or region), the *average number of working hours per week for children* (also below a certain age), *days missed in school per child within last month*,

Table 6: Core impacts and indicators on education

Impacts on education	Indicators
Educational attainment of children	- Percentage of children age 6 - 16 (or other age) attending school - Days missed in school per child within last month
o Child labour	- Percentage of children below age 15 (or other age) engaged in income generation, housework, farming, or other economic activity as their main occupation - Average number of working hours per week for children below age 15 (or other age) - Days missed in school per child within last month due to labour - Percentage of children age 6 - 16 (or other age) attending school
o Duration of school attendance of children	Highest education levels attained by children and young adults in household (below age 20)

or the *percentage of children attending school* can be applied here.

Whilst *educational attainment of children* looks more at whether children of a certain age visit school regularly, the subordinated impact of *duration of school attendance of children* looks at the overall time and the education levels that children reach. As an indicator, we suggest the measurement of *highest education levels attained by children and young adults (below age 20)* in the household. This indicator is most likely not measurable on a short-term or even mid-term scale. It is, therefore, important to take this indicator into

account only when looking at a long-term scale.

10.4.5. Dimension 5: perceptions / psychological issues

Whilst the dimensions of impact described above have a financial dimension, microinsurance might well have an impact on the perceptions and psychological state of the programme members and the corresponding household members (table 7). Due to changes in financial protection that microinsurance might cause, it can be expected that the household members change their perceptions about their financial vulnerability. In a positive

Table 7: Core impacts and indicators on perceptions / psychological issues

Impacts on perceptions / psychological impacts	Indicators
Peace of mind effect/ perception of financial vulnerability	- Changes in risk prioritisation through risk ranking exercise - Perception about confidence in the future - Perception about financial vulnerability

Table 8: Core impacts and indicators on social life and community

Impacts on social life and community	Indicators
• Social capital	• Total amount of money currently borrowed from others without interest • Total amount of money currently lent to others without interest

case, this might lead to what is often called the “peace of mind effect.” We suggest using *changes in risk prioritisation*, as well as *perceptions about confidence in the future* and *financial vulnerability*, as indicators that can all be assessed using scales.

10.4.6. Dimension 6: social life and community

Microinsurance has the potential to affect the social life of both the insured and the uninsured (table 8). For example, microinsurance might help empower insured women with regard to their position in community/society and in the household. The extent of out-migration might change, since the households face fewer financial shocks due to insurance. In contrast, particularly for the non-insured community members (but also for the insured), financial and social support after a shock event might be less due to crowding-out effects that microinsurance may have. Of all these effects, the participants of the Delphi study scored only *social capital* as a potential core impact, both for the insured and for the uninsured. Social capital is a complex construct that is difficult to measure.

In the context of microinsurance, the *amount of money borrowed from others without interest*, as well as the *amount of money lent to others without interest* are suggested as indicators and seem to be helpful to assess social capital.

10.5. Equity as impact

Very often, microinsurance schemes do not aim only at welfare improvements of poor communities in general, but to impact those subgroups that seem to be particularly poor and vulnerable, i.e., they aspire to equity between subgroups. These target subgroups might, for example, include extremely poor households, children, women, elderly, or religion- or occupation-based subgroups. For example, a specific interest of a health microinsurance programme might be to improve the access of women to health-care services and to adapt it to that of men. It could also be to adapt the educational attainment of girls to that of boys.

Furthermore, poorer and less poor households might participate differently in or benefit differently from a microinsurance scheme. Sometimes poorer households find it difficult to

afford participation in the insurance scheme; sometimes more affluent households have potentially better risk management options. Also, obtaining benefits amongst the group of insured households can differ if, for instance, better-off households find it easier to obtain benefits for one reason or the other. Understanding such effects is helpful for designing insurance schemes that deliver their benefits regardless of the economic situation of the participating household.

10.6. Transforming indicators into evaluation tools

All impacts, outcomes, and their indicators given above are considered to be core and suggested for integration into any impact assessment in microinsurance. However, measuring indicators requires a research tool designed to include appropriate questions. Moreover, obtaining the right information usually requires conducting fieldwork—although in some cases other data sources might be available—and asking the right questions in a way that can be understood and answered by the respondents. Whilst both aspects might seem straight forward initially, a number of aspects have to be considered: How many questions to include in the questionnaire? Who is the right respondent in a household and what time of the day offers the highest likelihood of meeting her—and, potentially,

what time of the year? How does one present questions in a way that interviewer and respondent have a common understanding of the underlying concept?

All questions have then to be accomplished in a questionnaire. However, designing a questionnaire to yield high quality data is not as easy as it may at first appear. When a questionnaire is administered, an intricate and subtle process starts, with the intent of providing useful and accurate information transfer from the respondent to the inquirer, as Foddy (2008) explains in detail. A series of questions need to be posed in a clear, comprehensive, and appropriate manner so that the respondent can formulate, articulate, and transmit the answers effectively. This section describes challenges and options when turning indicators into research tools and considerations for applying those tools in practice.

Usually several questions are needed to capture a single indicator

10.6.1. Formulating the right questions

Unfortunately, an indicator that is to be measured often does not match one single question in a questionnaire. In fact, for measuring one indicator, one might need several questions to be answered, whilst, at the same time, questions might overlap for different indicators. Take, for example, the *total number of visits to inpatient services (with at least 24 hours of hospitalisation) per household member within the last month*, which has been proposed above as an indicator measuring the core outcome of *health-care utilisation (needs-based)* for health microinsurance. Investigating this variable has direct implications for what a survey will need to capture: for example, the total number of household members and the visits to inpatient services within the last month prior to the survey. To explain differences amongst the households covered in a survey, the evaluator would additionally want to know the insurance status, probably some proxy for the household's wealth or income, education levels, number of illness episodes, distance from providers, and so on. This includes basically everything that is likely to influence the variable under scrutiny. The number of aspects that should be covered in the survey quickly expands with

this approach. Many aspects overlap for different indicators and have to be compressed into the minimum number of questions needed to answer the research questions of interest and to measure corresponding indicators.

10.6.2. Defining the respondents and formulating the questions right

Once the set of "right questions" is identified, phrasing these questions is the next step, interlinked with defining the respondent(s). The type and phrasing of survey questions used in a survey will play a role in producing relevant or unbiased survey responses. Another consideration to take into account and define is whom the survey should ideally address.

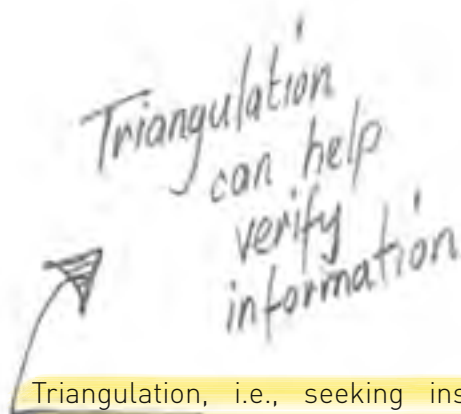
Responses to questions in a questionnaire might differ depending on the household member(s) replying to the questions. *Some members of a household have better insights into certain aspects than others.* For example, in developing country contexts, men often possess a better overview over the household's income, whilst women are said to have better knowledge about health problems in the household. This means that for questions referring to risk taking behaviour and financial protection, men might be the respondents of choice, whilst, in contrast, women might be able to give more accurate

and detailed answers on questions regarding impacts like health and education. However, interviewing different people in the household on issues they might know best is challenging in the practical implementation of a survey in the field. Not all household members might be available at particular times of the day or year (e.g., those who work outside the home will not be available in the afternoon; those who migrate for work may be away for months at a time). Deciding to only interview these particular persons will increase the logistical effort for the survey team by requiring them to make multiple visits to the same household. In such cases, it might be necessary to adapt the questionnaire and its questions to the persons available. Moreover, even if the required person is available, it is not always the case that the researcher will be able to talk to her. For example, in some cultural contexts, it might be difficult to talk to a household's daughter-in-law without including the mother-in-law as well.

In many countries, the latter is the more senior in the hierarchy of the family and she might not allow her daughter-in-law to talk alone to strangers—or she may be convinced that she, as the elder woman and higher-ranking family member, has more important things to say than her daughter-in-law and answers instead of her.

In most studies, it will be more feasible to choose a preferred respondent for the entire questionnaire and record the characteristics of the respondent properly. This allows for later checking on differences in responses induced by different respondent groups. However, this is not necessarily required, and other approaches exist. As Collins et al. (2009) show in their *Portfolios of the Poor*, it can be excellent to talk to as many households members as possible, as they have done in their study.⁸

People within a household differ on the depth of their knowledge on certain topics ---> Choice of respondent challenging when covering different topics



Triangulation, i.e., seeking insights regarding the same question from different sources of information, can also be fruitful. For example, if you want to examine the percentage of children attending school in the community, as is proposed as an indicator to assess the impact of microinsurance on the educational attainment of children, it might be worthwhile talking to both parents and school teachers. If you want to examine indicators related to health-care utilisation, it might be insightful to not only talk to household members, but also to medical staff from the local health post and the district hospital.

The way of formulating and phrasing a question is strongly related to the kind of household members or community members identified as potential respondents. A questionnaire that addresses the mother-in-law of the household may look different than a survey tool that is designed for the household head. It is important to ensure that the participants do, at least, understand the questions in the way that you want. This might be

a challenge and might have different dimensions like—amongst others—a cultural, an intellectual, as well as a practical one. Take the following examples:

- How many people live in your household? This question is probably most basic for all potential questionnaires, since many indicators do not look at the whole household, but at particular attributes *per household member*. For example, this does apply to the indicator mentioned above on *total number of visits to inpatient services (with at least 24 hours hospitalisation) per household member within the last month* used to assess the outcomes of health-care utilisation for health microinsurance. For this, it is crucial to know the number of household members. However, the concept of household might differ in different contexts, and often extended families living from various incomes might live under one roof. Are extended families to be considered a single or separate household? Are families' main breadwinners, who migrate for work most of the year, part of the household or not? And what about nephews from rural areas staying with their better-off uncles during their education? In order to yield more accurate replies, specification of the concept of household could be introduced in the question. A commonly used phrase for

clarifying the concept of household is *people sleeping under the same roof and eating from the same pot*, defining the economically relevant unit *household* better.

- Estimating the household's income: Households in the informal economy often have fluctuating and seasonal incomes, originating from a mix of income sources. Obtaining an overview can be difficult. For example, farm income is largely reinvested into agricultural inputs and, thus, not disposable income; seasonal income is difficult to capture, and so on. Therefore, for the core impacts and outcomes described above, income as an indicator is avoided as much as possible. Instead, an indicator such as expenditure is used and can serve as proxy for income and to assess a household's wealth. However, seasonality patterns also play a role. Food expenses can vary depending on the availability of a household's own crops or spending on religious or social aspects might differ across seasons. **Regardless of the approach taken, it is important to define clear time-reference periods, which are, as little as possible, prone to recall bias, i.e., the human tendency to forget things that have taken place at some time in the past.** Moreover, indicators assessing the households' assets are recommended for evaluating its living standard, since

Defining an adequate recall period is important for reliable results

these are usually more static over time and less subject to seasonal changes as the indicators described above.

- Inquiring about illnesses in a household is also a challenge, as it depends on the household's perception of an illness, and specifying what constitutes an illness is difficult. A precise definition might not be too important, however, in the context of most surveys: the *perception* of being sick might provide the necessary information about a household's potential desire to see a doctor or not and thus actually deliver relevant results for the specific survey focus. Inaccuracy in reporting by larger and smaller households might differ: a respondent from a two person household is more likely to recall the various illness episodes than a respondent of an eight person household. This potential bias, however, is hard to avoid. In order to reduce recall bias, regardless of

household size, it is also important to choose a reasonable reporting period, e.g., illnesses in the household within the last month. Recalling health events longer back might actually work only for larger events like hospitalisation. Furthermore, there are some illnesses for which it is difficult to inquire due to social stigma. Information about the prevalence of human immunodeficiency virus / acquired immunodeficiency syndrome (HIV/AIDS) in a survey is certainly a challenge.

In many parts of the world, people—in particular in remote areas—do not speak the official main language of the country in their daily life, but rather a dialect or other local language. This should be considered when designing the survey tools. If questions need to be asked of the participants, or a discussion is moderated, this should always be done in the local language that the people speak. However, this language might change from village to village. Slight variations in meanings of words can induce unwanted connotations. For instance, a translation of the word “community” into Oriya, a language mainly spoken in Orissa—a state in Eastern India—gave a different connotation to the respondents in one region of that state when it was translated by a native speaker from a different region. In previous civil unrests in that region, that specific translation

carried the connotation of belonging to a specific political community—a meaning not intended in the survey. However, it may be difficult to find local researchers who speak this language well and are suitable to conduct your study. So, this is, consequently, a practical challenge that has to be overcome for your study. Pretests might be helpful for this.

Also, under cultural and religious aspects, alignment of the survey tool is necessary. For example, fasting periods for different religious groups need to be taken into account when asking about the average number of meals eaten per day in the last month or the frequency of eating meat, as suggested as indicators for the quantity and quality of food. Depending on the particular country and region and cultural and religious background, questions that are too direct might need to be avoided for particular topics. In particular, if they are dealing with taboos (for example, abortion or contraception). There is also a need to adapt your language, or the language of your tools, to the education levels of respondents. In developing country contexts, most evaluators are significantly better educated than the participants in their trial, who often have no more than primary school education—if any at all. Keeping your sentences and questions short and simple, and avoiding overly complex words and

concepts, will help avoid confusion at the other end. For example, the impact of microinsurance on peace of mind is suggested for assessment by indicators like *changes in risk prioritisation through risk ranking exercise, perception about confidence in the future, and perception about financial vulnerability*. For all three indicators, at a first glance, rankings seem to be most appropriate, as suggested in the tables above. However, sometimes the use of scales (such as a Likert scale, requiring the respondents to choose one of five ranks) overburdens the respondents. Finding simple concepts or supporting visual aids is key. Using a language and approach that is understood is not only important for quality answers, but also for gaining trust.

10.6.3. Knowing the microinsurance scheme not only on paper but in practice

In order to best design surveys, it is also important to know the manner in which the microinsurance scheme works *in reality* and not only rely on how it works on paper or how it has been planned to work. Take, for example, a health

microinsurance scheme that initially required en bloc affiliation of the whole household. When examining equity of health and health care amongst different household members, it is important to know whether this requirement has been put into practice without exceptions, or whether affiliation of single household members was possible, or other exceptions to the rule were made. This may not only influence the content of your questions, but also helps avoiding misunderstandings on both sides. Therefore, to look at the scheme in practice and also at the context of the scheme is invaluable for interpreting results, drawing conclusions, and better identifying aspects which potentially influence the scheme's impact.

10.7. Conclusions

Experts in the microinsurance arena consider the core impacts and outcomes described in this chapter as critical for impact evaluations in microinsurance. We hope that this set of core impacts, outcomes, and indicators can find their way into many impact assessments in microinsurance and

Knowing how the scheme works in reality – not on paper – is key for drawing right conclusions

help enrich our understanding on those issues quickly.

The set of standard indicators should remain dynamic and adaptive, developing over time. Moreover, we are convinced that other core effects exist that have not yet been explored or identified, but are important.

Such a standard set of indicators, when turned into a tool, needs to be adapted to cultural, religious, social, and intellectual circumstances of the respondents. A survey tool assessing the impact of microinsurance should never be translated one-to-one from one study or one study site to the other, but needs to take into account the particular situation of the community, as well as the subgroups the scheme is targeting.

We believe that having a set of standard indicators will help evaluators better design their studies and create improved comparability between different studies (e.g., for the purpose of meta-analysis). Both better studies and improved comparability can lead to gaining more and deeper insight into the change in welfare that microinsurance can provide for the poor and the challenges related to this.

References

Appui au développement autonome (ADA), Belgische Raiffeisenstichting (BRS), and Microinsurance Network. 2011. Social Performance Indicators for Microinsurance. <http://www.microfact.org/social-performance/>

Collins, D., J. Morduch, S. Rutherford, and O. Ruthven. 2009. *Portfolios of the poor: How the world's poor live on \$2 a day*. Princeton, NJ, and Oxford, UK: Princeton University Press.

Foddy, W. 2008. Constructing questions of interviews and questionnaires: Theory and practice in social research. Cambridge, UK: Cambridge University Press. Reprint.

Hintz, M. 2010. *Micro-impact: Deconstructing the complex impact process of a simple microinsurance product in Indonesia*. Saarbrücken: Südwestdeutscher Verlag für Hochschulschriften.

International Labour Organization (ILO). 2012. What is child labour? <http://www.ilo.org/ippec/facts/lang--en/index.htm>

Organisation for Economic Co-operation and Development (OECD). 2002. Glossary of key terms in evaluation and results based management, from the Development Assistance Committee (DAC) Working Party on Aid Evaluation. <http://www.oecd.org/dac/evaluation/2754804.pdf>

Operationalising impact evaluation: from theory to practice

Divya Parmar and Manuela De Allegri



11.1. Introduction

Impact evaluations in the field of microinsurance can employ different research designs. Impact evaluation designs can be experimental, quasi-experimental and non-experimental, qualitative, or a combination of these. Increasingly, quantitative approaches, whether experimental or quasi-experimental, are often combined with qualitative data collection and analysis processes. The resulting designs are described as mixed methods. Irrespective of the specific design adopted, the common objective across all impact evaluations is to assess the “causal” relationship between the programme and a set outcome of interest.

This objective clearly distinguishes impact evaluations from other forms of evaluations, like implementation evaluation, which focuses on how an intervention is being implemented, or performance evaluation, which focuses on assessing whether a programme achieved the objectives it originally set to achieve.

All research designs used to study impact evaluations have specific strengths and weaknesses. For example, propensity score matching can only reduce selection bias on a limited set of selective measurable indicators used for matching the controls to the cases, whilst difference-in-difference (DID)

designs can only control for selection bias due to time invariant indicators. Similarly, although randomisation is often considered to be the best available option to estimate impact, it may not always be the most appropriate design to apply in a particular situation (Shadish et al. 2002). For instance, it may not be politically feasible to offer microinsurance to some individuals whilst denying it to others who have similar needs and could equally benefit from the intervention. The lesson we have learned over the years is that the most appropriate design balances theoretical considerations with field circumstances.

Focusing now on the elements that ought to be considered when translating the theory of impact evaluation into practical fieldwork, this chapter unfolds over five sections. Each section covers one specific aspect relevant to designing, planning, or conducting an impact evaluation in the field of microinsurance. Firstly, we introduce and discuss the elements that need to be considered when choosing an adequate research design, depending on field circumstances. Secondly, we describe the resource requirements for conducting an impact evaluation and how those requirements may reflect on the choice of the research design. Thirdly, we address pragmatic considerations on sample size, tool development, and actual data collection and analysis

processes. Fourthly, we address the ethical implications of the various designs and briefly review procedures for ethical clearance in impact evaluation. Lastly, we address the interface and relationship between the research team and the concerned policy makers. Across sections, we draw from direct field experience to complement conceptual models with practical illustrations.

11.2. Selecting an appropriate impact evaluation design

Purpose of the impact evaluation

The ultimate aim of any impact evaluation is not only to determine whether, and by how much, the microinsurance programme yields an impact on a defined outcome through the estimation of causal relationships, but also to provide the necessary evidence base to inform further policy developments.

If a microinsurance programme yields

beneficial impacts, it is important to understand what factors in the environment enabled change to take place and to assess whether the programme can be scaled up or replicated elsewhere. Conversely, if the programme does not yield the expected impacts, it is important to identify potential bottlenecks and barriers to change that can be removed. Hence, good impact evaluations should employ mixed methods: experimental or quasi-experimental methods focus on capturing the impact, whilst qualitative methods can reveal how change took place (White 2009) in the light of the context in which the programme is embedded (Hintz 2010).

Besides its ultimate purpose to inform policy, there are several field-related factors that need to be considered when selecting an adequate impact evaluation design. The most important factors are briefly described and discussed below. These factors are not listed in terms of their relative priority or importance. They are also not mutually exclusive and, by and large, will depend on the same underlying conditions. Their role and the extent of their influence on the design of the impact evaluation largely depends on a case-by-case basis.

The launch of the microinsurance programme and its targeting strategies

We start by looking at how the microinsurance programme is launched.

→ From research to policy: results constitute the basis which informs policy

Two initial important elements to be considered:

- 1) How the microinsurance programme is launched: whether it is phased in over time, or launched everywhere at the same time
- 2) How beneficiaries are targeted and enrolled

The rollout strategy

Phasing the rollout of a microinsurance programme over time presents two advantages. On one side, it eases the early administrative burden on the programme management. On the other side, it offers those concerned with the impact evaluation the possibility of recruiting the control sample from areas/districts/subportions of the target population not yet affected by the insurance launch. In technical terms, the element of an impact evaluation that relies on a phased implementation of the microinsurance programme is referred to as *stepped wedge* (Brown and Lilford 2006).

The robustness of the analysis increases substantially if the phasing can take place according to a randomised design, meaning that areas/districts/subportions of the target population are randomly assigned to be included in the intervention at different time points. Stratified randomisation also represents a viable and,

at the same time, analytically sound design option. In this case, areas/districts/subportions of the target population are first stratified on the basis of a variable (i.e., characteristic) thought to be important, for example literacy rates or level of urbanisation. Then areas/districts/subportions of the target population are randomly selected from the composed strata so that the final sample includes representatives of all strata.

Selecting villages on the basis of political decisions or feasibility concerns represents the least ideal scenario for a phased rollout strategy. The lack of random assignment to the intervention, in fact, imposes an important threat to the validity of the overall study because it introduces a source of bias in the estimation of the impact effect. In situations where no randomisation is possible, however, a phased rollout of the intervention according to political or feasibility criteria

*Insurance rollout
should ideally allow
for randomisation*

is still preferred over a situation with no phased rollout at all. Even in the absence of randomisation, a phased rollout strategy leaves open the possibility of identifying comparable controls, whilst interventions that affect entire regions/districts/target populations at once make it much more difficult for the researcher to identify comparable controls.

Impact evaluations of microinsurance programmes often set the unit to be randomised at a group level (e.g., a village, an agricultural cooperative, a loan group, etc.), rather than at an individual level. This randomisation strategy is known as *cluster randomisation*. There are two strong arguments that support cluster randomisation in the context of microinsurance programmes. Firstly, although the benefits of the microinsurance are for those who buy the microinsurance, others may experience the impacts within that cluster. This may be because the programme “leaks”, contaminating those who are not supposed to receive it, thereby weakening any estimate of treatment difference (Flory 2011). This is particularly relevant for microinsurance programmes as they operate in areas where people are already commonly engaged in informal risk sharing.

Randomising insurance access at the individual or household level may actually lead to underestimating the true effects of insurance due to the potential

Individual randomisation
may lead to
underestimating
insurance
effects

risk shifting from those without access to the microinsurance to those with access. Such a case was observed by Flory (2011) in Malawi, where the poorest benefitted from a microinsurance programme—even though they had not enrolled—because of interhousehold transfer practices. Secondly, individual randomisation requires that the implementing organisations have a complete listing of all potential beneficiaries and are able to manage and control the process of enrolling single members in a randomised manner over an extended period of time. These conditions can rarely be met by organisations operating in low-income settings, especially in rural areas, where much of the bookkeeping is still done exclusively manually. In addition, individual randomisation does pose some ethical challenges since it diverts from the communal orientation of traditional societies. Conflict may arise when two neighbouring individuals or households

are enrolled into the same programme at very different time points.

Cluster randomisation is better suited to reduce the risk of contamination across the intervention and the control arms, since it uses existing barriers (e.g., geographical, group, etc.) to define the very same allocation of the intervention (De Allegri et al. 2008). Continuous monitoring is required on the part of the impact evaluation team to ensure that the intervention is correctly allocated to the selected individuals or clusters. Mistakes in the allocation of the intervention can jeopardise the estimation of the impact at a later stage. The implications that cluster randomisation bears on the sample size are discussed later in this chapter.

A stepped wedge cluster randomised design was applied to evaluate the impact on access to health-care services and financial protection of a health microinsurance programme in the Nouna Health District in rural Burkina Faso. In this study, the 41 villages and seven urban sectors of the target region were formed into 33 clusters. Small neighbouring villages that shared common ethnic and demographic characteristics were grouped together to form a single cluster. Health insurance was offered in a phased manner, so that each year households in an additional randomly selected 11

clusters received the offer to enrol in the scheme. By year three, the entire target region had received the offer to enter the health insurance programme (De Allegri et al. 2008). The design allowed for minimal contamination across clusters and a robust estimation of the impact of the scheme on the outcomes of interest. At the same time, the stepped wedge design facilitated the scheme implementation, by allowing the implementing agency to focus their social marketing campaign on a restricted number of villages per year.

Targeting strategies

Another important feature of a microinsurance programme that needs to be considered when selecting an adequate impact evaluation design is the targeting of the beneficiaries. Programmes that target the beneficiaries according to an arbitrary cutoff point open the possibility for the impact evaluation team to apply regression discontinuity designs (RDD). For instance, this may be the case when implementing agencies decide to offer microinsurance only to the poorest 20% of the entire population. For RDD, it is assumed that the households who marginally miss the cutoff (for instance, households that lie in the 22% poorest category) are very similar to the targeted households and therefore can be used as controls. RDD represents a viable

design option when enough people fall in the category just above the cutoff point. A commonly used approach is to look at clusters of settlements that are divided by administrative boundaries. The administrative boundary is seen as a random cutoff, where seemingly similar households are offered microinsurance only if they fall on one side of the boundary. In such situations, the remaining households that stay close to the intervention households but are not offered the microinsurance can be used as controls. There are not many examples where RDD has been used to evaluate microinsurance programmes, but there are few examples from similar interventions, like the publically-funded health insurance programme in Colombia called *Régimen Subsidiado* (Miller et al. 2009). In the programme, only people below a certain threshold of the poverty index were eligible to receive full public subsidy to purchase health insurance. By comparing the eligible people with those that narrowly missed this threshold, impact of this programme was evaluated.

DID methods can be applied if the microinsurance programme is launched according to a cluster randomised design or if neighbouring areas are similar enough to the intervention area to be used as controls. Quimbo et al. (2010) applied a DID approach to evaluate the impact of a health insurance

programme for poor children in Philippines. The evaluation study included 30 public hospitals. These hospitals were divided into intervention and control groups, matched according to demand-and-supply characteristics like household income, number of beds, average costs, etc. Both exit and follow-up home interviews were collected at baseline (round one) and then again after the programme was implemented for two years (round two). Identical baseline and round two data were collected from the intervention and control groups for the DID method. After the launch of the microinsurance programme, applying DID can be often impossible or tricky, as the researcher has to depend on existing surveys for baseline data.

If all the beneficiaries are targeted at once, then propensity score matching (PSM), with all its limitations, may be the only available option left as DID or RDD will not be applicable. The researchers have to depend on the “eligible but not enrolled” group to be potential controls for those that enrol for the microinsurance programme (Trujillo et al. 2005; Gnawali et al. 2009).

An intrinsic challenge in the selection of the most appropriate study design rests in the difficulty of reconciling the highest research standards (to which a sound impact evaluation ought to abide) with field requirements and

conditions. Those directly concerned with the implementation of a micro-insurance scheme may be more concerned with programme start-up and survival than with producing evidence on expected and unexpected impacts to inform the wider policy arena. Therefore, they may be reluctant to agree to randomisation, stepped wedge designs, and RDD.

In our experience, **continuous and open dialogue with those in charge of implementation is the only strategy for reconciling implementation needs with those of a sound impact evaluation.** At times, when this fails, research teams in charge of the impact evaluation have no other choice if not selecting the one study design that, although not perfect, best exploits existing field conditions. The objective is always to select the design that presents the fewest threats to validity in the light of existing field conditions. During this process of negotiations between impact evaluators and implementing agencies, the former may often find themselves explaining over and over the concept

of impact evaluation, as this is not so clearly differentiated from other forms of evaluation amongst people working away from research. Likewise, evaluators have to be open to working within the framework of an emergent design, meaning they must be willing to readjust design decisions as field conditions unravel. Emerging designs are proper for qualitative studies. Therefore, they do represent a fundamental epistemological challenge for the positivist impact evaluator, who learned to rely exclusively on experimental and semi-experimental models. Nevertheless, years of reconciling research rigor with field needs have convinced us that often no other way is possible than letting the mind be open to adjust design decisions, within the limits of rigorous research, to emerging field conditions.

Who is commissioning and conducting the impact evaluation?

The ability to negotiate with the implementing agency the most robust study design largely depends on who is

Continuous dialogue between researcher and implementer is key for a mutual understanding and a good cooperation.

commissioning and who is conducting the impact evaluation. Impact evaluation teams may find the dialogue over research designs to proceed better when the need for an impact evaluation has been identified by the agency implementing the microinsurance program and/or when agency itself commissions the impact evaluation. In such cases, implementing agencies are very open to adjust their implementation design to allow for robust estimation procedures because they have a primary interest in the use of the results that will emerge from the evaluation. Large international organisations, such as the International Labour Office (ILO) or the World Bank, often enjoy the privilege of commissioning, and, in some instances, even managing directly both the implementation and the impact evaluation of interventions. This allows them to shape the rollout of programmes in such a way as to derive sound evidence on their impacts afterwards.

Researchers from academic institutions often find themselves facing the exact opposite scenario, being called in by a third party, frequently a ministry or another high governmental institution, to assess the impact of interventions run by local and international non-governmental organisations (NGOs) and/or consultancy groups. In such instances, the risk that the implementing agency perceives the impact evaluation, and,

in turn, the research team conducting it, as a potential threat is great, since its reputation and its ability to acquire future contracts may be at stake. As mentioned above, continuous dialogue is the only means to overcome initial scepticism towards the impact evaluation and the team conducting it. Still, at times this initial resistance cannot be fully overcome, the result being that the research team will have to find its analytical way around field conditions far from enabling an optimal impact evaluation design.

The number of resources that can be mobilised to conduct an impact evaluation also plays an important role in shaping design decisions. Impact evaluations that can count on substantial ad hoc funding can engage teams of expert researchers who can guide proper design decisions. This is often not the case when, due to lack of adequate funding, the impact evaluation is planned and conducted directly by the implementing agency. At times, this may result in impact evaluations that rely almost exclusively on routine monitoring and evaluation data to estimate effects.

Timing of the impact evaluation

Dialogue between the implementation team and the impact evaluation team should take place as early as possible, preferably during the conceptualisation

of the microinsurance programme. Early exchange allows for the impact evaluation to be factored into the implementation design from its very onset, ensuring that adequate conditions are in place to allow for a robust estimation of the effects of the intervention. A sound impact evaluation is one that relies on data collection that takes place both in the intervention and control areas at both baseline and follow-up (multiple follow-ups are possible). If the impact evaluation is not factored into the implementation design early enough, it may become impossible to produce a robust evaluation at a later point in time, either because the programme is launched in a way that there are no appropriate controls (e.g., the microinsurance programme is launched in the entire target region simultaneously) or because relevant data has not been collected from relevant controls already at baseline.

Field circumstances are such that, at times, impact evaluations are decided only once a programme is already in operation. In such situations, information may be available on the areas/districts/population targeted by the microinsurance programme, but most likely not on any comparable controls. To overcome the estimation problems resulting from the lack of data on controls, researchers are advised to search for alternative data sources (e.g., household surveys, census,

demographic health surveys, facility-based data, and programme data) applicable to the control areas/districts/population. If no alternative data source is available, then the researcher has to derive an estimation of the effect by relying on cross-sectional data collected both from the intervention and control individuals and/or households. Propensity score matching, with all its limitations, is probably the only viable option in these situations (Gnawali et al. 2009).

11.3. Resource consumption

General remarks

The different impact evaluation designs all imply different resource consumption requirements. Therefore, selecting the most appropriate design also requires an assessment of the specific resources available. Resources are usually more abundant when funds are made available from a third party, either in the form of a research grant or in the form of a contract established directly between the impact evaluation team and the agency commissioning the study. Such contracts, however, are very rarely directly managed by the agency implementing the microinsurance programme, since such agencies usually operate in conditions of financial limitations and would be unlikely to divert funds from implementation towards research. Such contracts are

more common when the impact evaluation is commissioned by a governmental or supra-governmental institution to an external evaluation team (either a research institution or a consulting agency). Therefore, if on one side, earmarked contracts increase the resources available for field research, on the other side, they limit ownership of the impact evaluation on the part of the implementing agency. The implications of having an ad hoc research team conduct an impact evaluation have already been discussed in detail earlier.

Research team

Team requirements vary substantially depending on the design and on the scope of the single impact evaluations.

Experimental designs that rely on randomisation are very demanding on the organisations implementing microinsurance, as they require a constant managerial and administrative engagement to ensure the correct allocation of the intervention.

Quasi-experimental and non-experimental designs are less burdensome from this point of view. Irrespective of the design that is applied, all quantitative impact evaluations require that people with specific expertise in statistics are an integral part of the research team. Decisions on design,

randomisation, and sampling can only be achieved with the support of expert statisticians. This is especially important, as discussed in detail later, given the frequent absence of formal and complete population data to serve as a sampling frame in most settings where microinsurance is implemented. Likewise, irrespective of the design, the support of statisticians and econometricians is essential during the analytical phases. The need for support during the analytical phases is greater in case of complex quasi- and non-experimental designs, combining multiple designs into one to compensate for lack of randomisation and/or baseline data (Shadish et al. 2002).

Impact evaluations applying mixed methods require an even larger set of expertise to come together in the design and analytical phases. Mixed methods designs, in fact, rely on the application of multiple methods at once and, as

Experimental designs are demanding for the implementer

such, multiple people have to engage in decision making throughout the length of the evaluation. This means that multiple traditions of conducting research are merged into one single study. The process can be extremely challenging, as it entails open dialogue across people with radically different scientific mind-sets. In the experience of the authors, the key success factor when conducting a mixed methods impact evaluation is the mediating capacity of the team leader. Ideally, the team leader herself should be well acquainted with multiple scientific traditions and respect them all equally to be able to facilitate dialogue in teams where ethnographers and econometricians need to reconcile diverging scientific paradigms into a single study.

Likewise, the scope of the impact evaluation determines the range of expertise that will need to be brought into the research team. The scope of an evaluation is determined by the number and typology of outcome indicators that need to be assessed. Complex impact evaluations that aim to capture changes on multiple dimensions (e.g., agriculture, health, poverty) often require large multidisciplinary teams to be involved already at the stage of study design and tool development. We, the authors, have been involved in impact evaluations, bringing together fifteen scientists from at least ten different disciplines. Working on an

impact evaluation with a broad scope of analysis may entail having to facilitate dialogue between clinicians, psychologists, economists, political scientists, geographers, and agricultural experts. In our experience, this dialogue is complex and not always easy, but very rewarding and enriching in the end.

Field team and data collection processes

The human resources needed to design and plan an impact evaluation and to analyse the data emerging from the field are only a minimal part of overall human resources needed to conduct such an effort successfully. The people actually in charge of data collection in the field are key to any impact evaluation. Data collection teams vary substantially across impact evaluations. Much of the variation can be explained in relation to the nature and scope of the single evaluations. For instance, assessing the impact of a health microinsurance scheme on health indicators may require that people with some clinical training handle the data collection processes, whilst this may be totally irrelevant when assessing the impact on poverty reduction of crop microinsurance. Likewise, purely quantitative versus mixed methods designs imply reliance on very different sets of expertise.

The overall educational levels in a given country also influence the selection and composition of the field data collection team. We have worked in settings where field interviewers could, at the very best, be expected to hold a high school degree, as well as in settings where a college degree was the minimum prerequisite to being enlisted as a potential field interviewer. It goes without saying that the initial educational level of the staff recruited to work on field data collection plays a very important role in the overall process. Interviewers with lower education may require longer training sessions. Yet, these may be the people who really know the communities where data collection is taking place and may therefore be much better placed to engage their respondents than their better-educated, often urban, counterparts. In general, it is easier to identify and recruit interviewers to work on quantitative data collection tools than to identify interviewers with sufficient experience to be able to conduct qualitative in-depth interviews and/or facilitate focus group discussions (FGD).

Field data collection often absorbs most of the financial resources available for an impact evaluation. In recent years, the use of digital devices to collect data has shown potential to reduce quantitative data collection costs substantially, whilst also ensuring greater quality of the data being collected. Digital devices, in fact, allow the research team to insert very accurate filters and logical checks in the structure of the questionnaires being used, limiting the possibility of error by the single interviewers. In our experience, two rounds of data collection on a sample of approximately 1200 households are sufficient to recover the costs of purchasing 15 tablets.

*Recruiting interviewers
for quantitative
fieldwork is
easier than for
qualitative*

Digital devices can speed up the process of obtaining data

Through the use of mobile phone lines and digital devices also facilitate immediate data sharing between the field and the impact evaluation team. This process allows the impact evaluation team to carry out constant quality checks and to feed preliminary results to the implementing and commissioning agency much faster than traditional pen and paper surveys. Field data entry represents an alternative to digital data collection, which also allows using time more efficiently and controlling for completeness and quality of data. In this model, a data entry clerk accompanies the interviewers to the field to input data in the computer one day after the actual data collection activity has taken place, at the latest. This allows the data entry clerk to identify missing and inconsistent information in time to be able to verify it directly in the field.

Digital solutions to facilitate qualitative data collection are, unfortunately, not as advanced as those supporting quantitative data collection. In-depth interviews and FGD are easily recorded on tape or using MP3 recorders, but skilled transcribers are still needed to report the recorded text on paper. A number of software solutions are available to automate this process, but in our experience, none of this software provides a valuable alternative to the traditional method of transcribing interviews. The available software, in

fact, mostly operates only in few languages (English, French, and Spanish) and largely relies on voice recognition. Therefore, it is of little or no use at all to researchers working with communities in low-income settings where these languages are not spoken. Similarly, automated translation devices are not sufficiently advanced to afford researchers a faster, yet accurate, translation of transcribed text from a local language into the language of analysis.

Irrespective of whether digital devices or traditional pen and paper surveys are used for field data collection, interviewers are usually organised in microteams of two to three interviewers and one supervisor. When the scope of the impact evaluation entails collecting data from multiple sources (for instance in a health facility and at the household level), it is usually more effective to create specialised microteams than to train all interviewers on all relevant data collection tools. Microteams can be composed of quantitative and qualitative interviewers, targeting different respondent constituencies and data collection needs at the same time. We have widely worked with this system of multiple specialised microteams, targeting communities at once (and thus keeping transport costs low) to maximise the amount and quality of data collected in a day from multiple sources.

Before embarking in the actual data collection, field interviewers receive training (normally two to seven days, depending on the complexity of the data collection tool) and assist the research team in piloting the tool. The pilot phase is essential for identifying potential pitfalls in the data collection tool in time for the research team to modify questions and adjust them to the specific setting where the impact evaluation will take place. Research teams wishing to rely on digital data collection may need to consider one to two additional days of training. The cost for these additional days of training is quickly recovered through faster field processes, higher-quality data (requiring minimal cleaning), and the absence of data entry at a later stage.

Costs for field data collection vary dramatically across settings and it would be unfair for us to pretend that we can provide the reader with realistic price estimates applicable across settings. In general, in low- and middle-income countries, the cost of an interviewer (and similarly of a data entry clerk) varies from US\$5 to US\$50 per day. In most countries, interviewers will expect to receive additional compensation if the data collection takes place in a location that does not allow them to return to their homes in the evening. Accommodation and per diem costs can, at times, be higher than the actual compensation received as interviewers.

On average, supervisors cost 20% to 40% more than simple interviewers. In many resource-poor countries, transport costs are an important cost driver when conducting an impact evaluation. This is due to the poor conditions of the roads which require data collection teams to hire high-quality vehicles, and to the extreme high price of fuel in some settings.

Data requirements also tend to differ depending on the research design. Simple randomisation and RDD designs generally rely on cross-sectional data and data are collected from both the intervention and the control groups. PSM is generally applied to cross-sectional data, but need larger sample of *eligible but non-enrolees* to ensure that researchers can find enough controls that closely match with the enrolees. Stepped wedge cluster randomisation, on the other hand, has the largest data requirement, as data needs to be collected from the intervention and control groups at every step of the cluster randomisation. For example, if the phased rollout takes three years to implement, like the health microinsurance programme in Nouna District discussed before, then data needs to be collected for at least three years. Whilst designing the impact evaluation study, data requirements and duration of the study need to be taken into consideration. These factors will further

have financial implications for the impact evaluation study.

11.4. Sampling

Sampling is probably one of the most challenging operational aspects facing teams conducting an impact evaluation. Different elements ought to be considered and reconciled into one coherent, yet feasible, sampling strategy. This process inevitably rests on an open dialogue between the implementing agency that knows the population it serves with its microinsurance programme, the commissioning agency that often defines the outcomes of interest to be observed, the research team that operationalises the outcomes of interest into measurable variables, and specialized statisticians and econometricians. As mentioned earlier, a rigorous impact evaluation always requires the contribution of experienced statisticians. Even in the absence of a comprehensive research team in situations where the implementing agency conducts its own impact evaluation, the participation of a statistician is essential. This section identifies some of the key aspects to consider in the sampling process, but can by no means replace the contribution of a statistician in the field.

The first step is getting agreement amongst the various stakeholders on which outcome of interest will be

observed. Whilst this may appear to be a simple task to accomplish, in reality this is far from being the case. Implementing and commissioning agencies often aim to produce changes that cannot so easily be observed and measured in the field within the timeframe of an impact evaluation (normally two to three years). Let us take, as an example, a health microinsurance scheme that aspires to reduce maternal mortality by including coverage of facility-based delivery amongst its services. Although maternal mortality rates are still very high in many low- and middle-income countries (LMICs), maternal mortality, per se, is a rare event. This makes capturing differences in maternal mortality due to an intervention extremely difficult, unless the impact evaluation team works in a setting where comprehensive population health surveillance data are available. Similarly, microinsurance targeting agricultural production ultimately aims at reducing poverty, but measuring changes in poverty levels is a very challenging task, since the mere definition of what constitutes poverty in a given setting comes into question.

In these situations, the responsibility of the research team is to engage the implementing and commissioning agencies in a dialogue that allows translating the ambitious objectives of the microinsurance programmes into observable and measurable outcome

Rare events are
difficult to measure
but proxies can help

indicators. For instance, increases in the utilisation of facility-based delivery can be taken as a proxy of expected reductions in maternal mortality. Changes in asset protection can be taken as a proxy of poverty reduction. In managing this dialogue, it is essential that the research team remains open to exploiting all opportunities for data collection available in a given setting. For instance, in some settings, it may actually be possible to measure changes in maternal mortality even in the absence of formal population surveillance systems. In some African countries, due to pressure to achieve the United Nations' Millennium Development Goal 5 (MDG5), chiefs and community health workers have been prompted to keep record of all maternal deaths, allowing research teams to access community-based data, which can be very helpful in assessing the impact of those microinsurance programmes that target women. Once the outcomes of interest have been identified and agreed upon by

all stakeholders, the research team needs to determine at what level these can be observed: institutional level (e.g., health facility, microinsurance implementing agency, agricultural cooperative, etc.), community, household, or individual level. A common mistake in sampling takes place at this stage, which we will illustrate with an example. Let us imagine that a health microinsurance programme seeks to improve access to care for children less than five years of age. Let us further imagine that the outcome of interest has been operationalised as *proportion of children less than five years of age with fever who report to a health facility*. Then, at this stage, the research team incorrectly samples households or mothers instead of sampling children with fever. The mistake derives from confusing the object of the sample itself (in this case, children less than five years of age with fever) with the strategy needed to identify and/or to collect information on the actual sample (in this case, households or mothers). This mistake can easily jeopardise the analysis of the entire impact evaluation, if, for instance, 400 mothers instead of 400 children with fever were sampled for the interview. This indicates the need to clearly align the outcome of interest with the applied sampling strategy.

Parallel to identifying the level at which the outcomes of interest can

be observed, the research team also needs to define with the implementing agency the size of the change that the microinsurance programme is expected to produce. Together with the size of the target population and the study design, the value of the expected change (i.e., the expected difference between treated microinsurance recipients and untreated microinsurance non-recipients) is fundamental for the calculation of the relevant sample size.

At this stage, the statisticians have all the information needed to assist the impact evaluation team in the calculation of an adequate sample size. In LMICs, although not always complete, the data compiled at the institutional level is often sufficient to identify an adequate sampling frame for samples drawn at this level (e.g., health facility, microinsurance implementing agency, agricultural cooperative, etc.). On the contrary, drawing community-based samples (at the household or individual levels) can be very challenging in the absence of a comprehensive population surveillance system. In these situations, researchers lack the needed sampling frame and may need to invest first in enlisting all households and individuals in a given community before being able to draw a representative sample from the same. Alternatively, the impact evaluation team may opt for a more pragmatic strategy

Some experimental designs can rely on smaller samples

and identify the complete sample from a number of selected communities using the “spin the bottle method” or “random walk” (Milligan et al. 2004). The trade-off between the two strategies just described is in terms of accuracy and costs, with the former method being more accurate, but substantially more costly. Experimental designs can rely on smaller samples and can, therefore, substantially reduce data collection costs. In addition, although not ideal, experimental designs can lead to a valid estimation of the effect even in the absence of baseline data, offering yet another opportunity to curb field costs. Experimental designs that rely on cluster rather than individual randomisation, however, require larger samples to allow for analysis to account for intra-cluster correlation (the probability that sampled units within a cluster are more similar to one another than sampled units across clusters). In addition, impact evaluations that rely on cluster randomisation design need to ensure that a sufficient

number of clusters are identified and factored into the design. The power of the estimation is directly proportional to the number of clusters included, as well as to the number of units sampled within each cluster.

Compared to experimental designs, quasi- and non-experimental designs, such as RDD and DID, require larger samples and baseline data. The larger sample is justified by a need to rule out the possibility that factors external to the microinsurance programme are responsible for the observed change and this can only be done by controlling for a large number of possible confounders. Propensity score matching can be applied on cross-sectional data, but due to its need to match intervention and control units at the analytical level, it also requires larger samples.

One last important element to consider is potential attrition over time. Impact evaluations frequently rely on longitudinal data collection methods, interviewing the same communities, households, or individuals several times before and after the launch of the microinsurance programme. If resources are available, it is best to include in the original sample a larger number of communities/households/individuals than the minimum number identified by the statisticians. This strategy safeguards the power of the estimation in case a large number of

sampled units should be lost to follow-up. If this strategy is not feasible, the impact evaluation team should be ready to consider replacing households lost to follow-up in a sample. Again, this process needs to be assisted by experienced statisticians and demographers.

11.5. Ethical considerations

In recent years, it has become impossible to conduct an impact evaluation without first obtaining ethical clearance from an ethics committee affiliated with an academic institution or institutional review board (IRB). IRBs are independent ethics committees or ethical review boards, designated to approve, monitor, and review biomedical and behavioural research involving humans. Requirements vary greatly across boards, but, generally, impact evaluation teams are required to compile documentation, including a detailed description of the research protocol, study tools, informed consent forms, and the resumes of all concerned investigators. Most are also required to describe the process of selecting respondents and ensuring that consent is taken from the respondents or from their legal representatives. Depending on the complexity of the research protocol, ethical clearance is obtained in a period of time that ranges from three to six months. Most IRBs charge for their services. Public

Clearance by an ethical review board is usually required in the country of the evaluation team and the country hosting the evaluation (if different)

and academic IRBs generally charge a nominal fee ranging from US\$2000 to US\$3000 per submission. Commercial IRBs can have substantially higher prices.

Impact evaluation teams affiliated with academic institutions are normally required to obtain clearance from the ethical review board at the same institution for which they work. Most universities in Europe and in North America have developed internal structures that work very rigorously to comply with highest international ethical standards. Impact evaluation teams who are not affiliated with an academic institution normally obtain ethical clearance from commercial ethical review boards. This chapter purposely does not include any reference to any commercial IRBs because we do not wish to advertise one service provider over another. We leave it to the interested reader to look for such services online, making his/her own decisions on the rigour of the process proposed to obtain clearance.

Most LMICs have also established their own independent IRB. Application procedures and payment conditions vary greatly across countries. Impact evaluations that are funded and conducted exclusively at the local level can apply directly to country-specific IRBs. Impact evaluations that receive international funding and/or are supported by international teams normally have to apply for ethical clearance both in the country where the impact evaluation team is based and in the country hosting the impact evaluation.

It is often difficult to understand the need for ethical clearance in the case of studies that do not deal with medical products or otherwise potentially toxic substances. The implicit assumption of impact evaluation teams is that their work will result in no harm. Research teams involved in impact evaluation, however, need to reflect carefully on such assumptions to guarantee that their work results in no harm and that, if possible at all, their work yields benefits for those with greater needs. Even impact evaluation rooted in the

social sciences can raise a number of justifiable ethical concerns, especially when some form of randomisation (whether at the community or at the individual level) and/or purposeful targeting takes place. Both randomisation and purposeful targeting into microinsurance entail that a selected group of communities/households/individuals enjoys the benefits of a programme, whilst others do not. Step wedge randomisation represents an effort to reduce potential sources of ethical conflict, since it ensures that the benefits accrued by one group will be shared with all others within a foreseeable period of time. Both randomisation and purposeful targeting have to be communicated very clearly to the concerned communities. Risks and benefits of such designs have to be discussed openly and in simple terms to allow target populations to appreciate the long-term benefits and support the intervention (Marshall 2007).

An additional ethical concern arises from the fact that, through the various study components embedded in an impact evaluation, extensive data on socioeconomic profile, health, and economic behaviours of communities/households/individuals will be collected and stored in databases. To allow for the identification of the same sampled units over time, in the case of longitudinal studies, data needs to be stored so that it can be traced back to

specific households and/or individuals. This creates a privacy protection concern. To protect privacy, access to the full dataset containing actual names and/or otherwise identifiable ID codes is normally restricted to one or two people in each impact evaluation team. The dataset is then usually cleaned of all personal information before being used for analysis by the rest of the study team. A small compensation, usually in-kind, is frequently offered to households or individuals who participate in long or repeated interviews. In traditional communities with a strong collective orientation, this compensation needs first to be discussed with the community leaders. If community leaders are not consulted, compensation can create conflicts in the community.

*Consider consulting
community leaders when
intending to offer
compensation to
respondents*

11.6. Dialogue with stakeholders and knowledge transfer

At different points throughout our discussion, we have drawn the reader's attention to the need to engage in an open dialogue with the implementing and commissioning agencies as the only means to ensure a sound, yet feasible, impact evaluation. In the last section, we will draw attention to the overall sociopolitical context within which impact evaluation of microinsurance programmes takes place. We explained at the very beginning that impact evaluations are not conducted with the sole goal of assessing a causal relationship between a given intervention and a set outcome of interest. Through the identification of this causal link, impact evaluations ultimately intend to influence and shape policies. Furthermore, sound impact evaluations respond to the ultimate single objective of improving the living conditions of the concerned communities.

Striving to influence policy requires acknowledging that impact evaluations do not happen in a sociopolitical vacuum and that impact evaluation teams need to engage with a number of politically relevant stakeholders, beyond those directly implementing and/or commissioning the evaluation. Relevant stakeholders include

policy makers, national and international agencies involved in neighbouring and overlapping programmes, and, last, but surely not least, the communities themselves. It is important that dialogue with all concerned stakeholders is initiated before the impact evaluation even takes place and is maintained throughout the course of the fieldwork. Consensus can be created by incorporating the concerns of all relevant stakeholders into the design of the impact evaluation and by ensuring that results are fed back at every stage of data collection. Effective knowledge transfer takes place only when communication strategies are adjusted to the specific constituency being addressed at a specific point in time. In addition, open dialogue allows the impact evaluation team to understand what other programmes may take place in the same region, accounting for potential confounding factors when estimating the effect of the microinsurance intervention.

Dialogue with communities and policy makers is crucial for acceptance and uptake of research results

Impact evaluation teams can rely on informal community gatherings to exchange with the communities themselves. For instance, in the Nouna Health District in rural Burkina Faso, the Centre de Recherche en Santé de Nouna (CRSN) holds regular meetings with village chiefs and traditional authorities to ensure that the content of the research activities it leads is properly communicated to the communities and their concerns incorporated into future research. This constant, open dialogue empowers the population, who feel that the work of the CRSN is not only feeding the interest of the political decision makers, but is addressing actual concerns at the community level. The health microinsurance programme, amongst others, was a result of previous studies conducted by CRSN that identified lack of financial resources as one of the leading causes explaining low health-care utilisation. Empowerment translates into community support to research, opening the door to experimental, quasi- and non-experimental designs that are appreciated for their potential to improve the living conditions of the population in the longer term.

Policy briefs, official dissemination meetings, and other formal knowledge brokering activities are more often used to exchange information with policy makers and concerned national and international non-governmental agencies. Scientific publications, either in

the form of peer-reviewed articles or discussion papers, represent only a limited source of knowledge transfer from the academic to the policy making community. Scientific publications are largely needed because they allow for information to be shared beyond the borders of the setting where the impact evaluation took place. As such, they may create interest in microinsurance in other settings or countries, but the academic language they use may at times be inadequate to inform policy decisions directly at the local level.

11.7. Conclusions

Impact evaluations should be designed well in advance, ideally already whilst the microinsurance programme is being designed. Early planning allows the impact evaluation team to integrate the evaluation in the launch of the programme and develop the best available methodology to measure impacts. It provides time to plan and collect baseline data prior to the start of the programme.

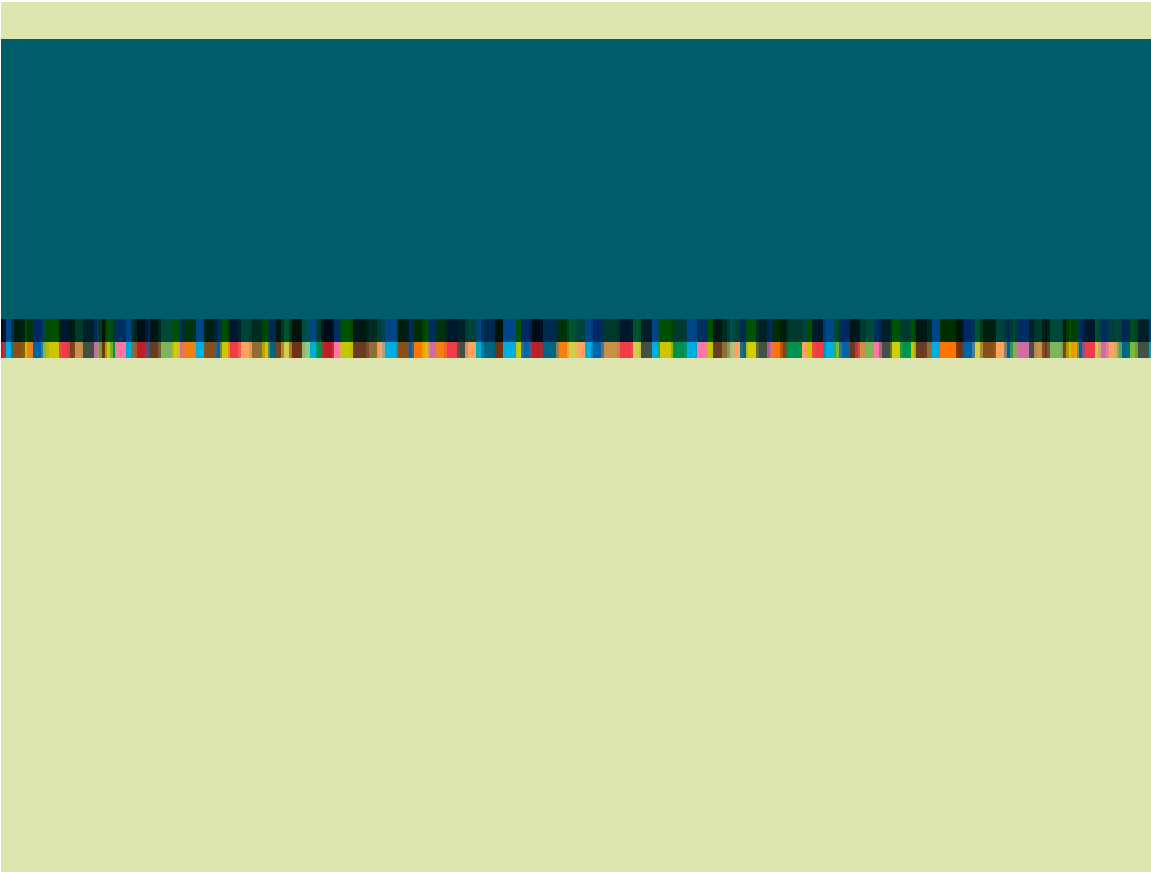
A team best suited to evaluate impacts must consist of field experts and experienced statisticians. Regular consultations with the microinsurance implementing team will ensure that field realities are taken into consideration in the design of the evaluation. The implementing team should be consulted when the research methodology and the sampling strategy are being

developed, as they should fit with the field setting whilst allowing for the best evaluation design to be implemented.

When designing, conducting, analysing, and disseminating the results of the impact evaluation, it is vital that the team adheres to all ethical guidelines and regulations. Impact evaluations not only provide evidence for the existence of impact, but also provide an opportunity to guide policy and influence the living conditions of the communities. This requires that the impact evaluation results be disseminated to all stakeholders—not only policy makers and international organisations—including communities that are at the centre of these programmes.

References

- Brown, C. A. and R. J. Lilford. 2006.** The stepped wedge trial design: A systematic review. *BMC Medical Research Methodology* 6:54.
- De Allegri, M., S. Pokhrel, H. Becher, et al. 2008.** Step-wedge cluster-randomised community-based trials: An application to the study of the impact of community health insurance. *Health Research Policy and Systems* 6:10.
- Flory, J. A. 2011.** Micro-savings & informal insurance in villages: How financial deepening on safety nets of the poor—a natural field experiment. Working Paper No. 2011-008. Becker Friedman Institute for Research in Economics, Chicago. <http://dx.doi.org/10.2139/ssrn.1969504>
- Gnawali, D. P., S. Pokhrel, A. Sié, M. Sanon, M. De Allegri, A. Souares, H. Dong, and R. Sauerborn. 2009.** The effect of community-based health insurance on the utilization of modern health care services: evidence from Burkina Faso. *Health Policy* 90(2-3):214-22.
- Hintz, M. 2010.** Understanding the context is understanding the impact: Evidence from a qualitative microinsurance impact survey in Indonesia. In: *Microinsurance: An innovative tool for risk and disaster management*, ed. E. Morelli, G. A. Onnis, W. J. Ammann, and C. Sutter, 109-135. Davos: Global Risk Forum GRF.
- Marshall, P. A. 2007.** Ethical challenges in study design and informed consent for health research in resource-poor settings. Special Topics in Social, Economic and Behavioural (SEB) Research Report Series no. 5. Geneva: World Health Organization.
- Miller, G., D. M. Pinto, and M. Vera-Hernandez. 2009.** High-powered incentives in developing country health insurance: Evidence from Colombia's Regimen Subsidiado. National Bureau of Economic Research (NBER) Working Paper 15456. Cambridge, MA. <http://www.nber.org/papers/w15456>
- Milligan, P., A. Njie, and S. Bennett. 2004.** Comparison of two cluster sampling methods for health surveys in developing countries. *International Journal of Epidemiology* 33:469-476.
- Quimbo, S. A., J. W. Peabody, R. Shimkhada, et al. 2010.** Evidence of a causal link between health outcomes, insurance coverage, and a policy to expand access: Experimental data from children in the Philippines. *Health Economics* 20(5):620-630.
- Shadish W. R., T. D. Cook, and D. T. Campbell. 2002.** *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.
- Trujillo, A. J., J. E. Portillo, and J. A. Vernon. 2005.** The impact of subsidized health insurance for the poor: Evaluating the Colombian experience using propensity score matching. *International Journal of Health Care Finance and Economics* 5(3):211-39.
- White, H. 2009.** Theory based impact evaluation: Principles and practice. *Journal of Development Effectiveness* 1(3):271-284.





Results of Impact Assessment Studies

Drawing conclusions

Karlijn Morsink and Peter Geurts



12.1. Introduction

When writing up conclusions of research, one ought to always address at least the following three questions:

- 1) What is the concrete answer to each of the research questions?
- 2) What does this mean for the general objectives (contribution to science, policy, and practice)?
- 3) What are the strengths and limitations of this study?

Therefore, we will start with a discussion of research questions and objectives of studies. To draw conclusions about impact, both the effect of the intervention on the studied impact measure should be understood, as well as the mechanisms through which the impact arises. It is important to understand not only the average impact the intervention has on the insured and non-insured, but also the impact on communities, households, or individuals with different characteristics: that is, the distributional impact. These two aspects will be discussed as well. Finally, conclusions of studies

should always be interpreted in terms of the strengths and limitations of the chosen research design. These can be expressed in terms of four validities: internal, external, construct, and statistical conclusion validity. Following a discussion of these four types, we will discuss the implications a particular research design has on the validity of the conclusions.

When drawing conclusions from research that has investigated the impact of microinsurance, it is firstly necessary to define what is meant by *impact*. Here, impact is defined as the final result of an intervention on the wellbeing of a community, household, or individual.

12.2. Finding concrete answers to research questions

The first step in drawing conclusions is to refer to the research questions. Research questions can be descriptive and explanatory in nature. Descriptive questions aim to describe the variables

In drawing conclusions it is important to understand not only the average but also the distributional impact (= impact on different client groups)

that are measured. They often start with phrases such as:

- How many...?
- What are...?
- How often...?
- What percentage...?

Two sample descriptive questions from the area of microinsurance are as follows:

- 1) What percentage of rural Filipino households received a microinsurance claim payout?
- 2) How many households with microinsurance have sold production assets to cope with a health shock compared to households without microinsurance?

Explanatory questions address the causal relation between variables. They start with phrases like:

- What is the effect of...?
- How does...?
- Why...?

In the case of microinsurance, two possible explanatory questions could be:

- 1) How does agriculture insurance effect on-farm risk management activities by small farmers in Ethiopia?

- 2) Why does typhoon insurance change the consumption smoothing activities that households employ? (Morsink 2012)

Research questions about impact always look at the impact, *ceteris paribus*, of an intervention on a certain impact indicator, measured through outcome measures and are, therefore, always explanatory. However, to provide a valid answer to the explanatory question, it is often necessary to answer several descriptive questions in advance that describe all relevant variables in the specific domain of the study. For the explanatory questions described above, the following descriptive questions need to be answered:

- 1) *What on-farm risk management activities do small farmers in Ethiopia undertake?*
- 2) *Which consumption smoothing activities do households employ?*

12.3. General objectives of impact assessments in microinsurance

Beyond the type of question, the objectives of the study must be considered. Scientific studies generally address theoretical issues. However, in most cases, impact studies are more practice oriented and theory is applied to the practical problem addressing the impact of an intervention. Some

examples of objectives include investigating the effect of insurance on improvements in health (Lei and Lin 2009), improvements in health outcomes and economic well-being (Aggarwal 2010), and a theoretical problem, such as the role of production risk in determining the demand for credit (Giné and Yang 2009). These objectives can be broadly or narrowly defined, but are generally not at an operational level. When drawing conclusions, the translation from the operational results to the broader objectives of the study is necessary. For example, when examining the effect of a health microinsurance scheme on the health of the insured, observations regarding health-care treatment seeking behaviour might be relatively easy to receive (e.g., higher health-care utilisation), but do not necessarily give insight regarding the impact of the scheme on health, since there is no general correlation between health care and health.¹ When relating results to the research objectives, it is important to realise that **impact evaluations, especially for policy makers and practitioners, are often about “what works”, focusing on results that imply “evidence-based conclusions that will have immediate policy use”** (Harrison 2011,

626). The implication is that not only is it important to understand which factors influence a certain impact, but also which factors can be manipulated, and to what extent. Take, for example, a study that shows that age, insurance literacy, and trust in the insurance provider have an effect on the impact indicator. Age is a concrete fact and cannot be manipulated, but age groups can be separately targeted, whilst insurance literacy and trust can be directly manipulated. However, the question is: if one wants to achieve a higher impact, which one provides the more efficient and effective investment?

12.4. Specification of theory matters for conclusions

When investigating the impact of microinsurance, it is essential to apply theory about the change that the intervention (microinsurance) can bring with respect to the relevant impact

For policy makers impact evaluations are often about what works

¹ Compare, for example, Jowett, Deolalikar, and Martinsson (2004, 855), who observe a positive effect of health insurance on health treatment seeking behaviour, but come to the conclusion that this does not provide an answer regarding the effect of insurance on health outcomes, which was the objective of the study.

indicators that fit with the objectives of the study. Using theory allows for the development of predictions/hypotheses. On the contrary, lack of theory will lead to a haphazard choice of hypotheses or a lack thereof. The theoretical predictions will allow for the inclusion of potentially confounding variables that may influence the treatment and impact. Even if the study design randomly assigns the intervention to a treatment and control group, the theory is necessary to develop hypotheses about heterogeneous effects/distributional impacts and interpret average effects that are assessed in randomised experiments.

The following section describes an example that demonstrates the importance of using theory about potential microinsurance impacts to develop hypotheses, and to test these hypotheses with the appropriate controls. Without theory, it is difficult to draw conclusions about the impact of microinsurance.

Example: Protecting current consumption or future income?

Shocks lead to fluctuations in consumption due to uncertain expenses incurred by households. This fluctuation prevents households from maximising utility and, therefore, if they are risk averse and rational decision makers (who attempt to maximise

utility), they are assumed to be willing to insure in order to smooth consumption (Pratt 1964; Arrow 1965). In this way, microinsurance smoothed consumption by providing a payout ex-post. One specification of this theory is that an increase in the number of insured households increases welfare (assuming that insurance leads to more consumption smoothing and this effect leads to more welfare).

A further refinement of this theory would not only consider the payout, but would also consider the effect of the payout on the protection of assets. Households generally own a variety of assets and, if they do not have insurance, they may deplete these assets to smooth their consumption (Alderman and Paxson 1994; Morduch 1995). However, these assets are not equally important for the household's future welfare. For example, production assets or human assets (knowledge) are needed for future income, whilst savings are more easily replaceable. Therefore, to create a more complete understanding of the impact of microinsurance on future welfare, the payout should not only be included as a measure of microinsurance impact, but it is also necessary to assess the assets that the payout protects. This is markedly relevant if the objective of the study is to consider the impact on poverty. Dercon and Hoddinott (2004) show that the lowest income

households, in particular, may have to resort to the depletion of the most important assets for future welfare—taking children out of school or selling productive assets—simply because they do not have access to other assets, such as savings or informal risk sharing networks.

12.5. Looking at distributional impacts

When investigating the impact of microinsurance on poverty and inequality, as is often the case for policy-oriented impact research, it is important to consider heterogeneous or distributional impacts. The distributional impact is the impact of microinsurance on (groups of) households and individuals with certain characteristics. It is important to use theory to develop hypotheses about heterogeneous impacts, chiefly because it can contribute to an advancement of scientific knowledge about mechanisms leading to effects. In a context

of policy for poverty reduction, this is particularly relevant because policy objectives for microinsurance are often related to outreach of insurance to previously uninsured households. However, extremely poor households may also be households that refrain from taking up microinsurance. If statements are made about average effects of microinsurance, these may significantly overstate or understate the effect of microinsurance on the intended target group.

There is ample evidence that existing welfare distributions, by influencing access to financial services, may have consequences for the impacts of financial services on certain (groups of) households or different (groups of) household members. For example, Greenwood and Jovanovic (1990) show that high-income households are better positioned to take advantage of financial services than low-income households are. They show that, after the introduction of the financial services, there is initially an increase in the inequality of impacts that then decreases over time. When drawing conclusions, it is important to consider that average effects observed in the impact study may well hide inverse effects for different groups of households or household members. The following examples illustrate potential distributional impacts of

Highlight whether results differ for different groups

microinsurance based on factors influencing microinsurance demand.

- Microinsurance has the potential to stimulate investments in higher-risk, higher-return activities, which is especially needed for risk averse households. However, evidence suggests that especially risk averse households (which are also the poorer households) are less likely to take up microinsurance (Giné et al. 2008; Clarke and Kalani 2011; Cole et al. 2013). Even if potential impacts of microinsurance are high, expected demand for this particular group is notably low, which may imply relatively low impact of microinsurance on poverty reduction for this group.
- Credit constraints are another factor that may influence the impact of microinsurance. Households with credit constraints are already less likely to invest in economic opportunities because of lack of credit for investment. If these credit constraints also lead to less take-up of microinsurance (because of inability to pay for insurance premiums), then the impacts of microinsurance may be relatively low for credit constrained households.
- The ex-ante effect of insurance, which allows people to shift from income smoothing to consumption smoothing (e.g., increased investment in agriculture under

insurance), is assumed to arise from an increased feeling of security (Dercon et al. 2008). Trust in the credibility of the insurer may impact the ex-ante effect of insurance. Households with low trust in the insurer, even if they have insurance, may not feel secure about the insurance paying out. In this case, the ex-ante effect of the insurance for households with low trust may be extremely low.

- Distributional impacts are also important because of the apparently strong effects of social capital and networks on insurance uptake (Jowett 2003; Cai 2012). However, there is no conclusive evidence about the mechanisms leading to these observed effects. Because social networks are also the vehicle for informal insurance, it is important to understand why these effects arise and whether they influence the impact of the insurance. For example, it can be imagined that a household with a strong social network is more likely to get access to microinsurance. If certain households are excluded from these networks, the impact of microinsurance may be particularly low for these households. These effects may be even stronger if the exclusion from the networks also prevents the households from accessing informal risk sharing opportunities. The

households may then be left exceptionally vulnerable to shocks.

- Hintz (2010) points to the additional financial burden that microinsurance may have for female members in order to pay the premiums, whilst male members of the household are benefiting from the payout. In this example, the welfare effect of microinsurance on the overall household (average effect) may be positive or zero, hiding a heterogeneous effect for the male and female household members (negative for women, positive for men).

The above examples of the distributional effects of microinsurance show that theoretically relevant factors about insurance are essential to interpret microinsurance impacts and increase our scientific knowledge. Furthermore, theory is important from a policy perspective (focusing on welfare, poverty, and inequality) because theory allows for the investigation of distributional/heterogeneous impacts of interventions.

12.6. Strengths and limitations of microinsurance impact assessments

When drawing conclusions, one needs to present a discussion of the research design and its implications for the validity of the results. Imagine, for example, a study that concludes,

When reporting results the 4 types of validity should be considered: internal validity, external validity, construct validity, and statistical validity

based on a regression analysis, that health microinsurance adds 12% to the likelihood that female clients in rural Nigeria visit the doctor when they are sick. The confidence we have that the impact (visiting the doctor) is caused by the insurance and not by some other factor (e.g., those who enrolled were generally more cautious about their health) is an example of internal validity. The extent to which the findings can be generalised to other areas, clients, and products (e.g., men in Latin American cities) is an example of external validity. The confidence we have that the indicator (the response to a survey question about seeing the doctor) accurately represents the intended concept (seeking health treatment in case of illness) is an example of construct validity. The confidence we have that the result—the statistical effect of 12%—is derived from the correct

application of the adequate statistical tools and techniques is an example of statistical conclusion validity.

In other words, whilst internal validity refers to the causal relationship between variables, external validity is “the validity of inferences about whether the cause effect relationship holds over variation in persons, settings, treatment variables and measurement variables” (Shadish, Cook, and Campbell 2002, 38). Construct validity refers to “[e]nsuring that the variables measured adequately represent the underlying realities of...interventions linked to processes of change” (Leeuw and Vaessen 2009, XV), or ascertain that you measure what you intend to measure (as it is defined in theory). Statistical conclusion validity concerns the quantitative techniques to ensure the degree of confidence about the existence of a relationship between intervention and impact variable and the magnitude of change (Leeuw and Vaessen 2009, XV).

12.6.1. Four types of validity

When drawing conclusions from research as a whole, all four types of validities need to be optimized. However, focusing on one type of validity often implies relaxation of other validities. In this respect a trade-off has to

be made.² This implies that conclusions should pay ample attention to the consequences of these trade-offs for the validity of the results. In the following section, we will look more carefully into the four measures of validity and discuss them for different research designs.

Internal validity

As mentioned above, internal validity is “about proving causality, i.e., proving whether observed covariation between A (the presumed treatment) and B (the presumed outcome) reflects a causal relationship from A to B” (Shadish, Cook and Campbell 2002, 38). Or, to make it simple, did A cause B? Take, for instance, a research question that addresses the impact of health microinsurance on health outcomes. Expected utility theory predicts that

*Internal validity:
did A cause B?*

² See also the discussion of using a mixed-methods approach by Leeuw and Vaessen (2009).

Randomised experiments produce the highest internal validity

people with a higher probability of experiencing a loss (sick people) are more likely to take up insurance. If health is not included in the study as a control variable (which in practice is difficult to do), the study may conclude that people with microinsurance are significantly more likely to get sick than people without insurance (reflecting the fact that sick people are more likely to self-select into the insurance than healthy people). However, it would not be internally valid to conclude that take-up of microinsurance causes a decrease of health status. Experiments with random assignment can avoid this problem by randomly distributing sick and healthy people to treatment and control group.

As another example, empirical evidence from developing countries suggests that risk averse people are less likely to take up microinsurance (Cai et al. 2010; Clarke and Kalani 2011; Dercon et al. 2011; Cole et al. 2013). Imagine an observational study where the effect of microinsurance on reduction of out-of-pocket payments is studied on a sample of households with and without insurance (self-selected). Failure to control for risk aversion may understate the effect of microinsurance on

out-of-pocket payments because less risk averse households are more likely to take up microinsurance and less likely to take preventive measures. The lack of preventive activities may lead to a higher demand for health care and higher out-of-pocket payments. In the observational study, this would not have led to problems with internal validity if risk aversion and prevention activities had been included as control factors. If they had not been included, however, observed effects would have been understated. Here again, in an experiment where microinsurance is randomly assigned, risk aversion and preventive activities would have been assumed to have been randomly distributed over the treatment and control group. Hence, their effects would have appeared in the error term, automatically leading to an internally valid conclusion about the average effect in the sample.

External validity

With external validity, one considers the generalisability of results. Policy makers, donors, and practitioners who are interested in the impact of microinsurance on poverty and vulnerability are often interested in the

External validity: are the results also valid for other settings? (generalisability)

transferability of the scheme to other settings (e.g., other regions or other socioeconomic target groups). In many cases, they are not interested in the particular scheme under study, but in microinsurance as an intervention in general. For example, imagine a study of a self-managed, community-based microinsurance scheme that concludes the scheme hardly increases health-care utilisation. High levels of social control in the community prevent sick people from visiting the doctor because of the existing belief that using the health fund implies spending the money “of the others” or fleecing the other community members. This result is specific to the type of scheme (self-managed, community-based) and the specific community (high levels of social control). The results are, therefore, difficult to generalise to other schemes or communities with different levels of social control and, hence, the external validity of the results is low.

Another limitation to external validity is variation in the treatment variable: the insurance scheme. Let's take a look at a few studies investigating the impact of health insurance on out-of-pocket expenditures. Wagstaff and Pradhan (2005) show a Vietnamese health insurance scheme reduces household out-of-pocket expenditure for health. Chankova et al. (2008) found that health insurance reduces out-of-pocket expenditure for inpatient care. Finally, Jütting (2004) found reduced out-of-pocket expenditures for poor people who are members of community-based health insurance schemes. However, contradictory evidence also exists. Wagstaff (2007) and Wagstaff et al. (2007) found no impact of insurance on out-of-pocket expenditures and Chankova et al. (2008) found that health insurance does not reduce out-of-pocket expenditure for outpatient care. Some of these authors, in their conclusions, provide explanations for findings which appear to contradict theoretical predictions about the effect of insurance on out-of-pocket expenditure. For example, Chankova et al. (2008) assume that differences in impact of outpatient and inpatient care on out-of-pocket expenditures are caused by the benefit package and availability of co-payments. Wagstaff et al. (2007) think that out-of-pocket expenditures are not reduced because health care is sought more often by people who are insured, health care

*Internal and external validity are
often competing*

provided by the health-care facility is more expensive since the insurance has been introduced, and not all health-care costs are always fully covered by the insurance payout. These examples show that the specifics of the insurance product and context are essential for interpreting results.

It should be noted that internal and external validity are often competing: a high degree of internal validity can be achieved by highly controlled experiments. However, due to the relatively artificial setting required, it is difficult to generalise or transfer results to other settings, natural ones in particular.

Construct validity

Construct validity refers to, "[e]nsuring that the variables measured adequately represent the underlying realities of... interventions linked to processes of change" (Leeuw and Vaessen 2009, XV). Drawing an internally valid conclusion

*Construct validity:
is the theoretical construct
indeed answering the
key question correctly?*

about the wrong concept will only blur our understanding of the impact. For example, a common measure for ex-post welfare effects is out-of-pocket payments. Whilst the amount of out-of-pocket payment may be similar for two households, for one household, it may come from their savings, whilst another household may have sold a cow. In terms of the impact on poverty reduction, the effect of the insurance for these households may be different. Therefore, if one is interested in studying the impact of microinsurance on poverty reduction, out-of-pocket payment is an indicator with low construct validity.

Another example is where measurement of direct welfare effects caused by microinsurance show a positive effect. If, at the same time, microinsurance crowds out other informal insurance mechanisms, the overall welfare effect may be small, or even zero if microinsurance fully substitutes informal insurance. A valid construct would not be the effect of microinsurance on welfare, but the effect of all insurance mechanisms (formal and informal) on overall welfare.

Statistical conclusion validity

Statistical conclusion validity concerns the application of correct quantitative techniques to ensure the degree of confidence regarding the results, in

Statistical validity: can the statistical method chosen create a clear relationship between intervention and impact, and state the magnitude of change?

particular, the existence of a relationship between intervention and impact variable and the magnitude of change (Leeuw and Vaessen 2009, XV). For example, when ordinary least square (OLS) estimations are used, the statistical conclusion validity is low and potentially leads to biased estimates when the dependent variable is a dichotomy, often the case in impact studies (Harrison 2011). In this case, a causal analysis should be applied to take into account the dichotomy of dependent variable, e.g., a logistic regression. Other examples are studies which conduct regressions (such as OLS, probit, and logistic regressions), but fail to address issues with self-selection and endogeneity leading to biased estimates. Another important element of statistical conclusion validity is the consideration of the size of the effects in relation to other variables included in the model, and the contribution of the effect to the likelihood that a unit in the population, for example, a potential insured household experiences the studied impact. Since statistical validity refers to statistical methods only, this type of validity cannot be applied to qualitative methods, or to quantitative yet non-statistical methods.

12.6.2. Validity for different kinds of research designs

In the validity of a research design as a whole, all four validities (in case of explanatory studies) or three validities (in case of qualitative studies or quantitative descriptive studies) need to be optimized. However, in reality, focusing on one type of validity often implies relaxation of other validities. The manner in which this is done depends on the question the study attempts to address and how state-of-the-art the existing theory used is. In any case, these choices and their implications for the validity of the study's conclusions need to be carefully discussed. In the following section we will discuss

Strengthening one validity often implies the relaxation of other validity types

the potential strengths and limitations of the main research designs (experiments, quasi-experiments and observational studies, and qualitative approaches) that ought to be discussed when drawing conclusions.

Experiments

Experiments are studies in which interventions are deliberately introduced to observe their effects. In impact studies of development interventions, a special kind of experiment is often applied, namely a randomised controlled trial (RCT). In an RCT design, firstly, a population is selected from which subjects are sampled. This is followed by a random assignment of the sampled subjects to either experimental (treatment) or control group conditions. The random assignment to either treatment or control conditions is assumed to lead to a random distribution of other omitted, potentially confounding factors over the treatment and control group, leading to an assumed zero effect of these factors on the outcome indicators. If potential threats to internal validity, such as spillover effects or contamination, have been accounted for in the research design, the internal validity for an RCT is high. High internal validity often comes at the expense of external validity, which is often low for an RCT. Low external validity can be problematic from a policy perspective

because results cannot be generalised to other contexts.

A comparison of the mean in the treatment group and the mean in the control group will lead to the average treatment effect for the population sample that has been selected before randomisation into the treatment and control condition. Therefore, the choice of the population and the method of sampling from this population will play a role in the size of the observed effect and, thus, merits a critical discussion in the conclusions. For example, an RCT investigating the effect of randomly assigned crop insurance will lead to different results if the sample population is drawn from villages where the predominant source of income comes from a recently opened mine than when the sample population is drawn from a village that has agriculture as the main income source.

Although RCTs allow for high levels of internal validity with regards to average effects (hence their popularity), without an understanding of underlying mechanisms leading to these effects, they are unlikely to lead to advancements in scientific knowledge or contribute to policy development (Deaton 2010). In addition, the average effects are difficult to interpret if theory about the effect does not contain relevant confounding factors (Imbens and Wooldridge 2009). This is certainly the

case for microinsurance, where there are many unresolved scientific challenges, such as the negative effect of increasing risk aversion on insurance demand (Giné et al. 2008; Ito and Kono 2010; Clarke and Kalani 2011; Dercon et al. 2011; Cole et al. 2013), a potential non-monotonic effect of risk aversion on insurance demand (Clarke and Kalani 2011; Dercon et al. 2011) or compositional and interaction effects (Dercon et al. 2011; Morsink 2012).

Even though this does not change the result about the average effect from an RCT because the confounding variables are assumed to be randomly distributed over the treatment and control group, the average effect may be “hiding” effects for specific groups of people with certain characteristics (the confounding factors). This is less of a problem if the effect is known (negative effect for women for paying a premium and positive effect for men who receive

the payout) (Hintz 2010). The problem increases if effects of confounding variables and potential interactions are not understood (for example, observed negative effect of increasing risk aversion on insurance demand).³

Quasi-experiments and observational studies

In the case of rare risk events, resource constraints, or other limitations, quasi-experiments may be the most appropriate research design. Quasi-experiments are similar to experiments in that they also take a sample from the population and have treatment and control groups, but are “quasi” because the treatment and control conditions are not randomly assigned. As a result, unknown, omitted factors that are part of the control or treatment group may influence the observed effects of the treatment on the outcome, resulting in biased estimates with potentially lower internal validity. Therefore, omitted variable bias ought to receive significant attention when discussing the sampling procedure in the conclusions of quasi-experiments. For example, a common method to select the treatment and control groups in quasi-experiments is to use procedures of matching, such as

Results reporting average effects may hide the effects on specific groups of people

3 See, for a discussion of the importance of theoretical models for interpreting RCT results, Pawson and Tilley (1997); Deaton (2010); and, for examples of RCTs with theoretical models underlying them, Duflo, Hanna, and Ryan (2008); Todd and Wolpin (2006); Attanasio et al. (2010).

propensity score matching (PSM)⁴ on single or double differences (DD),⁵ or through stratified random sampling. For example, Lei and Lin (2009) and Aggarwal (2010) use PSM, and Jowett et al. (2004) use a stratified random sample to investigate the impact of health insurance. When drawing conclusions, it is important to realise that the internal and external validity of the results of PSM and DD depend on the extent to which theory—supported by previous empirical evidence—is used to construct the treatment and control group. For example, neither Lei and Lin (2009) nor Aggarwal (2010) explain how matching criteria were derived from theory, potentially threatening the internal validity of the results (Shadish et al. 2002, 164).

When drawing conclusions from quasi-experiments, reference should be made to the statistical checks for potential omitted variable bias such as econometric techniques of Heckmann (1978; 1979) or Altonji, Elder, and Taber (2000). When these problems

are adequately addressed, the study's internal validity comes close to that of an RCT. Harrison (2011, 631) shows, based on Benson and Hartz (2000, 1878) and Concato et al. (2000, 1887), that estimates of treatment effects in well-designed observational studies do not overestimate the treatment effects in comparison to RCTs.

Quasi-experiments do not rely on random assignment, implying that respondents have chosen insurance voluntarily or have decided to take it up, which leads to problems with self-selection. However, the fact that data is collected from a sample that is representative of the population in the natural setting implies that external validity is often higher. For example, the objective of the study conducted by Lei and Lin (2009), applying PSM, is to contribute to the question of whether

*The natural settings
in which quasi-experiments
are carried out may lead
to higher external validity*

4 In propensity score matching a statistical comparison group is constructed based on a theoretical probability of participating in the treatment, using observed characteristics. Respondents are then matched on the basis of this probability, or propensity score, to non-respondents. The average treatment effect of the programme is then calculated as the mean difference in outcomes across these two groups (Khandker, Koolwal, and Samad 2010, 53).

5 Double-difference (DD) methods, compared with PSM, assume that unobserved heterogeneity in participation is present, but that such factors are time-invariant. Data on project and control observations before and after the programme intervention can cancel out this fixed component (Khandker, Koolwal, and Samad 2010, 71).

the Chinese health insurance scheme (NCMS) adds more service and better health, given China's objectives to create universal health coverage. In the selection of their sample, they cover nine provinces and four counties per province, which vary in terms of geography, economic development, public resources, and health indicators, with the objective to create a representative sample of the target group of NCMS. This sampling strategy leads to potentially high external validity, which is remarkably relevant in this study that has an objective of drawing conclusions about the Chinese target population (Lei and Lin 2009, S40).

Qualitative methods

Qualitative methods, such as focus group discussions, participatory research, key-informant interviews, and case studies are well suited for exploring the measurement of variables and mechanisms and for hypotheses development. Qualitative methods are valuable alone or as a supplement to quantitative research (Leeuw and Vaessen 2009, XIV). When drawing conclusions from qualitative research, these studies can provide high levels of construct validity. Furthermore, although internal validity is generally lower than with RCTs, it can be argued that their internal validity is high with respect to the single observation or case study. However, external validity

is especially low because studies are not based on a probability sample. The low external validity also threatens the value of high internal validity because no statements can be made about how specific or general the internally valid observation is. Statistical conclusion validity does not apply to qualitative research. Despite the limitations of qualitative studies, they are, nevertheless, valuable for developing hypotheses about impacts of microinsurance on the lives of the insured and uninsured households. *Portfolios of the Poor* (Collins et al. 2009) demonstrated the complexity of seemingly straightforward questions about the financial lives of the poor and the importance of gaining an in-depth understanding of these questions. Hintz (2010) already found evidence, in his evaluation of Payung Keluarga in Indonesia, of the fact that microinsurance can impact social and human assets. Hintz (2010) points to the impact of microinsurance on extra burden for female members to repay, whilst male members of the household are benefiting from the payout. Another example is the potential of increased spending on social assets because payouts increase power of some households in communities relative to others (Hintz 2010). To be able to draw internally valid, generalisable conclusions about these effects, the impact of the insurance on these factors would have to be quantitatively investigated.

Triangulation

Triangulation is a method in which information is gathered on the same concepts via different designs or observational methods. A few examples of triangulation are: 1) combining laboratory experiments with field experiments or observational studies, 2) combining interview data with archival data, and 3) combining surveys with content analysis of written documents. These additional triangulations, when conducted properly, can increase internal, external, and construct validity in all research, but are notably common in qualitative research (Yin 2003).

When drawing conclusions, one needs to specifically address how the triangulation contributes to increasing a specific validity. For example, Lei and Lin (2009) attempted to triangulate by conducting multiple estimation strategies combining different quantitative data sources: household data from a household survey, information from a key-informant in the community-level government, and county-level information. They estimated the impact of the health insurance programme on preventive health-care utilisation (Table IV, panel 3, Lei and Lin 2009, S37) and found a positive, significant effect for all estimation techniques (OLS, individual Fixed Effects (FE) estimation, instrumental variables, and PSM). The fact that all estimates, despite using different data sources,

showed similar and significant results increased the internal validity of this finding. However, if the different estimations show contradictory results, this will reduce the internal validity of the findings. For example, Lei and Lin (2009) also estimated the impact of the health insurance on the probability of visiting folk doctors. Although all estimations showed a negative effect, two out of five estimations were insignificant (Table IV, panel 8, Lei and Lin 2009, S37). The internal validity in the latter example is, thus, not as strong as the internal validity in the previous example.

Morsink (2012) provides another example of triangulation. She uses administrative data from an insurance company with local government data on typhoon experiences in villages to support the empirical analysis of survey data about the causal effect of peer experiences with insurance claim payments on demand for insurance. Internal validity of the results increased because the insurance and local data were used to confirm the fact that experiences of peers (of a household) with insurance claim payments preceded the households' insurance purchase.

12.7. Conclusion

This chapter is intended for policy makers and practitioners to better understand and assess the conclusions that

are drawn (or can be drawn) from impact studies. To do this, it is not only necessary to discuss results in terms of the four validities of the research design, but they should also be explicitly linked to the objectives, research question, and theoretical model or framework.

Research conclusions need to address at least the following three questions:

- 1) What is the concrete answer to the research question?
- 2) What does this mean for the general objectives (contribution to theory and practice)?
- 3) What are the strengths and weaknesses of the study in terms of validities?

Research questions about microinsurance are always explanatory, but descriptive studies (qualitative studies) are often necessary to explore measurement of variables and develop hypotheses. Research questions about impact are often defined in relatively broad terms and have to be put into practice before they can be measured. When drawing conclusions, the results related to operational variables have to be translated back to more general concepts and objectives. Impact studies are about what works, which implies that the conclusions should not only pay attention to factors influencing a certain

impact, but also to which factors can be manipulated and to what extent.

Interpreting research results in terms of the theoretical model or framework (and potentially conflicting ones) is essential because similar observations, from different theoretical specifications, can imply different impacts. Furthermore, theory specifies potential confounding factors which are important for understanding mechanisms underlying observed impacts and heterogeneous/distributional impacts for households with specific characteristics. The latter is very important from a development policy perspective, where the interest is often about the impacts on the poor or previously uninsured. Seemingly positive average effects of insurance may hide contradictory effects for households with certain characteristics. Hence, distributional impacts deserve close attention when drawing conclusions from studies.

The validity of the results should be carefully discussed, both in relation to the research designs as well as in relation to theory. This discussion ought to address internal, external, construct, and statistical conclusion validity, taking the strengths and weaknesses of different designs into account.

References

- Aggarwal, A. 2010.** Impact evaluation of India's 'Yeshasvini' community-based health insurance programme. *Health Economics* 19:5-35.
- Alderman, H. and C. H. Paxson. 1994.** Do the poor insure? A synthesis of the literature on risk and consumption in developing countries. In: *Economics in a Changing World. Vol 4: Development, Trade and the Environment*, ed. E. L. Bacha, 48-78. London: Macmillan.
- Altonji, J. G., T. E. Elder, and C. R. Taber. 2000.** Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools. Paper presented at the American Economic Association Winter Meetings. Boston.
- Arrow, K. J. 1965.** Aspects of a theory of risk bearing. Presented as part of the Yrjo Jahnsson Lectures, Helsinki, Finland.
- Attanasio, O., C. Meghir, and A. Santiago. 2010.** Education choices in Mexico: Using a structural model and a randomized experiment to evaluate Progreso. IFS Working Papers W10/14, Institute for Fiscal Studies. doi: 10.1093/restud/rdr015/
- Benson, K. and A. J. Hartz. 2000.** A comparison of observational studies and randomized, controlled trials. *New England Journal of Medicine* 342(25):1878-86.
- Cai, H. 2012.** Social networks and the decision to insure: Evidence from randomized experiments in China. Job Market Paper. University of California, Berkeley. http://ecnr.berkeley.edu/vfs/PPs/Cai-Jin/web/JMP_JingCai_012812.pdf.
- Cai, H., Y. Chen, H. Fang, and L. Zhou. 2010.** Microinsurance, trust and economic development: Evidence from a randomized natural field experiment. National Bureau of Economic Research (NBER) Working Paper no. 15396. Cambridge, MA. <http://www.nber.org/papers/w15396>.
- Chankova, S., S. Sulzbach, and F. Diop. 2008.** Impact of mutual health organizations: Evidence from West Africa. *Health Policy and Planning* 23(4):264-276.
- Clarke, D. and G. Kalani. 2011.** Microinsurance decisions: Evidence from Ethiopia. Mimeo. Department of Economics, Centre for the Study of African Economies, University of Oxford, U.K.
- Cole, S., X. Giné, J. Tobacman, P. Topalova, R. Townsend, and J. Vickery. 2013.** Barriers to household risk management: Evidence from India. *American Economic Journal: Applied Economics*, American Economic Association Vol. 5(1): 104-135.
- Collins, D., J. Morduch, S. Rutherford, and O. Ruthven. 2009.** *Portfolios of the Poor: How the world's poor live on \$2 a day*. Princeton, NJ, and Oxford, UK: Princeton University Press.
- Concato, J., N. Shah, and R. I. Horwitz. 2000.** Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New England Journal of Medicine*. 342(25): 1887-92.
- Deaton, A. 2010.** Instruments, randomization, and learning about development. *Journal of Economic Literature* 48(2):424-455.
- Dercon, S., J. W. Gunning, and A. Zeitlin. 2011.** The demand for insurance under limited credibility: Evidence from Kenya. http://www.aae.wisc.edu/mwiedc/papers/2011/Zeitlin_Andrew.pdf
- Dercon, S. and J. Hoddinott. 2004.** Health, shocks, and poverty persistence. In: *Insurance Against Poverty*, ed. S. Dercon, 124-136. Oxford, UK: Oxford University Press.
- Dercon, S., M. Kirchberger, J. W. Gunning, and J. P. Platteau. 2008.** Literature review on microinsurance. Microinsurance paper no. 1. Geneva: International Labour Organization Microinsurance Innovation Facility.
- Duflo, E., R. Hanna, and S. Ryan. 2008.** Monitoring works: Getting teachers to come to school. CEPR Discussion Papers 6682. Centre for Economic Policy Research, London.
- Giné, X., R. Townsend, and J. Vickery. 2008.** Patterns of rainfall insurance participation in rural India. *World Bank Economic Review* 22(3):593-566.

- Giné, X. and D. Yang. 2009.** Insurance, credit, and technology adoption: Field experimental evidence from Malawi. *Journal of Development Economics* 89(2009):1-11.
- Greenwood, J. and B. Jovanovic. 1990.** Financial development, growth, and the distribution of income. *Journal of Political Economy* 98: 1070-1107.
- Harrison, G. W. 2011.** Randomisation and its discontents. *Journal of African Economies* 20(4):626-652.
- Heckman, J. 1978.** Dummy endogenous variables in a simultaneous equations system. *Econometrica* 46:931-961.
- Heckman, J. 1979.** Sample selection bias as a specification error. *Econometrica* 47:153-161.
- Hintz, M. 2010.** Understanding the context is understanding the impact: Evidence from a qualitative microinsurance impact survey in Indonesia. In: *Microinsurance: An innovative tool for risk and disaster management*, ed. E. Morelli, G. A. Onnis, W. J. Ammann, and C. Sutter, 109-135. Davos: Global Risk Forum GRF.
- Imbens, G. W. and J. M. Wooldridge. 2009.** Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47(1):5-86.
- Ito, S. and H. Kono. 2010.** Why is the take-up of microinsurance so low? Evidence from a health insurance scheme in India. *Developing Economies* 48(1):74-101.
- Jowett, M. 2003.** Do informal risk sharing networks crowd out public voluntary health insurance? Evidence from Vietnam. *Applied Economics* 35:1153-1161.
- Jowett, M., A. Deolalikar, and P. Martinsson. 2004.** Health insurance and treatment seeking behavior: Evidence from a low-income country. *Health Economics* 13:845-857.
- Jütting, J. P. 2004.** Do community-based health insurance schemes improve poor people's access to health care? Evidence from rural Senegal. *World Development* 32(2):273-288.
- Khandker, S. R., G. B. Koolwal, and H. A. Samad. 2010.** *Handbook on Impact Evaluation: Quantitative Methods and Practices*. Washington, D.C.: World Bank.
- Leeuw, F. and J. Vaessen. 2009.** Impact evaluations and development: NONIE guidance on impact evaluation. Washington, D.C.: The Network of Networks on Impact Evaluation (NONIE). http://siteresources.worldbank.org/EXTOED/Resources/nonie_guidance.pdf
- Lei, X. and W. Lin. 2009.** The new cooperative medical scheme in rural China: Does more coverage mean more service and better health? *Health Economics* 18:S25-S46.
- Morduch, J. 1995.** Income smoothing and consumption smoothing. *Journal of Economic Perspectives* 9(3):103-114.
- Morsink, K. 2012.** Weathering the storm: The demand for and impact of microinsurance. PhD Thesis, University of Twente (The Netherlands).
- Pawson, R. and N. Tilley. 1997.** *Realistic Evaluation*. London and Thousand Oaks, CA: SAGE Publications.
- Pratt, J. W. 1964.** Risk aversion in the small and in the large. *Econometrica* 32(1):122-136.
- Shadish W. R., T. D. Cook, and D. T. Campbell. 2002.** *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.
- Todd, P. E. and K. I. Wolpin. 2006.** Assess Mexico: Using a social experiment to validate a dynamic behavioral model of child schooling and fertility. *American Economic Review* 96(5):1384-1417.
- Wagstaff, A. 2007.** Health insurance for the poor: Initial impacts of Vietnam's health care fund for the poor. Impact Evaluation Series No. 11, World Bank Policy Research Working Paper 4134, Washington, D.C.

Wagstaff, A., M. Lindelow, G. Jun, X. Ling, and Q. Juncheng. 2007. Extending health insurance to the rural population: An impact evaluation of China's new cooperative medical scheme. *Journal of Health Economics* 28(1):1-19.

Wagstaff, A. and M. Pradhan. 2005. Health insurance impacts on health and nonmedical consumption in a developing country. World Bank Policy Research Working Paper 3563, Washington, D.C.

Yin, R. K. 2003. *Case study research: Design and methods*. 3rd edition. Thousand Oaks, London, New Delhi: SAGE Publications.

Reporting and disseminating findings

Heidi McGowan



13.1. Introduction

Microinsurance impact assessments—and the lessons they impart—serve purposes ranging from fundraising to completing programmatic evaluations and advancing the academic discourse. They cannot accomplish these objectives—or, by extension, influence the microinsurance community's knowledge or practices—if they are not effectively reported or if information about their context, methodology, and results is not properly disseminated. To maximise stakeholders' awareness and understanding of impact evaluations and their implications, this chapter presents techniques for creating comprehensible and engaging reports and cost-effective ways of promoting them. As such, it contains sections which discuss the following:

1. The functions of microinsurance impact assessment reporting
2. Content and formatting options
3. Troubleshooting solutions for common issues
4. Methods for developing and implementing successful dissemination strategies

13.2. Reporting functions

Three factors fundamentally shape the optimal content and formatting choices of a microinsurance impact assessment report: its intended purpose,

target audience, and situation-specific requirements, like length restrictions.¹ Along with other aspects of conducting impact evaluations, these factors are best identified during a study's planning phases—because understanding *why* and *for whom* assessments are performed constitutes the basis for doing everything effectively, from choosing the most relevant evaluation questions and metrics to completing fieldwork, undertaking analysis, and producing and disseminating information about reports. Because multiple stakeholders with different needs (including time constraints and familiarity with research designs or analytical techniques) could often benefit from (or require) information about the same assessment, organisations typically produce several reports individually tailored for their particular consumption. To avoid publication bias, however, evaluators must clearly note that each report refers to the same study.²

13.2.1. Purposes

The purpose of a microinsurance impact evaluation encompasses the objectives it seeks to achieve. Examples of purposes include:

-
- 1 The term “report” denotes the numerous ways information about microinsurance impact evaluations and their results can be communicated, including presentations, journal articles, online content, and even street theater.
 - 2 The term “evaluators” signifies people that produce and disseminate reports about microinsurance impact assessments, regardless of their involvement with designing or performing the studies in question.

- *Determining resource allocation.* Impact information about various (and potentially competing) interventions can help organisations apportion their money and staffing resources more effectively.
- *Influencing stakeholders.* Impact assessments shape stakeholders' perceptions and occasionally galvanise them into action.
- *Improving products or scaling up coverage.* With impact results, microinsurers can improve policy design, marketing, sales, and servicing processes—and, by extension, perceptions of client-value and take-up and renewal rates.
- *Educating current and prospective clients.* When presented in accessible formats and comprehensible terms, microinsurance impact evaluations can educate existing and prospective policyholders about potential shocks, appropriate risk management techniques, and ways that microinsurance builds financial inclusion, amongst other insights.
- *Fundraising.* Evaluations provide important support for fundraising purposes because they demonstrate organisations' ability to design, implement, and administer microinsurance schemes, achieve particular types of change, or even conceptualise and conduct impact assessments themselves.
- *Furthering the academic discourse.* Assessments impart information

about whether and under what conditions microinsurance interventions work, whilst advancing the field's understanding of both study designs and analytical tools.

Though not exhaustive, this list contains common and important reasons for performing microinsurance impact assessments. The degree of overlap between categories demonstrates the considerable interrelationships between them. Fundraising, for example, involves influencing stakeholders to make decisions about resource allocation. Reports can therefore have singular or multiple purposes: an article which communicates whether and how coverage offerings enable clients to manage risks, for instance, can concurrently document best evaluation practices for the wider microinsurance sphere.

Clearly articulating a commonly-understood purpose is the first step to reporting successfully. Throughout the report production and dissemination process, this articulation entails making content and formatting choices accordingly,

→ Reports should be tailored to the prospective audience

subject to the target audience’s characteristics and situation-specific requirements, as described below.

13.2.2. Target audiences

Target audiences comprise the intended consumers of reports. Here, table 1 enumerates some of this diverse spectrum of stakeholders, explains the significance that each frequently attaches to microinsurance impact assessment reporting, and presents characteristics to consider when creating reports for their perusal.

Like the purposes of evaluations, consumer categories overlap to some extent (practitioner associations, for instance, include members of the other groups mentioned). Additionally, because of considerable variation both across and within categories (and the fact that each target audience is unique), the entries are only generalisations. Evaluators should always individually assess audiences’ characteristics to determine which formatting and content choices would best facilitate accomplishing their reporting objectives.

Table 1: Prospective consumers of microinsurance impact assessments

Target audience	Significance of microinsurance impact assessment reporting	Characteristics to consider
Microinsurers and reinsurers	• Provides indicators of their products’ performance and those of competitors and comparable organisations elsewhere • Supplies information helpful for improving or scaling up offerings, designing new coverage, obtaining financing, (re-)allocating organisational resources (for example, between different initiatives) and determining whether institutional goals were met • Facilitates maintaining accountability to funders, clients, and members of clients’ communities	Understand microinsurance concepts—though potentially less well in field offices or at community-run schemes—but probably not information about impact assessments; reinsurers’ actuarial training, however, might simplify grasping some of these ideas.
Technical assistance providers	• Determines the effectiveness of their recommendations • Apprises them of what products or related processes, like marketing, work or don’t work under specific circumstances, thereby enabling them to offer more effective services	Technical assistance providers offer a wide array of services that require diverse qualifications; hence—and whilst most probably understand the tenets of microinsurance—generalising about their knowledge of impact assessments is difficult.
Regulators and other policy makers	• Highlights the sufficiency of existing microinsurance-related laws and practices • Indicates whether and what type of changes are necessary	Because the educational backgrounds, technical knowledge, and professional experience of these civil servants vary widely both amongst and within countries, generalising about their acquaintance with microinsurance or impact assessment is again problematic.

Target audience	Significance of microinsurance impact assessment reporting	Characteristics to consider
Practitioner associations	• Confirms or refutes the accuracy of associations' prevailing views • Presents opportunities to develop and refine initiatives, like pursuing advocacy work	Most association members are familiar with microinsurance, but their familiarity with impact assessment designs and analytical techniques varies considerably.
Donors and investors	• Helps to allocate resources • Monitors the performance of—and indicates whether midcourse corrections are required for—ongoing projects • Evaluates the success and cost-effectiveness of completed projects	Some donors and investors specialise in microinsurance, impact assessments, or both; others possess numeracy via project financing expertise that would facilitate learning about microinsurance and impact assessment methodologies.
Academics and students	• Provides information about microinsurance's antipoverty affects—or lack thereof—and place amongst other economic development interventions • Furthers their knowledge of assessment designs and analytical methodologies	Academics and students are amongst the best-acquainted with impact evaluations; whilst their knowledge of microinsurance might be less extensive, they have resources at their disposal to learn more. These resources, such as subscriptions to journals and index databases, additionally afford academics and students unparalleled access to published impact assessments.
Clients and community members in existing and prospective microinsurance markets	• Educates them about risk management, financial literacy, and economic development more broadly • Creates accountability and fosters trust by demonstrating microinsurers' ability or inability to offer both quality coverage and candid communication with customers	Clients and community members are liable to have less educational experience, limited numeracy—and sometimes literacy—and, depending upon the local availability of coverage, no familiarity with impact assessment or microinsurance.
General audiences	• Introduces the concept of shocks and their consequences for low-income people in developing countries • Explains how—and how successfully—microinsurance and other risk management techniques militate against the ramifications of shocks	Members of this group are likely unacquainted with either impact evaluations or microinsurance.

13.2.3. Situation-specific requirements

Situation-specific requirements constitute restrictions to reports' content and formatting options, including stipulations about the medium or language of presentation, length, and the inclusion or arrangement of particular elements. These limits can be

generated both internally (by organisational resource constraints) or externally (by target audiences or dissemination channels like peer-reviewed journals). Regardless of what evaluators deem important to communicate, reports might have to address certain questions, implement the potentially restrictive recommendations of style guides or stay within particular page

limits. Because these constraints are often non-negotiable they must be integrated into existing efforts to achieve reporting objectives.

13.2.4. Content and formatting options

This section presents and explains a variety of content and formatting choices and describes how to use them effectively. Whilst reports are generally produced at the conclusion of impact assessments, finalising content and formatting decisions beforehand both simplifies the report creation process and circumvents the tendency to “cherry-pick” items for inclusion ex-post. Ultimately, complete reports convey information about the microinsurance interventions in question, the assessment and analytical methodologies employed, the data obtained, and the results and their implications in ways that allow consumers to determine the evaluation’s susceptibility to bias and generalisability (or internal and external validity, respectively). As hallmarks of the scientific method, transparency and the ability to replicate the study are essential. Reports must provide the information and access to documentation necessary to allow consumers to reproduce either the conclusions ascertained with the data collected or entire evaluations themselves. Whilst certain formats and target audiences

preclude presenting in-depth technical material, evaluators can maintain inclusiveness, completeness, and uniformity by incorporating links to relevant resources into reports where applicable.

13.2.4.1. Format

Table 2 presents seven formats for microinsurance impact assessment reporting, lists key advantages and disadvantages of each, and references actual examples. The formats overlap to some extent (interviews can be transcribed in magazines, for example), and more benefits and drawbacks additionally exist than are stated (vulnerability to hacking, for example, is a problem for websites), though the reasons provided are believed to be especially salient. Regardless, evaluators are encouraged to brainstorm other potential advantages and disadvantages of using the formats they choose within the contexts of their individual studies.

Transparency and the ability to replicate the study are essential

Table 2: Formats for microinsurance impact assessment reports

Format	Benefits	Drawbacks
Book	Enables extended consideration of study	Expensive to publish; more time-consuming to produce
Peer-reviewed academic article	Provides prestige and scientific imprimatur; meets standards of rigour	Generally requires special expertise to prepare; can be inaccessible/incomprehensible to non-academic audiences; publication process is often lengthy and sometimes requires relinquishing copyright
Practice-oriented report	Comprehensible to a larger number of stakeholders than academic articles	Usually not indexed by databases and therefore more difficult to locate; corresponding assessment may be less robust or rigorously detailed; if in-house, may contain confidential information that cannot be widely shared
Non-academic article, for example, published in print or online	Widely accessible; more comprehensible to non-practitioners	Explaining impact assessment concepts in common language can be difficult
Presentation, including interviews	In-person element; allows for interaction with audience or interviewer and use of multiple modes of communication, such as speaking and slides	Potential time limitations; often dependent upon electricity or technologies that aren't universally available or reliable in developing country settings
Audiovisual programming, like documentary or dramatisation	Provides potentially more engaging alternative to written formats; accessible to illiterate audiences	Occasionally requires access to expensive or unreliable technologies; potentially costly and time-consuming to produce
Webpage	Gives study a location in cyberspace; can provide links to related material; can capitalise on host site's reputation/popularity	Vulnerability to hacking

Besides presentations and audiovisual programming, these formats are noteworthy for their potential inaccessibility to current and prospective microinsurance clients and members of their communities (who often lack the monetary, technological or educational wherewithal to obtain and interpret them). Because these stakeholders are arguably microinsurance's most important, this chapter recommends distilling the essence of evaluations and their outcomes into novel formats

that overcome such obstacles—including posters, banners, comics, wall paintings, flip charts, street theater, interactive games, and even text messages.

13.2.4.2. Content

Determining what to include in microinsurance impact assessment reports is complicated: distinct reporting purposes, target audiences, and situation-specific requirements each

encourage the choice of diverse and sometimes incongruous elements. Because most reports created are academic or practice-oriented papers, this subsection provides a step-by-step guide for constructing them. Specifically, it combines recommendations from diverse sets of multidisciplinary guidelines for writing reports with suggestions unique to microinsurance and quantitative impact assessment methodologies to present a checklist for producing comprehensive peer-reviewed articles.³ As necessary, the chapter provides modifications for reports oriented toward practice.

Widely-implemented reporting criteria encourage standardisation, which promotes transparency and simplifies the interpretation and comparability of reports. This consistency facilitates the production of high-quality systematic reviews and meta-analyses, which integrate information from multiple studies to ascertain the existence of patterns, contradictions, or other relationships that illuminate an intervention's effects and avenues for additional investigation. Because microinsurance impact assessment reporting is relatively nascent, quickly establishing and formalising the use of guidelines will enable more studies to

be reported accurately and completely going forward.

This subsection derives its recommendations from the Consolidated Standards for Reporting Trials (CONSORT) Statement, which suggests 25 items for inclusion in reports about health-related randomised controlled trials (Schulz, Altman, and Moher 2010). Since its debut in the 1990s, standards for reporting other types of studies have proliferated. This subsection also incorporates guidelines from Strengthening the Reporting of Observational Studies in Epidemiology (STROBE),⁴ Transparent Reporting of Evaluations with Nonrandomized Designs (TREND),⁵ International Committee of Medical Journal Editors (ICMJE, formerly the Uniform Requirements for Manuscripts Submitted to Biomedical Journals, or URM),⁶ together with information drawn from my own experience. From this consolidated material, six components are recommended for inclusion in peer-reviewed microinsurance impact assessments. These components are: front matter, introduction, methods, results, discussion, and sources of complete evaluation-related information.

3 Whilst some of these recommendations apply to reporting on qualitative studies, Malterud (2001) and Tong, Sainsbury and Craig (2007) provide more information in this regard.

4 See von Elm et al. (2007) for more details on STROBE.

5 See Des Jarlais et al. (2004) for more details on TREND.

6 <http://www.icmje.org/>.

1. *Front matter*

Front matter comprises either a separate cover page or the space before an article's introduction. It includes:

- The piece's title
- The authors' names, affiliations, and contact details
- Sources of funding and other non-monetary support
- Key words
- An abstract

An abstract enables readers to quickly grasp details about the design, conduct, and outcomes of evaluations, situate them in context, and understand their importance. As such, it contains:

- The study's main hypothesis or research question
- A synopsis of preceding research
- The primary variable of interest, defined clearly
- The evaluation's location
- A description of the study's subjects and how they were selected
- The interventions and analytical approaches employed
- Significant problems encountered
- The results and their interpretation

Despite this multiplicity of elements, an abstract is typically constrained to a small space (for example, by a journal's limitation on word count) and otherwise meant to be succinct. As

a result, authors should never include content extraneous to their reports.

Besides these components, a practice-oriented report also presents recommendations (sometimes with corresponding budgetary ramifications), often using a one page synopsis called an executive summary, which accommodates the additional material instead.

2. *Introduction*

An article's introduction should explain the background and rationale (or objectives) for performing the study concerned. Besides establishing context by providing a literature review of related research, this includes presenting the main evaluation question and any pre-specified hypotheses. A general outline of the ensuing paper should follow.

3. *Methods*

The methods section discusses how the corresponding study was conceptualised and conducted, together with its pre-specified analytical approach. It must contain enough information to facilitate the evaluation's replication or enable readers to verify the results obtained with the original data collected. As such, it includes the following nine components:

- A description of the study's design, with justifications for less common or unintuitive choices given the assessment's circumstances.
- Information about the study's location and implementation team, including the number and type of settings involved, a description of those (for example, loan officers or survey enumerators) who administered the evaluation, and the specifics of any training given to them.⁷
- An explanation of participant enrolment criteria and processes, including a comprehensive description of the study's eligibility requirements and methods and dates of recruitment (for example, self-selection in response to radio advertising). Authors should additionally confirm whether and how subjects' informed consent was obtained.
- All determinants of sample size and structure, such as the desired significance level and statistical power. Besides the resulting target sample size and allocation type (such as cluster sampling), authors should also describe how and why stratification was implemented, if applicable. Every source of data (for example, census results or preliminary surveys) consulted to estimate these parameters should be listed.
- A description of the interventions administered to each group, together with all follow-up conducted. This includes information about how and when each intervention was implemented and all measures taken to minimise contamination, non-compliance, and attrition.
- The definition of pre-specified primary and secondary variables of interest, including how, when, and by whom they were evaluated.
- Changes made to any aspect of the study after commencement, with reasons. These include alterations to the evaluation's design, methods, eligibility criteria, interventions, variables of interest, data collection approaches, and analytical techniques, together with a full accounting of all modifications introduced. Likewise, any occurrences that may have influenced the study's results (such as floods which affected data collection efforts) should be reported.
- The methods used to estimate and compare primary and secondary variables of interest across groups.
- Any methods of additional analysis with reasons for performing them and confirmation of whether they were pre-specified.

7 Many aspects of an assessment's location, such as climate or prevailing cultural norms, can affect its applicability to other contexts. To enable readers to effectively assess an evaluation's relevance, authors must explain each of these features and their significance.

4. Results

The results section describes the actual study participants, traces their

Provide detailed information about the study participants before the introduction of insurance, if possible

progression through the evaluation, and provides estimates of the variables of interest, other pre-specified metrics, and any unexpected effects detected. Accordingly, it contains the following five components:

- *Sample constitution.* Authors should provide the number of people assessed for eligibility and specify how many were rejected, how many were accepted but declined to participate, and how many were accepted and actually allocated. Any discrepancies between the originally planned and actual sample sizes (for example, due to the unanticipated loss of funding) should be explained.
- *Participant progression.* Description of the participants' progression through the evaluation, with specific information about and reasons for losses (such as those deemed unreachable) and exclusions (such as those determined to be ineligible after allocation) during the implementation, follow-up, and data analysis phases.
- *Baseline data by participant group.* Because not even randomisation guarantees similarity across allocated groups (especially for smaller sample sizes), authors should include a table that summarizes the demographic characteristics of groups at baseline. They should likewise briefly describe its contents, paying particular attention to the more pertinent variables and salient features like unexpected values or noteworthy discrepancies between groups.
- *Estimated results.* Estimated results for every planned primary and secondary variable of interest across all participant groups. An outcome summary including each measurement's mean and standard error should accompany every result where possible, together with the number of study subjects included in the corresponding analyses and information about whether they were evaluated with their original assigned groups. Effect sizes with confidence intervals (and, optionally, p-values) should also be presented. To avoid selective reporting, authors should provide results for all pre-specified variables of

interest and not just “interesting” or statistically significant ones.

- *Unexpected effects.* Authors should define and quantify (by stating how many participants were impacted and to what extent) any unexpected effects observed. They should additionally explain how data concerning these effects were collected and detail any analyses subsequently performed.

5. Discussion

The discussion section evaluates a study’s validity and explains and contextualises its results, with practice-oriented reports additionally providing policy or other recommendations as appropriate. As such, it contains the following three elements:

- *Strengths and weaknesses.* An objective examination of the study’s strengths and weaknesses, including a presentation of its limitations and the methods used to address them. Authors should further discuss the results’ precision (or lack thereof) and potential sources of ambiguity (such as inexact instruments used to measure the variables in question).
- *Generalisability.* An exploration of the evaluation’s generalisability. Once internal validity is established, authors should review determinants of the study’s applicability to

other contexts, including its recruitment period and processes, eligibility criteria, participants, location, interventions, implementation, follow-up, and variables of inquiry.

- *Interpretation and contextualisation of results.* After summarizing the evaluation’s key findings and potentially exploring the mechanisms underscoring them, authors should compare them to the outcomes of similar studies (including those with opposing conclusions), and applicable systematic reviews or meta-analyses. Finally, they should describe the assessment’s overall implications in terms of research and practice. Practice-oriented reports can further provide policy suggestions, with or without accompanying budgets.

*Compare findings
to other studies
and extract policy-
relevant lessons*

6. Complete sources of information related to evaluation

To assure transparency, authors must tell readers where and how to obtain the documentation pertaining to their assessments, subject to applicable policies.

13.3. Troubleshooting

Creating reports can be difficult: besides decisions related to format and content, evaluators frequently confront a range of other choices on matters—some sensitive—involving ethics and their evaluations. This section troubleshoots five such problems, including technical difficulties, disappointing results, conflicts of interest, issues involving privacy and proprietary information, and insufficient resources. Whilst struggles with report preparation are commonplace, evaluators can militate against their effects to still produce high-quality outputs.

13.3.1. Difficulties

Evaluations are routinely compromised by factors—like higher-than-anticipated attrition rates or incomplete enumerator training—that impact both the quality and quantity of data collected and the analyses subsequently performed. Whilst concealing or whitewashing these events may prove tempting when preparing reports,

There is an ethical obligation to disclose details about all difficulties faced during an evaluation

evaluators are ethically obligated to themselves, their target audiences, and the overall research community to candidly present and discuss problems pertaining to their assessments. Failure to do so can have serious professional ramifications and—if left undiscovered—repercussions for the field's understanding of microinsurance and, by extension, subsequent research, teaching, and practice. Conversely, presenting and exploring the effect of difficulties cements evaluators' credibility, provides potentially important insights into more effectively designing and conducting subsequent studies, and facilitates communicating whether and how microinsurance interventions actually work.

13.3.2. Disappointing results

Evaluations frequently demonstrate negative, insignificant, or non-existent

Negative or insignificant results are also important and should not be withheld!

effects. Academic journals, however, are widely acknowledged to preferentially publish papers with strong positive results—a phenomenon called publication bias.⁸

Publication bias has two damaging repercussions, both of which impact the results of systematic reviews and meta-analyses. For example, by leading evaluators to shelve studies with other outcomes (a practice known as the “file drawer” problem), it prevents legitimate knowledge from reaching audiences that might benefit. Additionally, it promotes professional misconduct and the production of spurious reports by encouraging researchers to engage in inappropriate activities to obtain publishable results. Even reports in other formats are affected, albeit to a different degree. For example, donors frequently strive to fund positively impactful projects; evaluations that demonstrate otherwise can precipitate reduced or withdrawn support and damage to the recipient organisation’s reputation as an implementer of effective interventions.

Given this reality, it’s understandable that evaluators with negative, insignificant, non-existent or even unexpected findings (vis-à-vis the existing literature or the assessment community’s a priori perceptions) would withhold or “readjust” the studies involved. To promote good evaluation practice, facilitate the release of accurate information, and enable the preparation of more accurate systematic reviews and meta-analyses, evaluators should disregard such impulses and promote their reports as written.

13.3.3. Privacy and proprietary information

Assessments routinely encounter and, through reporting, must occasionally reveal private or proprietary information, including study participants’ personal details and confidential aspects of microinsurers’ products and processes. Such disclosures should only be made when absolutely essential, with the consent of the parties involved, according to applicable laws and other policies, and after considering the potential consequences (such as competitors appropriating divulged business practices).

⁸ A related problem is redundant publication, or the practice of illegitimately republishing the same results as original material.

Evaluators should additionally only release the minimum amount of information necessary (though never subjects' actual names) and exclusively to those who must know it (for example, by producing separate reports or using redaction). Where cultural norms or legal protections are more lenient or non-existent, evaluators must uphold global best practices—not least to facilitate the performance of subsequent research.

13.3.4. Conflicts of interest

Conflicts of interest can arise when a study's researchers or writers (or their respective employers) have financial or other relationships with individuals or organisations that might improperly influence their work. In the microinsurance realm, examples include researchers with ownership stakes (for instance, through stocks) in microinsurers or reinsurers whose products are being evaluated, or relatives that could benefit materially (for instance, by obtaining business) through the selection of particular study sites or populations. Whilst potential and existing conflicts do not necessarily presage impropriety, they must be handled appropriately to assure the integrity of a report's conclusions.

All actual and prospective conflicts of interest must be identified during

Tackle potential conflicts of interest during the planning stage of an evaluation, then eliminate them and/or disclose them in the report

a study's planning phases. In response, evaluators can preemptively eliminate them, decide to disclose them, or exclude the parties involved from participating in the evaluation. If disclosure is chosen, several ways exist for divulging the information. Most academic journals, for example, have specific policies (usually provided on their websites) governing how conflicts must be reported. Without explicit requirements, evaluators should describe remaining conflicts in a separate notification section at the end of their reports. Given the ramifications for assessments' integrity, eliminating and then properly reporting remaining conflicts of interest is particularly important.

13.3.5. Documentation and plagiarism

Properly attributing ideas is a fundamental aspect of good reporting: besides delineating the record upon which assessments are based, it reflects evaluators' integrity and, thus, credibility. Unfortunately, confusion about how and when to cite sources is widespread. Whilst journals and other publishers often require the use of specific style manuals (which provide standardised citation techniques and instructions, as listed on the websites of the publications concerned), evaluators are not always bound by these restrictions. In such cases, the *Publication Manual* of the American Psychological Association (APA) describes when and how to create citations (APA 2010).

Plagiarism, on the other hand, is the act of knowingly or inadvertently appropriating or closely approximating other people's words or ideas without attribution. Whilst attitudes towards plagiarism vary across cultures, the practice is considered a serious intellectual offense, with potentially steep consequences (such as the loss of employment and professional standing). As such, evaluators should properly cite (for example, using the *Publication Manual* described above) all influences on their work.

13.3.6. Insufficient resources

The lack of money, time, technical expertise, and technology can make preparing and disseminating high-quality reports difficult. Whilst these issues should be anticipated and addressed during an evaluation's planning process, several resources are available for unanticipated complications of this nature. For problems involving technical expertise, the consultant rosters, advisory services, and small grants (for purchasing external capabilities) offered by the International Initiative for Impact Evaluation (3ie) and the ILO's Microinsurance Innovation Facility may prove helpful. Similarly, numerous local and international organisations that focus on evaluation, like the Evaluation Cooperation Group, provide links to resources, including funding opportunities, available consultants, and instruction in assessment and analytical techniques. Finally, both Bamberger (2006) and Bamberger et al. (2004) provide suggestions and case studies for conducting faster and lower-cost evaluations that preserve resources for subsequent report-writing and dissemination activities.

13.4. Dissemination

Dissemination strategies facilitate publishing and publicising information about microinsurance impact assessment reports. As such, they contain

three components, all discussed below: making arrangements for publishing and archiving, conducting promotional activities, and completing related organisational tasks (including creating dissemination work plans and budgets). Like other aspects of evaluations, they should be developed during a study's planning phases.

13.4.1. Publishing and archiving

Optimal publishing and archiving arrangements depend upon several factors, including a report's format, purpose, target audience, and the strength and uniqueness of its methodology and findings. Consistent with this chapter's focus, this section presents options relevant to peer-reviewed journal articles and practice-oriented reports below.

13.4.1.1. Academic journals

When selecting journals for article submission, considerations beyond the scope and history of each publication include reputation, turnaround time, and policies regarding issues like copyrights and reader accessibility. This information is available through journals' own websites and *Ulrich's Periodicals Directory*.⁹ Editage's *How to Choose Journals for Submitting Your*

Paper,¹⁰ Wikipedia's *Journal ranking* entry,¹¹ and Suber's *Open Access Overview*¹² provide the details necessary to make informed choices in this regard.

The following, non-exhaustive list of journals have previously published microinsurance impact assessment reports (Magnoni and Zimmerman 2011):

<i>Applied Economics</i>	<i>Journal of Health Economics</i>
<i>BMC Health Services Research</i>	<i>Journal of Risk and Insurance</i>
<i>British Medical Journal</i>	<i>Pediatrics</i>
<i>Health Affairs</i>	<i>PLoS ONE</i>
<i>Health Economics</i>	<i>Social Science and Medicine</i>
<i>Health Policy</i>	<i>World Bank Economic Review</i>
<i>Health Policy and Planning</i>	<i>World Development</i>

Though these titles reflect the predominance of quantitative evaluations and assessments of health-related microinsurance plans, the wide variety of current pipeline studies indicates that more diverse publications will print microinsurance impact assessment articles going forward (Radermacher et al. 2012; Magnoni and Zimmerman 2011).

Before submitting to journals, some evaluators present their articles at

10 <http://www.editage.com>.

11 http://en.wikipedia.org/wiki/Journal_ranking.

12 <http://www.eartham.edu/~peters/fos/overview.htm>.

9 <http://www.serialssolutions.com/en/services/ulrichs/>.

academic conferences or release them as “working papers” to obtain feedback and increase their chances of weathering peer review. Academic conferences are refereed venues to which scholars submit generally shorter papers. Because some publish articles in proceedings that preclude authors from subsequently submitting to journals, evaluators should check conferences’ individual policies before proceeding. Likewise, the term *working papers* refers to preliminary reports that are posted online or circulated informally.

When ready to apply for publication, evaluators should follow the manuscript preparation and submission guidelines on their chosen journal’s website. The potential outcomes (following publications’ sometimes lengthy turnaround times) are typically accepted, “revise and resubmit” (through which comments from editors or referees are addressed), or rejected. To avoid the serious professional breach of redundant publication, evaluators should never simultaneously submit to multiple journals.

13.4.1.2. Practice-oriented papers

Practice-oriented papers are published by the organisations which conduct or fund evaluations, sometimes as part of an existing series (like the Microinsurance Innovation Facility’s *Microinsurance Papers* series) and nearly always

online. Whilst larger establishments often have in-house publishing departments with submission requirements and standardised outputs, smaller organisations can exercise leeway in this respect.

13.4.1.3. Archiving

Online archives or databases provide reports with external visibility and a permanent internet presence, situate them alongside similar studies, make papers published informally easier to locate and offer the administering institution’s imprimatur. OpenDOAR provides a searchable list of academic depositories, though authors should clarify copyright issues before contributing.¹³ In the microinsurance and impact assessment realms, the

Check options for
archiving
your publication
online

¹³ <http://opendoar.org>.

Microinsurance Network, 3ie, Micro-Insurance Centre, and the World Bank also maintain databases or publication lists; please see the organisations' websites for submission guidelines.

13.4.2. Promotional activities

Beyond publishing and archiving, promotional activities facilitate effectively disseminating information about reports. After describing how to select them, this subsection presents a diverse array under the following four broad and sometimes overlapping categories: events, audiovisual programming, written communications, and networking. For evaluators whose reports have policy making implications, it additionally provides information about toolkits and case studies for influencing decision makers.

Two important considerations are central to selecting suitable promotional activities: what information to disseminate and how to communicate it successfully. Deciding whether to enlist partners like donors or practitioner networks is also important because additional organisations can be better placed to reach certain audiences or contribute resources like labour. However, collaboration may also have drawbacks, including sometimes relinquishing creative control. For more information about choosing suitable dissemination activities, please

see Carpenter et al. (2005). With these concerns in focus, prospective promotional activities are listed below.

13.4.2.1. Events

Evaluators can organise their own events including everything from conferences and seminars to workshops, trainings, discussions, performances, launches, and fundraisers—or participate in established ones. Regarding the former (and especially for producing large, complex, or costly events), those without access to planning expertise can consult guides like Kilkenny's (2011). In contrast, existing events related to microinsurance are listed on the Microinsurance Network's online calendar; of these, the most prominent is the International Microinsurance Conference, hosted annually by the Microinsurance Network and Munich Re Foundation.

13.4.2.2. Audiovisual programming

Audiovisual formats include interviews, news briefs, dramatisations, and sound or video recordings of actual events. Because producing them occasionally requires access to specialised and potentially expensive equipment or expertise, evaluators can economise by generating interest from organisations like radio stations that incur the costs of using production capacity themselves (though this approach sometimes

necessitates relinquishing control over how reports are portrayed]. *Working with the media: A best practice guide*, by the Economic and Social Research Council (ESRC) [n.d.], describes how to develop broadcasting contacts and provides tips for presenting research findings through television and radio.

Whilst audiovisual materials can be disseminated via offline conduits like television or theater, the most common and (depending upon the target audience) widely-accessible method is online. Organisations frequently post productions to their websites and external platforms like the sites of partner organisations, professional associations, and even networking and entertainment venues. For example, the ILO's Microinsurance Innovation Facility provides dozens of video interviews in its Knowledge Centre, whilst the Micro Insurance Academy maintains a YouTube channel.

13.4.2.3. Written communication (both online and in print)

Magazine and newspaper articles, website and blog postings, press releases, newsletter features, brochures, and factsheets are some of the many diverse (and comparatively cheap) ways of disseminating information through written communications. In this domain, two widely-read newsletters are the Microinsurance

Network's *Microinsurance Trilogy* and the Microinsurance Innovation Facility's *Knowledge Flash*.¹⁴ For access to the larger microfinance community, evaluators can post content on the Consultative Group to Assist the Poor's Microfinance Gateway.

13.4.2.4. Networking

Social networking, electronic mailing lists, and interpersonal communications comprise the final group of promotional activities discussed. Regarding the former, table 3 lists major groups on LinkedIn, Twitter, and Pinterest that are specific to microinsurance. Because social networking platforms are dynamic, evaluators should verify the status of existing ones and search for others before proceeding.¹⁵ In the microinsurance and impact evaluation spaces, most organisations (including the ILO's Microinsurance Innovation Facility, Microinsurance Network, and 3ie) also maintain their own e-mail lists. E-mail lists independent of institutions abound as well, with India's Development Creative being a good example.¹⁶

14 A list of other microinsurance-related newsletters is available through the *More resources* section of the Microinsurance Network's website (<http://www.microinsurance-network.org/links.php>).

15 Whilst many microinsurance organisations have Facebook pages, Facebook has not yet gained traction among practitioners for disseminating information; neither have Mendeley or Zotero, which provide "academic social networking" services.

16 Compare <http://groups.google.com/group/dev-creative>.

Table 3: Major groups in social networks that are specific to microinsurance

Social network	Microinsurance-related groups/hashtags/boards
LinkedIn	Access to Insurance Initiative; Centre for Insurance and Risk Management; International Microinsurance Conference—Experts and Feedback Group; Microinsurance (operated by the ILO's Microinsurance Innovation Facility)
Twitter	#microinsurance @MIFacility; @MILKnowledge; @CGAP; @ICMIFMicroinsurance; @NetworkFlash
Pinterest	Microinsurance Network (microinsurance)

Interpersonal communications, including physical and virtual meetings, calls, and private e-mails, texts, and instant messages are additional—and more personalised—ways of spreading information about reports. For example, they can be used to complement or reinforce other dissemination activities by targeting influential or well-connected people. Two strategies for communicating effectively in this respect include customisation and having recipients indicate “implementation intentions”, like committing to distribute reports to colleagues, which researchers have found increases their likelihood of following through (Gollwitzer and Sheeran 2006).

13.4.2.5. Impact toolkits and case studies

For evaluators whose reporting and dissemination strategies include influencing policy makers, 3ie and the Overseas Development Institute's step-by-step *Policy impact toolkit* and accompanying case studies provide the instructions and insights necessary to

successfully reach and persuade decision makers in developing countries worldwide. Similarly, the World Bank's *Making smart policy* paper presents case studies of 12 impact evaluations that were leveraged for policy making purposes (Bamberger and Kirk 2009). Because influencing policy sometimes also requires cultivating and channeling public engagement, the Participation Compass¹⁷ describes methods for involving people to achieve a wide variety of objectives and provides links to research and case studies.

13.4.3. Work plans and budgets

Work plans and budgets can help evaluators successfully coordinate and allocate funding for the activities mentioned above. For example, a good work plan specifically and incrementally delineates each of the tasks involved in a publishing or promotional endeavor, names the individuals responsible for completing them, and provides corresponding timeframes and dates

17 <http://participationcompass.org>.

that include leeway for delays or any retrenchment that might occur. Similarly, effective budgets apportion funding—including both direct and overhead expenses—for each of these line items, with a margin added for unanticipated costs. If inflation is problematic, and the sums and time spans involved warrant, evaluators should adjust their estimates accordingly or perform the computations in a foreign—but potentially more stable—currency.

Work plans and budgets also function as yardsticks for evaluating dissemination strategies themselves, for instance by indicating whether planned activities were completed in a timely and cost-effective manner. Quick and inexpensive ways of appraising dissemination activities abound and can provide implementing organisations with valuable insights for future efforts or indications that retrenchment is necessary. For example, evaluators can survey event attendees to gauge what they learned; similarly, after issuing press releases, evaluators can track related news coverage through services like Google News (ESRC, n.d.). By monitoring dissemination efforts in this way, organisations can gather the information necessary to ensure that reports are impactful.

Together, these dissemination strategies will enable evaluators to publish and promote reports effectively

and efficiently. Combined with the report writing and troubleshooting approaches described above, they comprise a comprehensive guide for successfully imparting information about microinsurance impact assessments to diverse target audiences and for wide-ranging purposes. Over time, this information will increase the microinsurance and wider economic development communities' understanding of risk management practices and, by extension, their ability to empower low-income people in developing countries to more capably militate against life's uncertainties.

References

- American Psychological Association (APA).** 2010. *Publication manual of the American Psychological Association*. Washington, D.C.: American Psychological Association.
- Bamberger, M.** 2006. Conducting quality impact evaluations under budget, time and data constraints. Washington, D.C.: Independent Evaluation Group and Poverty Analysis, Monitoring and Impact Evaluation Thematic Group, World Bank.
- Bamberger, M., J. Rugh, M. Church, and L. Fort.** 2004. Shoestring evaluation: Designing impact evaluations under budget, time and data constraints. *The American Journal of Evaluation* 25(1):5-37.
- Bamberger, M. and A. Kirk.** 2009. Making smart policy: Using impact evaluation for policy making. Doing Impact Evaluation Series, No. 14. Washington, D.C.: World Bank.
- Carpenter, D., V. Nieva, T. Albaghal, and J. Sorra.** 2005. Development of a planning tool to

guide dissemination of research results. Agency for Healthcare Research and Quality (AHRQ) Publication No. 05-0021-4. Vol. 4. Rockville, MD. <http://www.ahrq.gov/qual/advances/planning-tool.htm>

Des Jarlais, D. C., C. Lyles, N. Crepaz, and the TREND Group. 2004. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health* 94:361-366.

Economic and Social Research Council (ESRC). n.d. *Working with the Media: A Best Practice Guide*. Swindon, England: ESRC. http://www.esrc.ac.uk/_images/Working_with_the_Media_tcm8-2674.pdf

Gollwitzer, P. M. and P. Sheeran. 2006. Implementation intentions and goal achievement: A meta-analysis of effects and processes. *Advances in Experimental Social Psychology* 38:9-119.

Kilkenny, S. 2011. *The complete guide to successful event planning*. Ocala, Florida: Atlantic Publishing Group.

Magnoni, B. and E. Zimmerman. 2011. Do clients get value from microinsurance? A systematic review of recent and current research. Microinsurance, Learning and Knowledge (MILK) project, MicroInsurance Centre. http://www.microinsurancecentre.org/resources/documents/doc_details/811-do-clients-get-value-from-microinsurance-a-systematic-review-of-recent-and-current-research.html

Malterud, K. 2001. Qualitative research: Standards, challenges and guidelines. *The Lancet* 358:483-488.

Radermacher, R., H. McGowan, and S. Dercon. 2012. What is the impact of microinsurance? In: *Protecting the poor: A microinsurance compendium*. Vol. II, ed. C. Churchill and M. Matul, 59-82. Geneva: International Labour Organization.

Schulz, K. F., D. G. Altman, and D. Moher. 2010. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ* 340:c332.

Tong, A., P. Sainsbury, and J. Craig. 2007. Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Healthcare* 19(6):349-357.

von Elm, E., D. G. Altman, M. Egger, S. J. Pocock, P. C. Gøtzsche, J. P. Vandenbroucke, and STROBE Initiative. 2007. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *The Lancet* 370 (9596):1453-7.

Systematic reviews

Howard White



14.1. Introduction

Systematic reviewing is a rigorous methodological approach for synthesising evidence. It is different from a traditional literature review as it is systematic in identifying which literature is to be included and how the data are synthesised.

Systematic reviews are an important part of evidence-based development. Every single impact study has limited external validity. We should not draw global policy conclusions based on findings about one specific intervention carried out in a particular time and place in a particular way. **Systematic reviews pull together all available high-quality evidence related to a particular question, and so provide more general answers to the policy questions of interest.**

This chapter goes through the stages of conducting a systematic review:

- setting the question
- the search strategy
- quality assessment
- data extraction
- synthesis

Systematic reviews help find more generalisable answers

14.2. Setting the question

The main issue in setting the question is the breadth of the question. We would all like to know the answer to the question, "Does microinsurance work?" But this question is rather broad, does not define what *work* means nor does it answer the question *works for whom?* A good review question is more specific.

A useful way to break down the component parts of the review question is through Cochrane's acronym, PICO:

- Population, which may well be low- and middle-income countries
- Intervention, which may be microinsurance, but is most usually more specific, such as health insurance or even catastrophic health insurance

Start with a precise research question

- Comparator, what is the comparison group, which will usually be “no insurance”
- Outcome, the outcome to be analysed such as health status (Higgins and Green 2011)

So, a suitable question might be, “How do health microinsurance schemes in low- and middle-income countries affect out-of-pocket expenditures, utilisation of health services, and health status?” An example of this is the systematic review on client-value of microinsurance by Microinsurance Learning and Knowledge (MILK)¹, which asked, “Do clients get value from microinsurance?” (Magnoni and Zimmerman 2011)

Some may still consider this question too broad. Several existing reviews are more specific with respect to population or intervention. For example: *Social health insurance for improving access to care for disabled and elderly people in developing countries* (Okebukola and Ogunsakin 2009) and *Community-based health insurance in low-income countries: A systematic review of the evidence* (Ekman 2004).

The question of how specific to make a review question is the choice between lumping and splitting (Gøtzsche 2000, cited in Grimshaw et

al. 2003). “Splitters” argue we should only compare studies that are very similar in terms of design, population, intervention characteristics, and outcome. “Lumpers” argue that broader reviews allow policy relevance since they compare a range of interventions to attain a common goal, allowing policy makers to select the most (cost-) effective intervention relevant to their context. Moreover, broadening review scope also enables generalisability to be assessed across a wider range of contexts and study populations (Shadish et al. 2002; Grimshaw et al. 2003).

Primary impact studies are still thin on the ground for many interventions in low- and middle-income countries.

This fact tends to support lumping over splitting, as questions that are too tightly defined will result in empty, or near-empty, reviews. Better to cover a larger range of interventions and outcomes, even if most of those are

*When few studies exist,
one may lump studies
covering different
contexts*

1 MILK is a project of the MicroInsurance Centre. See <http://ohg.cochrane.org/sites/ohg.cochrane.org/files/uploads/Risk%20of%20bias%20assessment%20tool.pdf>

empty, since at least more terrain has been mapped. When there is more evidence, the case for splitting is stronger.

14.3. The search strategy

The search for a systematic review attempts to locate all potentially relevant literature. The search should be clearly documented, and include both published and unpublished literature.

Five main approaches are used:

- electronic database searches
- screening websites of key agencies
- handsearches
- snowballing
- contacting leading experts

Electronic searches should cover key bibliographic databases, which are:

- multidisciplinary, such as Web of Science and Google Scholar
- specific to international development and microinsurance, including the Joint Libraries (JOLIS) database of the World Bank and the International Monetary Fund (IMF), the British Library of Development Studies (BLDS), and Eldis (hosted by the Institute of Development Studies (IDS)), the database of impact evaluations and systematic reviews of the International Initiative for Impact Evaluations (3ie), and the stocktaking initiative of the Microinsurance

Network's Impact Working Group (IWG)

- specific to social sciences, both general and discipline-specific, such as Social Science Research Network (SSRN), and IDEAS/RePEc and Econlit for economics
- subject-specific, specifying the search string to use in querying these databases is a task requiring some experience, so the assistance of a search specialist should be sought.

Language bias should be avoided. The exclusion criteria cannot rule out studies on grounds of the language in which they were written. A truly comprehensive search would include also search terms in other languages, notably Spanish, Portuguese, and French in international development, because of the sizeable body of primary studies in Latin America and Africa.

Many studies are identified via the screening of websites of key development and research agencies, such as the World Bank's Documents and Reports database, databases of independent evaluation departments of multilateral development banks, 3ie's impact evaluation database, and the websites of the Abdul Latif Jameel Poverty Action Lab (J-PAL), Innovations for Poverty Action (IPA), the World Bank's Development Impact Evaluation (DIME) database, and so on. The team should

be sure to include the websites of agencies specialized in the field, such as the Microinsurance Network, the Microinsurance Innovation Facility of the International Labour Organization (ILO), the MicroInsurance Centre and the Micro Insurance Academy (MIA), in the search.

Handsearching in libraries identifies studies that are poorly indexed. This step involves the handsearching of key journals—though this is becoming redundant as electronic searches include the capacity to search on key phrases and full blocks of text—and of library shelves.

Snowballing includes both bibliographic back-referencing (reviewing references of included studies) and citation tracking (reviewing references in which the included study has been cited). For example, Web of Science, Google Scholar, and IDEAS all allow citation tracking.

Given publication delays in the social sciences, the search should also include contacting key experts in the field for information on recent or ongoing studies.

14.4. Quality assessment

Properly conducted research will identify a lot of studies. The challenge is to identify, from these thousands

of studies, those to be included in the review. A systematic review requires clear inclusion criteria. Aside from being on the topic of the review, these criteria specify the range of acceptable analytic methods used in included studies.

Acceptable methods vary with the review question. Many reviews are “effectiveness reviews”, that is they wish to review evidence of the impact of the intervention. In that case, only studies with valid identification strategies, that is, experimental and quasi-experimental designs, should be included. The search will turn up many papers and reports that have the key words but, clearly, are not microinsurance impact evaluations. Examples would include advocacy material, process evaluations, and other project reports.

But reviews may also address questions such as, “Who buys insurance?”—what is sometimes called an analysis of barriers and facilitators. Such an analysis will draw on a broader range of evidence, both quantitative and qualitative. The methods inclusion criteria relate to how reliable this evidence is. This includes, for example, whether the data collection methods are clearly documented.

It's best to include only robust studies in a review and drop the rest



In all cases, a decision needs to be made on how high to set the bar for inclusion. Setting it too high can result in empty reviews. But setting the bar too low means that conclusions may be based on unreliable evidence.

This process has two steps:

- 1) Screening of titles, which usually allows for disposal of the large majority of the identified titles, as they clearly are not primary studies in the area of interest or they clearly will not satisfy the methods inclusion criteria.
- 2) Abstract or full text review of those titles which pass the screening stage. The number of papers subject to this closer scrutiny is usually in the hundreds. The number of studies ultimately included is usually in the range zero to one hundred, the larger giving the broader the scope of the review. This step is undertaken by two (or more) researchers working independently from one another, with a third to adjudicate in cases of disagreement. The results are summarized in a flow chart, which states how many studies were considered at each stage,

from the initial thousands down to the dozens or fewer finally included.

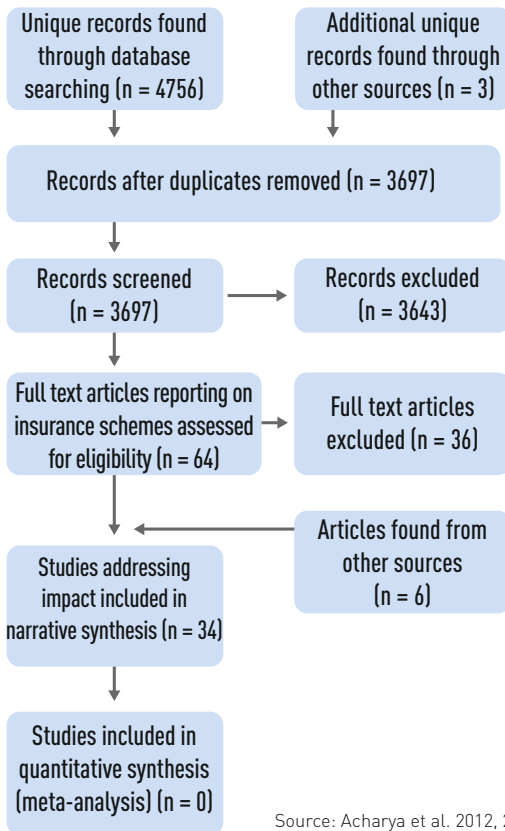
Figure 1 provides an example from a study of national health insurance in developing countries. An initial 4,759 studies were identified, reduced to 3,697 after removing duplicates all of which were screened. Full text review was conducted for 64, of which 34 were included in the narrative synthesis, but none for statistical meta-analysis (Acharya et al. 2012).

The included studies are subject to a more rigorous quality assessment to assess their risk of bias. There are several toolkits available to assist in analysing the risk of bias. These tools do vary in their assessment. For example, the medically-based Cochrane Collaboration only assesses randomised controlled trials as having a low risk of bias.² Taking into account the methods more commonly used in development impact evaluations, 3ie is working on a risk of bias tool.³

2 See <http://ohg.cochrane.org/sites/ohg.cochrane.org/files/uploads/Risk%20of%20bias%20assessment%20tool.pdf>

3 See http://www.3ieimpact.org/media/filer/2012/12/26/jorge_hombrados_and_hugh_waddington_conference-session12-b_3ie_dhaka_colloquium.pdf

Figure 1: Search results for review on national health insurance in developing countries



Source: Acharya et al. 2012, 22.

14.5. Data extraction

Reviews should collect extensive data from each study on populations, interventions and co-interventions, comparison conditions, outcomes, contextual factors, and other effect moderators.

These data should be collected in a systematic study from every study using pre-defined codes. A good systematic

review will code information about intervention design, beneficiary population, and so on. The codebook for data collection should be presented in the study protocol. Coding can sometimes be conducted using text analysis software such as Atlas.ti.

Some qualitative data may not be so amenable to coding. One possibility is to capture these data in matrices, with each row corresponding to a study and each column to an issue being examined. A short narrative summary should be entered into each cell.

For effectiveness review data, extraction involves “extracting effect sizes” into a standard form. An effect size is a statistical measure of the change in outcomes in the intervention group, over the comparison group. A good effect size estimate should be comparable across studies—that is, independent of units of measurement—and only reflect effect magnitude for each study, not other factors, such as sample size. The type of metric used depends on the outcome variable being measured. For continuous outcomes, like income or out-of-pocket health expenditures, we usually calculate the standardised mean difference (SMD), which measures the size of the intervention effect in terms of the number of standard deviations in the outcome variable. In the case of dichotomous outcomes—that is, when the outcome of interest is a categorical value

that can only take the value of 0 or 1—we calculate the risk ratio (RR) or odds ratio (OR), which measures the ratio between two proportions. In the case of using health care, it would measure the ratio between the outcome level in the treatment group (say, 23 percent = 0.23) and the outcome level in the control group (say, 16 percent = 0.16). However, for a variety of reasons, it is not always easy to extract effect sizes from quasi-experimental designs (Waddington et al. 2012).

14.6. Synthesizing the evidence

Synthesis of evidence on effects may be either quantitative, using meta-analysis, or narrative.

Meta-analyses

The traditional and, in international development, still most common, method of quantitative synthesis is vote-counting or goal-scoring—that is, adding up the number of studies finding positive, negative, and no impact. Simply put, vote-counting is inappropriate and can lead to misleading conclusions.

Consider two studies. One has a 95 percent confidence interval of the impact of microinsurance on health status of 0.98 to 1.22. The coefficient is an odds ratio, so the interval contains one. Hence, the finding is of no significant impact at the 5 percent level. A second study has a confidence interval of 1.05 to 1.15, which is thus significant. Both studies have a point estimate of 1.10, and their confidence intervals substantially overlap. But goal scoring would say that the two studies together give “mixed findings” with one study saying the intervention works, and another that it doesn’t. Suppose further that the former study has a sample size of just 78, and the latter study 3,216. Perhaps the lack of significance in the first study is simply because of small sample size (hence the large confidence interval).

The problem with vote counting is that it relies upon “one study, one vote”, taking account of neither the magnitude of the effect size nor its precision. And whilst vote counting by statistical significance may account for precision, it still fails to take into account the magnitude of effect (Littell et al. 2008).

➤ The simple counting of studies with specific findings is inappropriate and potentially misleading

Instead, meta-analysis should be conducted, which does allow for these factors. The general aim of meta-analysis is to combine data from several studies into one aggregated measure, using statistical methods for this. As Glass puts it in his paper coining the term meta-analysis, meta-analysis is the “[...] analysis of analyses. I use it to refer to the statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings” (1976, 3).⁴

In the example above, meta-analysis combines these two sets of results into a single point estimate with a new confidence interval. This confidence interval may be, for example, 1.07 to 1.13. The correct way of combining the evidence shows that the intervention does have a significant impact.

The assumption behind meta-analysis is that all studies are drawing samples from the same underlying population and exposing them to the same treatment. This assumption may be a bit of a stretch for the range of microinsurance interventions implemented by different agencies in very different settings in different countries; so, there is disagreement as to whether meta-analysis is justified.

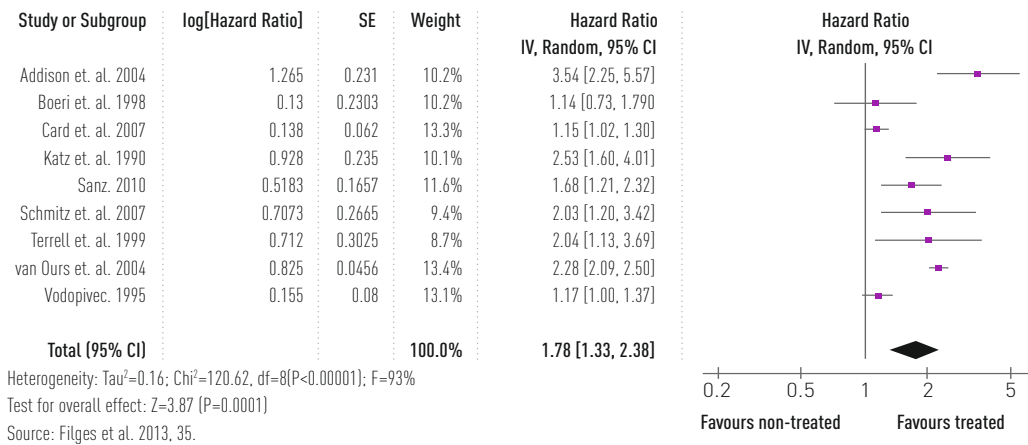
A careful statistical combination of studies can strengthen insights

The answer is surely to proceed with caution. We likely *do* want to know the answer to the question, “Does microinsurance work?” Any review contributing to that answer can helpfully summarise the evidence with the broad generalisations which meta-analysis allows. As for any data analysis, the reviewer has to be on the lookout for heterogeneity. Even if it can be shown that, overall, microinsurance “works”, that does not mean that all types of microinsurance work everywhere for everyone. Analysis of heterogeneity allows for a more nuanced analysis.

Results of a meta-analysis can be displayed in a forest plot (a type of graphical display). In a forest plot, a horizontal line depicts the confidence interval from one study. A vertical line represents either zero or one, depending how the dependent variable is defined. If the confidence interval crosses that line, the study finds no significant impact. At the bottom of the plot is the confidence interval from the meta-analysis,

⁴ Although Glass has introduced the term meta-analysis, the methodological approach was conducted and described long before, for example by Pearson (1904).

Figure 2: Forest plot for impact of limited unemployment benefit duration on employment



represented as a diamond shape. Even if a meta-analysis is not conducted, it can be useful to present a visual synthesis of different studies in a forest plot.

The example in figure 2 looks at the impact of limiting the duration of unemployment benefits for people re-entering the work force. The data illustrate the point made above, as three of the eight studies find no effect, which may be called mixed evidence. But, the meta-analysis shows that people are nearly twice as likely to re-enter the work force in the week or month unemployment benefit ends (Filges et al. 2013).⁵

⁵ This effect is also present the month before it ends, though not two months before. It has disappeared one month after the benefit ends, i.e., those who did not take up a job as the benefit was ending are not more likely to, once it has ended, if they were not able to do so in the month it ended (Filges et al. 2013).

Narratives

When statistical meta-analysis is not sensible or possible, then a narrative synthesis should be used. Narrative synthesis aims to arrive at some “over-arching theory that reconciles the findings” (Hunter and Schmidt 2004, 445). Littell et al. (2008) describe some of the approaches, cautioning that the intended objectivity of the systematic review approach should be retained through transparent decision rules, including how to weight studies in the synthesis. Moreover, narrative reviews should always distinguish between null findings that result from low power and null findings that reflect a genuine absence of treatment effects of policy relevant magnitudes. Unfortunately, many narrative syntheses produce no more than a summary of each included study, and the resulting lack

of actual synthesis is of limited use to practitioners.

A good narrative synthesis is that, a synthesis. There are two ways in which this is done. The first is by coding of intervention design etc. Hence reviews can make statements such as “X percent of microinsurance schemes are explicitly targeted at poor households,” which is useful in understanding the results. Coding may also pick up implementation issues related to impact, for example, “X percent of studies reported that focus groups provided evidence that people found the insurance premiums too high.” Note that being systematic means not “cherry picking” specific study findings from one or two studies.

The second approach relies on matrices described above. Having prepared a matrix, it is possible to read down the column to summarise the narrative findings related to that specific issue.

14.7. Conclusion: conducting reviews

Evidence-based policy requires drawing on all available evidence. Just as policy should not be based on anecdotal evidence, global policy should not be based on a single study of a single intervention. Systematic reviews provide the methodology to ensure that all relevant evidence is captured and considered for inclusion in an objective manner.

Tackling a systematic review should not be taken lightly. They differ from traditional literature reviews in several ways, and most usually are conducted by teams combining sector knowledge, expertise on statistical meta-analysis as well as qualitative synthesis, and a search specialist. Review titles and protocols can be registered with one of several international bodies, such as the Cochrane and Campbell Collaborations and the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) at the Institute of Education, University of London, which will review the various stages of the systematic review, publishing the resulting review in their library.

References

- Acharya, A., S. Vellakkal, F. Taylor, E. Masset, A. Satija, M. Burke, and S. Ebrahim. 2012.** Impact of national health insurance for the poor and the informal sector in developing countries: A systematic review. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London. <http://eppi.ioe.ac.uk/cms/LinkClick.aspx?fileticket=gGjtFFBPX7U%3d&tabid=3346>
- Ekman, B. 2004.** Community-based health insurance in low-income countries: A systematic review of the evidence. *Health Policy and Planning* 19(5):249-270.
- Filges, T., L. P. Geerdsen, A. D. Knudsen, A. K. Jørgensen, and K. Kowalski. 2013.** Unemployment benefit exhaustion: Incentive effects on job finding rates: A systematic review. *Campbell Collaboration* 2013:4. www.campbellcollaboration.org/lib/download/2579/.
- Glass, G. V. 1976.** Primary, secondary and meta-analysis of research. *Educational Researcher* 5:3-8.

Gøtzsche, P.C. 2000. Why we need a broad perspective on meta-analysis. *British Medical Journal* 321:585-6.

Grimshaw, J., L. M. McAuley, L. A. Bero, R. Grilli, A. D. Oxman, C. Ramsay, L. Vale, and M. Zwarenstein. 2003. Systematic reviews of the effectiveness of quality improvement strategies and programmes. *Quality and Safety in Health Care* 12:298-303.

Higgins, J. P. T. and S. Green, eds. 2011. Cochrane handbook for systematic reviews of interventions version 5.1.0 (updated March 2011). The Cochrane Collaboration. <http://www.cochrane-handbook.org/>

Hunter, J. E. and F. L. Schmidt. 2004. *Methods of Meta-Analysis: Correcting error and bias in research findings*. London: Sage.

Littell, J., J. Corcoran, and V. Pillai. 2008. *Systematic reviews and meta-analysis*. New York: Oxford University Press.

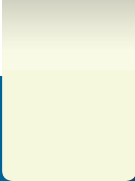
Magnoni, B. and E. Zimmerman. 2011. Do clients get value from microinsurance? A systematic review of recent and current research. Microinsurance, Learning and Knowledge (MILK) project, MicroInsurance Centre. http://www.microinsurancecentre.org/resources/documents/doc_details/811-do-clients-get-value-from-microinsurance-a-systematic-review-of-recent-and-current-research.html

Okebukola, O. P. and J. O. Ogunsakin. 2009. Social health insurance for improving access to care for disabled and elderly people in developing countries. Title registration form, the Campbell Collaboration. <http://campbellcollaboration.org/lib/pr>

Pearson, K. 1904. Report on certain enteric fever inoculation statistics. *British Medical Journal* 3:1243-1246.

Shadish W. R., T. D. Cook, and D. T. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

Waddington, H., H. White, B. Snilstveit, J. G. Hombrados, M. Vojtkova, P. Davies, A. Bhavsar, et al. 2012. How to do a good systematic review of effects in international development: a tool kit. *Journal of Development Effectiveness* 4(3):359-387.



Appendix 1: Guidelines for indicators

Karla Henning and Andreas Landmann

Introduction

These guidelines are intended as a helpful tool for readers of *The Practical Guide to Impact Assessments in Microinsurance* who are evaluating microinsurance products based on the core indicators described in chapter 10 of the Practical Guide. The main focus of the guidelines is on generating data for quantitative methods, as they are described in chapters 5 and 6 of the Practical Guide. However, this is not meant to exclude other ways to define, measure, and analyse data. Rather, we want to present *possibilities* when analysing questions related to microinsurance. In general, this work is a collection of illustrations, research ideas, hints, and warnings. As such, it might (hopefully) be helpful for some without being misleading for others.

Initially, we always present a definition of the indicator, using intuitive language. Alternative definitions may exist, and we leave it to the readers to choose what is appropriate in their context. Next, we lay out potential reasons explaining why the indicator could be affected by microinsurance.

Thinking about explanations can help to form expectations about which type of insurance should affect the indicator most. We list all plausible candidates (emphasising the most obvious in bold font) from the following list of insurance types: health insurance, life insurance, livestock insurance (indemnity or index), crop insurance (indemnity or index), property insurance, and other. We provide examples of how to measure these indicators quantitatively, mostly drawn from well-established global surveys such as the Living Standards Measurement Survey (LSMS) and the Demographic and Health Surveys (DHS) from the World Bank. This ensures comparability with other data sets and applicability of items in a wide range of geographic settings. For the analysis, we propose details, such as subgroup analysis to identify heterogeneous treatment effects, whenever our theoretical considerations lead us to suspect such effects. We finally hint at combinations between the indicator and other measures that might be interesting to analyse jointly.

Outcome: Risk taking behaviour

Indicator: Productive investment as percentage of total income of the household

Definition of the indicator

The indicator measures the share of total income of a household that is not saved or consumed but invested for future production and income generation. Productive investment usually takes place in the context of entrepreneurial activity, and typically consists of investment in more sophisticated productive assets, yielding higher returns.

Theory of expected effects

In theory, high consumption risk deters households from investing in riskier but more profitable activities (Rosenzweig and Binswanger 1993; Dercon and Cristiaensen 2011). Following this argument, microinsurance should enable those with insurance coverage to make higher return investment decisions. In this case, insurance can enhance the productivity and total household income in the long term as a consequence of the change in risk taking behaviour. Cai et al. (2009) state that clients of a Chinese government livestock insurance scheme significantly increase their acquisition of sows, both regarding quantity and quality of the seeds.

Applicable type of risk coverage

- Health insurance
- Life insurance
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- Property insurance
- Other

As productive investment is strongly connected to entrepreneurship, and agricultural production is the main entrepreneurial activity in developing countries, risk coverage by livestock or crop insurance are obvious candidates for creating impact on productive investment. Of course, other types of insurance could also affect the willingness to invest, but their effect channels are less salient.

Measurement

- How much did you spend last week/month to buy tools, equipment, buildings, land, vehicles, fertiliser, seeds, livestock, etc., for your business or agricultural production?

Note: For more precise results, separate question for every input of productive investment, ask for different

periods of the year (cropping seasons): e.g., How much did you spent in total for [include roster with options] during the last cropping season?

Analysis

As this indicator is strongly connected to entrepreneurship and agricultural production, the analysis should mainly focus on economically active individuals. It also may be interesting to analyse heterogeneous effects by risk aversion because the decisions of risk averse individuals should be more affected by insurance. Furthermore, in the analysis of this indicator, income and wealth differences should be controlled for

because these differences could be prevalent factors in higher productive investment. For instance, farmers with more land may be more inclined to buy insurance and also invest more in productive inputs as they are likely to be more affluent and may lose more. As their higher investment in productive inputs could be due to their affluence and not to insurance, it could be interesting to control for the size of landholding or the number of livestock (or business size in case of non-agricultural entrepreneurship).

Combine with

- Total amount of loans taken
- Total amount of saving

Outcome: Risk taking behaviour

Indicator: Total amount of loans taken

Definition of the indicator

The total amount of loans taken captures the current indebtedness of the individual. Loans can be taken from various formal and informal sources (different sources should be clearly identified in the analysis). The indicator refers to the total monetary amount of the loans as well as to the number of outstanding loans.

Theory of expected effects

As a direct effect it is expected that in the context of ex-post risk mitigation, fewer loans are taken up to cushion the shock. As microinsurance is intended to encourage riskier but also potentially more profitable production decisions, a higher amount of loans taken for productive investment in the (agricultural) business could be an indirect effect of microinsurance. Nevertheless, in contrast to this assumption, Gine and Yang (2007) found that rainfall indexed insurance reduced farmer's take-up of loans for purchasing more productive, higher-yielding hybrid maize and better groundnut seeds. As a potential explanation for this unexpected observation, it can be argued that the effects of microinsurance might take longer to materialise

for low-income and more risk averse households, and why more risk seeking or wealthier households may adapt their risk taking behaviour in a shorter period of time (Radermacher et al. 2012).

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Livestock and crop insurance can lead to a change in risk taking behaviour and a higher take-up rate of loans to invest in riskier, but also more efficient, production. On the other hand, a decrease of loan take-up, a method formerly used to compensate losses in agricultural production, can be expected as the risk is then covered by the insurance. Health and life insurance are expected to lead to higher take-up rates of loans, as in case of illness or death, the debt is not transferred to other family members and installments can still be covered due to the insurance pay out, compensating a loss in household income.

Measurement

- How many loans do you currently have (from informal lenders/formal lenders)?
- What is the total amount of your current debts?
- What was the amount of your loan repayments last week/month?
- For what purpose did you obtain the loan (esp. distinguish between business/farm use and personal use)?

Analysis

In the analysis of this indicator, it should be well identified for what purpose the loans are taken up. If they are taken as an insurance substitute (particularly in case of illness, death, crop failure, death of livestock, etc.), it is expected that the take-up of microinsurance leads to a decrease on the total amount

of loans taken. In the direct context of risk taking though, it is expected that insured individuals take up more and higher loans as uninsured individuals as insurance encourages them to engage in riskier and presumably more profitable production. The analysis should take account of a differentiation of contexts in which the loan was taken up. An important distinction has to be drawn between borrowing under stress (ex-post shock) and borrowing preemptively as part of an investment decision (Radermacher et al. 2012). Similar to instances of productive investment, it can be interesting to analyse heterogeneous effects by risk aversion.

Combine with

- Total amount of savings
- Productive investment as percentage of total household income

Outcome: Risk taking behaviour

Indicator: Total amount of savings

Definition of the indicator

The total amount of savings is the amount of household income not spent on consumption or investment but put aside for future use to cover recurring costs, unexpected expenditures, or consumption planned ahead (wedding, dowry, funeral, health costs, education expenses, farming inputs, etc.).

Theory of expected effects

Savings are an important risk mitigation tool for low-income households to protect themselves against shocks and stabilise cash flow. With insurance, precautionary savings are expected to decrease, whilst funds are expected to be preserved by insurance if a shock happens.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

All types of insurances listed are applicable in this case as savings can be intended to decrease a household's risk of suffering a financial shock in many dimensions (health, agriculture, property, etc.).

Measurement

- How much of weekly/monthly household income do you not consume or invest but put aside and save for future use? Provide roster with type of saving (private, institutional, savings group, etc.) and amount and purpose of savings (National Bureau of Statistics, Nigeria 2010).
- What is the amount of your current savings stocks?

Analysis

Due to the expected change in risk taking behaviour it may be interesting to analyse heterogeneous effects by risk aversion. Furthermore, different purposes and intentions behind the saving behaviour should be taken into consideration in the analysis as well as the form of saving (informal, institutional, savings group, etc.,).

Combine with

- Total amount of loans taken
- Productive investment as percentage of total household income
- Total amount of savings (ex-ante risk management strategy)
- Total amount of savings (ex-post risk management strategy)

Outcome: Risk management strategies (ex-ante)

Indicator: Total amount of savings

Definition of the indicator

The total amount of saving is the amount of household income not spent on consumption or investment but put aside for future use to cover recurring costs, unexpected expenditures, or consumption planned ahead (wedding, dowry, funeral, health costs, education expenses, farming inputs, etc.).

Theory of expected effects

Because vulnerable households are often reluctant to invest excess income productively, they often accumulate funds that can be accessed in the event of a shock. Savings are an important tool for low-income households to mitigate risk. With money set aside, households protect themselves against shocks and stabilise cash flow. Nevertheless, these precautionary savings yield only limited returns compared to savings that are invested on productive physical capital. With insurance, precautionary savings are expected to decrease, whilst funds are expected to be preserved by insurance if a shock happens.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- Other

All types of insurances listed are applicable in this case as savings can be intended to decrease a household's risk of suffering a financial shock in many dimensions (health, agriculture, property, etc.).

Measurement

- How much of weekly/monthly household income do you not consume or invest but put aside and save for future use? Provide roster with type of saving (private, institutional, savings group, etc.) and amount and purpose of savings and aggregate amounts.
- What is the total stock of savings you currently have? (National Bureau of Statistics, Nigeria 2010).

Analysis

Due to the expected change in risk taking behaviour, it may be interesting to analyse heterogeneous effects by risk aversion. Furthermore, different purposes and intentions behind the saving behaviour should be taken into consideration in the analysis as well as the form of saving (informal, institutional, savings group, etc.).

Combine with

- Amount of liquid assets
- Number of income sources per household
- Number of memberships in Rotating Savings and Credit Associations (ROSCAs), Accumulating Savings and Credit Associations (ASCAs) and other informal savings networks per household

Outcome: Risk management strategies (ex-ante)

Indicator: Amount of liquid assets

Definition of the indicator

Liquid assets are either cash or assets that can be converted into cash in a very short time period and with a minimum loss of value. They can also be considered as a form of saving, which is rapidly disposable. For a facilitated conversion into cash an important prerequisite is the relative ease in transfer between different ownerships.

Theory of expected effects

Liquid assets are an important tool of low-income households' risk-mitigation – to protect themselves against shocks and stabilise cash flow. With insurance, liquid assets intended for risk mitigation are expected to decrease, whilst funds are expected to be preserved by insurance if a shock happens.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

All forms of formal risk coverage via insurance can be applicable in this case if liquid assets are seen as an ex-ante risk management strategy. As they are by definition easily and quickly transferable into cash, they can compensate shocks that occurred due to all sorts of risks such as health, life, agriculture, or property.

Measurement

- Provide a roster with types of liquid assets, such as cash, bonds, deposits, gold, silver, minerals, etc., and consider culturally specific liquid assets, the current value of the [ITEM] the household owns, value of [ITEM] a year ago.
- How much has the household received from [ITEM] in the past 12 months (interest, dividends, profit, payments, etc.)? (National Bureau of Statistics, Nigeria 2010).

Analysis

One effect could be that insured individuals hold fewer liquid assets and use them for productive investment, savings, or other, more future related financial activities. In this context, it could be interesting to analyse heterogeneous effects by risk aversion, as the

risk averse might still hold on to informal ex-ante risk management strategies. Furthermore, it could be interesting to analyse heterogeneous effects with regards to the purpose of liquid assets, i.e., the intended use within a household for insured and uninsured (ex-ante risk management, funds of out-of-pocket expenses, savings, etc.).

Combine with

- Total amount of savings
- Number of income sources per household
- Number of memberships in ROSCAs, ASCAs, and other informal savings networks per household

Outcome: Risk management strategies (ex-ante)

Indicator: Number of income sources per household

Definition of the indicator

This indicator captures all sources of income in the household. These sources can be of both formal and informal nature. Potential sources could be jobs with regular or irregular wages, income from farming and livestock or from asset ownership (renting, borrowing, etc.).

Theory of expected effects

A common ex-ante risk management strategy of households is to diversify the risk of income shocks by increasing the number of income sources of the household. Thus, deficiency of one income source does not have catastrophic consequences for household income and the income shock remains rather small. Insurance should make income diversification less necessary.

Applicable type of risk coverage

- **Health insurance**
- Life insurance
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- Property insurance
- Other

All forms of formal risk coverage via insurance can be applicable in this case if a high number of income sources is seen as ex-ante risk management strategy. The highest applicability is given for those insurance types which cover risks that occur on a frequent basis with high probability (health insurance, crop and livestock insurance) rather than as a onetime future event (life insurance, funeral insurance, etc.).

Measurement

- Please list all sources of income of this household [provide roster by household members and sources of income] (National Statics Directorate Caicoli, Dili, Timor Leste and World Bank 2001).

Analysis

Under the assumption that informal ex-ante risk management strategies become crowded out by formal insurance, it might be that insured individuals reduce the number of income sources and focus on the most efficient ones with the highest revenue. Furthermore it can be expected that individuals are more willing to focus on rather risky and unstable jobs,

promising higher earnings rather than diversifying the risk of income loss by numerous jobs. Heterogeneous effects regarding the type of jobs could, thus, be an interesting issue for analysis. The number of income sources may, however, also depend on the characteristics and capabilities of the local labour market and the type of jobs available. In this regard, households displaying a particularly high or low degree of income source

diversification may also do this as a reaction to the structure of the job market and not due to their risk taking behaviour.

Combine with

- Total amount of savings
- Total amount of liquid assets
- Number of memberships in ROSCAs, ASCAs and other informal savings networks per household

Outcome: Risk management strategies (ex-ante)

Indicator: Number of membership in ROSCAs, ASCAs and other informal savings networks per household

Definition of the indicator

The indicator captures the number of memberships in different forms of savings networks per household. An informal savings group is a social organisation formed to help community members save money for specific purposes (either individual or community-level). The two most common examples are Rotating Savings and Credit Associations (ROSCAs) or Accumulated Savings and Credit Associations (ASCAs). ROSCAs function by taking monthly deposits from each member of a group and then giving the whole monthly sum to one member of the group. The recipient of the monthly sum is based on a predetermined rotation, ensuring each participant will eventually receive a large payout. ASCAs also require group members to make regular contributions. Instead of rotating payouts, the ASCA group fund is used to make loans that are paid back with interest. Loans are made either to group members or trusted third parties. After a certain period of time, the group fund and its interest are paid back to the original members (Anderson and Baland 2002).

Theory of expected effects

Informal savings networks are an important tool of low-income households' risk-mitigation – to protect themselves against shocks and to stabilise cash flow. With insurance, membership in these informal networks becomes less important in the context of risk management. Nevertheless, if membership is based on other intentions than risk mitigation, the number of memberships will rather remain stable.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

All forms of formal risk coverage via insurance can be applicable in this case if (numerous) membership(s) in informal savings association is seen as ex-ante risk management strategy.

Measurement

- Is anyone in your household member of an informal savings group? If yes, who and in which savings groups? Consider providing a roster with locally prevalent options.

Analysis

It could be interesting to analyse a potential change of purpose of memberships in informal savings networks under insurance as the savings left with the group may not be a part of

an ex-ante risk management strategy, anymore but rather a form of investment and financial diversification. In this context, it could be also interesting to control for heterogeneous effects regarding the amount of money put in the savings group.

Combine with

- Total amount of savings
- Total amount of liquid assets
- Number of income sources per household

Outcome: Risk management strategies (ex-post)/coping strategies in case of shocks

Indicator: Amount of formal loans taken in case of shock events

Definition of the indicator

This indicator captures the amount of formal loans taken up to mitigate a household shock. It, thus, refers to actions households take in order to cope with shocks after they have actually materialised. The indicator can refer to the monetary amount of the loans taken up in the context of a shock as well as to the quantity of loans.

Theory of expected effects

Alternative ex-post responses to shocks, like borrowing, can drain households of existing resources and place demands beyond the cash flow and savings capacity. Coping strategies involving borrowing and, thus, often exacerbate the pressures of debt. As a direct effect it is expected that, in the context of ex-post risk mitigation, fewer loans are taken up to cushion the shock under insurance. This assumption is due to the fact that loans are no longer needed (at least to the same extent) as a risk mitigation mechanism if the income shock can be cushioned by insurance. Thus, the intent of microinsurance here is to turn reactive ex-post risk management

practices into a proactive strategy of risk mitigation.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

All types of risk coverage listed can be applicable here as the indicator refers to ex-post risk management based on various risks. The take-up of loans after a shock occurred is a particularly relevant indicator in situations requiring relatively quick and high coverage of involved costs (health-care services, funeral costs, recovery of property, buying new agricultural inputs for the next cropping season, etc.).

Measurement

- Remember the last shocks that occurred to your household. What kind of loans did you take up after the shock happened? Provide

a roster with shock event, source of loan, amount of loan.

Analysis

The analysis should focus on households that experienced a shock. Ex-post borrowing in case of a shock is also referred to in the literature as “borrowing under stress” (Radermacher et al. 2012). In contrast to ex-ante borrowing, the purpose of ex-post borrowing is quite straight forward—the coverage of incurred costs. It is expected that the take-up of microinsurance leads to a decrease of the total amount of loans in both amount and quantity. As the decision to take up a loan and under which conditions can vary with regard

to the particularities of the shock event and its severity, the analysis should take account of the particular contexts in which the loans are taken up.

Combine with

- Amount of savings used in case of shock
- Amount of informal loans used in case of shock event
- Difference of total household expenditures before and after shock events (without paying back of loans and interest)
- Food intake (self-reported quality and quantity)
- Percentage of children taken out of school due to shock event

Outcome: Risk management strategies (ex-post)/coping strategies in case of shocks

Indicator: Amount of savings used in case of shock

Definition of the indicator

The amount of savings used in case of shocks refers to the part of total savings used to cope with the incurred expenses of the shock after it materialised.

Theory of expected effects

Alternative ex-post responses to shocks, like the use of savings, can drain households of existing resources and place demands beyond the cash flow and savings capacity. Savings are an important tool of low-income households' risk mitigation – to protect themselves against shocks and to stabilise cash flow. With insurance, savings are expected to be preserved if a shock happens. Thus, savings can be stabilised and used for more productive purposes.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Numerous types of insurance are applicable as savings and can be intended to decrease the severity of a financial shock in many dimensions (health, agriculture, property, etc.).

Measurement

- How much of your savings did you use to fund the shock?
- What is the amount of your current savings stocks? (National Bureau of Statistics, Nigeria 2010).

Analysis

Due to the expected change in risk taking behaviour, it might be interesting to analyse heterogeneous effects by risk aversion. Furthermore, different purposes and intentions behind the saving behaviour should be taken into consideration in the analysis as well as the form of savings (informal, institutional, savings group, etc.).

Combine with

- Amount of (formal) loans taken in case of shock
- Difference of total household expenditures before and after shock events (without paying back of loans and interest)

- Food intake (quality and quantity, self-reported)
- Percentage of children taken out of school due to shock event

Outcome: Risk management strategies (ex-post)/coping strategies in case of shocks

Indicator: Difference of total household expenditures before and after shock events (without paying back of loans and interest)

Definition of the Indicator

The indicator captures changes in total household expenditures due to a shock event. The difference between household expenditures before and after a shock event also sheds light on the intensity of the incident. Expenses for paying back loans and interest are not included.

- Crop insurance (indemnity or index)
- Property insurance
- Other

This indicator is applicable to all insurable risks and types of risk coverage as a change in expenditure can be a reaction to all sorts of shocks in order to compensate for the costs incurred.

Theory of expected effects

Consumption-smoothing is a typical household coping strategy in a household shock event. With insurance, expenditures are expected to remain stable if a shock happens, as there is no longer a need to reduce them as part of a risk management strategy. Absorbing the major costs of the shock events, micro-insurance avoids a sharp decrease of household expenditures, which could lead to a descent into (deeper) poverty.

Measurement

To measure this indicator, it would be ideal to compare data collected before, as well as after, the occurrence of the shock event. As this is rather an ideal setting, it may be more feasible to rely on precise memorisation of household expenditure before and after the shock by providing a roster with relevant options for expenditures in order to reach the highest level of accuracy possible.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)

Analysis

As changes in household expenditures are expected to vary according to the severity of the shock event and the amount of incurred costs, it could be advisable to control for and cluster

different types of shocks. Furthermore, it could be interesting to take a more detailed look into subcategories of total expenditures in order to retrieve relevant information about which types of expenditures/consumption change in particular.

Combine with

- Amount of formal loans taken in case of shock
- Amount of savings used in case of shocks
- Food intake (quality and quantity, self-reported)
- Percentage of children taken out of school due to shock event

Outcome: Risk management strategies (ex-post)/coping strategies in case of shocks

Indicator: Food intake (quality and quantity, self-reported)

Definition of the indicator

The indicator captures the households' self-reported food intake in both quality and quantity.

Theory of expected effects

As consumption-smoothing is a common mechanism of household risk mitigation in cases of income or health shocks, food security is a closely connected issue. As a consequence of a household shock, individuals have to cope with the expenses related to the respective event *and* continue to meet ongoing household needs of which expenditure for food ranks on the top. This double burden can easily lead to the decision to eat fewer meals or eat less nutritious food. This effect is particularly relevant for poor households that do not make use of any other coping mechanisms, and, thus, would need to reduce their consumption and their food expenditures in a shock event. As microinsurance aims at stabilising household income in a shock event, food intake in quantity and quality is expected to stay stable ex-post shock.

Applicable type of risk coverage

- **Health insurance**
- Life insurance
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- Property insurance
- Other

Health insurance can be an applicable type of risk management here, particularly if the insurance product encompasses additional interventions related to nutrition. Furthermore, in the case of self-subsistence of the household, agricultural insurance schemes such as crop or livestock insurance could be applicable. Agricultural schemes mitigate the risk of crop failure and/or livestock death, both which could have a negative impact on nutrition.

Measurement

Because self-reported and aggregated information of quantity and quality of food can be inaccurate, it is common practice to measure food consumption in the form of an extensive roster,

covering each food item separately by amount/quantity and price.

Example questions:

- I want to ask about all food consumed by your household, regardless of which person ate it. Has your household consumed [FOOD] during the past 7 days? Please exclude from your answer any [FOOD] purchased for processing or resale (National Statics Directorate Cai-coli, Dili, Timor Leste and World Bank 2001).
- In the past year, was there any month when your household food needs were not met?

Analysis

As consumption-smoothing, particularly smoothing of food expenses, is a mechanism usually only applied when there are no other alternatives, it mostly affects poor households. Thus, it could be interesting to analyse across different levels of poverty. It could be also interesting to analyse heterogeneous effects for different types of household

members, assuming that, in a shock event, a poor household would smooth food expenditures differently for its members, depending on their particular needs (especially for those household members with increased health risks such as pregnant women, children, elderly people, and sick individuals). Additionally, the analysis should differentiate between food quality and quantity, as these subindicators can lead to different assumptions.

Combine with

- Other indicators of risk management strategies (ex-post)/coping strategies in case of shock
- Total expenditures on food per person per week
- Frequency of eating vegetables or fruit (number per week)
- Frequency of eating meat (if people eat meat at all) (number per week)
- Average number of meals eaten per day in last month
- Number of days when food was insufficient for the household in last month

Outcome: Risk management strategies (ex-post)/coping strategies in case of shocks

Indicator: Percentage of children taken out of school due to shock event

Definition of the indicator

The indicator measures the percentage of children who have been taken out of school by their family in order to cope with the consequences and expenses of the shock event.

Theory of expected effects

In the case of a lack of alternatives to cope with a shock event, households may be forced to not only use their financial but also physical assets to cope with the corresponding consequences. Taking children out of school can be based on two intentions. One is to save money, if school fees and other expenses are needed to keep the children in school; the other intention is to take children out of school in order to put them to work in order to cope with the consequences of the shock. Seen from a long term perspective this coping mechanism is inefficient, as it impedes the educational and skill development of the children, which are essential for the future socioeconomic situation of a household. Under insurance, it is expected that

children remain in school after a shock event as educational expenses can still be covered and no additional manpower and support is needed in the household.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Numerous types of insurance are applicable in this case as taking children out of school can be a reaction to all sorts of shock events, placing the household in a difficult economic situation.

Measurement

The indicator can be measured both on the household level by survey questions as well as on the community/school level by using official statistics of school dropout rates after shock events took place.

Analysis

Taking children out of school is a typical coping mechanism used by poor to very poor households which lack alternative coping mechanisms. This indicator implies not only financial but also physical coping in the form of putting the children to work. Furthermore, it could be interesting to control for the age and sex of the children taken out of school. Taking children out of primary school, for example, can be more devastating for their educational

development than at a later stage of school. For some countries, studies show that girls are more likely than boys to be taken out of school to support the household, as the girls' educational prospects are less valued.

Combine with

- Other indicators of risk management strategies (ex-post)/coping strategies in case of shock
- Child labour measures

Outcome: Sale of assets for managing expenses related to shock event

Indicator: Total value of sold assets in case of shocks

Definition of the indicator

The indicator captures the total value of all assets sold in the event of shocks. All sorts of assets are relevant for this indicator.

Theory of expected effects

In order to cope with the corresponding effects and consequence of a household shock, selling assets is a common strategy. These assets can be both liquid and illiquid assets. Particularly high importance in this context can be attributed to productive assets (such as machinery, livestock, etc.) as the sale of these kind of assets has direct negative consequences for the households income opportunities. As the poorer have fewer (and only essential) assets that can be sold in order to cope with a shock, they suffer from a particularly high degree of vulnerability. Under insurance it is expected that the total value of sold assets is lower, as less or no assets need to be sold to cover the expenses incurred.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**

- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

Numerous types of insurances are applicable in this case as the sale of assets can be a reaction to all sorts of shock events, placing the household in a difficult economic situation.

Measurement

- Which assets did you sell after the last shocks? (Provide a roster with column for different shock events, assets sold and their values)

Analysis

For the analysis of this indicator, it can be interesting to compare the actual value of the assets and the received price during sale in the context of the shock, as this is expected to be much lower than the normal market price. Furthermore, it is advisable to distinguish between rather liquid assets and productive assets, as the sale of the latter can have extensive consequences for the household regarding its productive and, thus, future income potential.

Combine with

- Percentage of assets recovered/replaced after being sold in case of shock six months after shock event
- Amount of liquid assets
- Level of assets—housing conditions

Outcome: Sale of assets for managing expenses related to shock event

Indicator: Percentage of assets recovered/replaced after being sold in case of shock six months after shock event

Definition of the indicator

This indicator refers to the part of sold assets after the shock event that are recovered or replaced six months after the incident i.e., some land/livestock may be repurchased after insurance pay out, houses may be repaired.

Theory of expected effects

In order to cope with the corresponding effects and consequences of a household shock, selling assets is a common strategy. These assets can be both liquid and illiquid assets. Particularly high importance in this context can be attributed to productive assets (such as machinery, livestock, etc.) as the sale of these kind of assets has direct negative consequences for a household's income opportunities. As poorer households have fewer assets (or only essentials) that can be sold in order to cope with a shock, they suffer from a particularly high degree of vulnerability. The ability to recover or replace these (productive) assets within a certain time frame after a shock event occurred is thus an interesting indicator of the household's ability to recover and its degree of vulnerability.

Assuming that fewer assets have to be sold in a shock event to cover incurred costs under insurance, fewer assets might need to be recovered as, selling them in the first place had been prevented. If assets had to be sold despite insurance, it is expected that a higher percentage of those assets can be recovered or replaced as insurance decreases the degree of vulnerability and supports a household's ability to recover from the shock within a short time period. This effect could be even stronger if fewer productive assets need to be sold, so a household's production and income potential remains stable.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Numerous types of insurances are applicable in this case as the sale of assets and their recovery can be a reaction to all sorts of shock events,

placing the household in a difficult economic situation.

Measurement

- Of the assets you sold after the shock, how many have you been able to replace/recover within six-months of the shock event?

Analysis

For the analysis of this indicator, it can be interesting to compare the actual value of the assets and the received price during sale in the context of the

shock, as this is expected to be much lower than the normal market price. Furthermore, it is advisable to distinguish between the replacement (or recovery) of rather liquid assets and productive assets, as the recovery of the latter can have extensive consequences for a household regarding its productive and, thus, future income potential.

Combine with

- Total value of sold assets on case of shock
- Level of assets—housing conditions

Outcome: Reliance on informal risk sharing networks

Indicator: Total amount of borrowing for shock-related expenditures from informal networks

Definition of the indicator

The indicator captures the total amount of loans that are taken from informal networks and directly related to expenditures of the shock event. Thus, by definition, it only refers to borrowing conducted after the shock event and not before the respective incident. Informal networks can be self-help groups, saving networks such as ROSCAs or ASCAs, or family, friends, and other acquaintances. In many cases, informal risk sharing networks entail only partial risk protection as default of group members can occur if they are not able to repay into the risk sharing pool (Besley and Coate 1995).

Theory of expected effects

Borrowing from informal networks in a shock event might create social obligations and expectations. It may, in fact, result in costs for a household depending on the specific characteristics of an informal network. Informal networks composed of family and/or close friends are often rather altruistic and do not contain any strings attached. However, risk management via informal networks may be more

unreliable than formal mechanisms, as they depend on the liquidity and willingness to pay of the other network partners, who could suffer from a similar financial shock. Under insurance, a direct effect expected is that, in the context of ex-post risk mitigation, less borrowing from informal networks is conducted to cushion the shock (Dercon et al. 2012). However, not only a demand side effect can be expected but also a decrease in the supply of informal support for risk mitigation as with insurance in place, other individuals might be generally less willing to help (Hintz 2010).

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

All types of risk coverage listed are applicable in this case as borrowing from informal networks can be a reaction to all sorts of shock events, placing the household in a difficult economic situation.

Measurement

- How much did you borrow in total from informal networks (relatives, friends, local groups) in order to cover expenses related to the shock event? Provide a roster covering shock events, source of loan, amount of loan, and additional costs involved (interest or other obligations and expectations).
- The measurement of this indicator could also be conducted in a roster covering all forms of coping mechanisms used in informal risk management networks.

Analysis

It could be interesting to control for different informal sources of borrowing (could be more than one) and, if possible, the particular risk of default of the underlying networks. Furthermore, controlling for formal lending could also be of interest to find out more about the lending practice of a household (informal or formal) and whether the choice is due to independent preferences or lack of access to formal mechanisms.

Combine with

- Loans given to others
- Other indicators for reliance on informal risk sharing networks
- Total amount of loans taken (risk taking behaviour)

Outcome: Reliance on informal risk sharing networks

Indicator: Total amount of borrowing for other expenditures from informal networks

Definition of the indicator

The indicator captures the total amount of loans taken from informal networks that are not directly related to expenditures of the shock event, but are intended for other purposes. Thus, by definition the indicator can refer to borrowing from informal networks at any point in time, ex-ante or ex-post shock. Informal networks can be self-help groups, saving networks such as ROSCAs or ASCAs or family, friends, and other acquaintances. Particularly, informal risk sharing networks entail only partial risk protection as default of group members can occur if they are not able to repay into the risk sharing pool (Besley and Coate 1995).

Theory of expected effects

As mentioned in the description of the indicator, due to borrowing for shock related expenses from informal networks, a shift in purpose of borrowing could be expected for insured households. Borrowing from informal networks for other expenditures is expected to rise relatively to borrowing for shock-related expenses as the latter is crowded out by the insurance coverage. Borrowing from informal

networks is, thus, expected to shift from an ex-post risk management coping to a financial mechanism for other purposes such as consumption, productive investment, etc.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

All types of risk coverage listed are applicable in this case as borrowing from informal networks can be a reaction to all sorts of shock events, placing the household in a difficult economic situation.

Measurement

- How much did you borrow in total from informal networks (relatives, friends, local groups) in order to cover expenses other than those related to the shock event? Provide a roster covering the purpose of loan, source of loan, amount of loan, and additional costs involved

(interest or other obligations and expectations).

- The measurement of this indicator could also be conducted in a roster covering all forms of coping mechanisms used in informal risk management networks.

Analysis

It could be interesting to control for different informal sources (could be more than one) of borrowing and if possible the particular risk of default of the underlying networks. Furthermore, controlling for formal lending

could be also of interest in order to find out more about the lending practice of the household (rather informal or formal) and whether the choice is due to independent preferences or lack of access to formal mechanisms.

Combine with

- Other indicators concerning reliance on informal risk sharing networks, especially total amount of borrowing for shock-related expenditures from informal networks
- Total amount of loans taken (risk taking behaviour)

Outcome: Reliance on informal risk sharing networks

Indicator: Total amount of contributions received from family in case of a shock (as loan)

Definition of the indicator

This indicator encompasses all financial contributions received by family members in a shock event, given as loans to be repaid. The contributions as defined here refer to financial contributions which could be provided by family members living nearby or abroad (remittances).

Theory of expected effects

Contributions received by family in a shock event are common practice in many societies that make use of a particularly high degree of intra-family commitment and solidarity. Borrowing from informal networks in a shock event might create social obligations and expectations. It may, in fact, result in costs for a household depending on the specific characteristics of such informal network. Informal networks composed by family are often rather altruistic and do not contain any strings attached. However, risk management via informal networks may be more unreliable than formal mechanisms as they depend on the liquidity and willingness to pay of the other family members, who could suffer from a similar financial

shock. Under microinsurance, it is expected that fewer contributions by family are received as an insurance substitute intended to cushion the shock.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

All types of risk coverage listed are applicable in this case as contributions by family can be intended to decrease a household's risk of suffering a financial shock in many dimensions (health, agriculture, property, life, etc.).

Measurement

- What is the total amount of contributions you received from family members after the shock event as a loan? Include a roster with shock event, contributing family member, type of contribution (cash, other financial contributions), the total amount per contribution, and

conditions bound to the loan (interest rate or other obligations).

Analysis

The amount of contributions received by family in the case of a shock event, should be distinguished in the analysis from contributions received on a permanent or regular basis (i.e., remittances from family members living abroad, etc.). Furthermore, the

indicator should be analysed together with the related indicator capturing family contributions that are a gift and do not need to be repaid.

Combine with

- Other indicators concerning reliance on informal risk sharing networks, particularly total amount of contributions received by family in case of a shock (as gift)

Outcome: Reliance on informal risk sharing networks

Indicator: Total amount of contributions received by family in case of a shock (as gift)

Definition of the indicator

This indicator encompasses all financial contributions received by family members in a shock event, which are given as a gift and thus do not need to be repaid.

Theory of expected effects

Contributions received by family in a shock event are common practice in many societies that make use of a particularly high degree of intra-family commitment and solidarity. Receiving a contribution as a gift from an informal network in a shock event might create social obligations and expectations. It may result in costs for a household depending on the specific characteristics of the informal network. Informal networks composed of family are often rather altruistic and do not contain any strings attached. However, these contributions may be more unreliable than formal mechanisms as they depend on the liquidity and willingness to pay of the other network partners, who could suffer from a similar financial shock. Under insurance, it is expected that fewer contributions by family are received as an insurance substitute intended to cushion the shock.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

All types of risk coverage listed are applicable in this case as contributions by family can be intended to decrease a household's risk of suffering a financial shock in many dimensions (health, agriculture, property, life, etc.).

Measurement

- What is the total amount of contributions you received by family members after the shock event as a gift? Include a roster with shock event, contributing family member, type of contribution (cash, other financial contributions), the total amount per contribution, and conditions bound to the loan (interest rate or other obligations).

Analysis

The amount of contributions received by family particularly in the case of a shock event, should be distinguished in the analysis from contributions received on a permanent or regular basis (i.e., remittances from family members living abroad, etc.). Furthermore, the indicator should be analysed together with the related

indicator capturing family contributions that are a loan and need to be repaid.

Combine with

- Other indicators concerning reliance on informal risk sharing networks, particularly total amount of contributions received by family in case of a shock (as loan)

Outcome: Reliance on informal risk sharing networks

Indicator: Total amount of contributions received from informal or semiformal organisations, e.g. local church association, employer, etc. in case of shock (as loan)

Definition of the indicator

The indicator captures contributions from informal and semiformal organisations in cases of shock. These contributions could be cash or in-kind and are provided after the occurrence of the shock, most likely in the form of a onetime transfer. The contributions are given as a loan and need to be repaid within a certain time frame.

Theory of expected effects

In the absence of formal social security schemes, contributions from informal and semiformal organisations in the case of a shock are common practice. This is particularly evident in contexts where community ties are strong. Under insurance, it is expected that the amount of these contributions decreases as individuals can cope with the shock incurred expenses autonomously and are less dependent on informal or semiformal support and funds.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Numerous types of insurances are applicable in this case as formal and semiformal contributions can be intended to cover a household's expenses after all kind of shocks affecting household income and funds.

Measurement

- What is the total amount of contributions (loans) you received from informal and semiformal organisations after the shock? (This could be local church, employer, community organisations, etc.)

Analysis

Analysis and measurement should account for the different informal and semiformal sources for contributions and the respective conditions attached to the contributions given as a loan (repayment conditions, other obligations, or expectations attached).

Combine with

- Other indicators concerning reliance on informal risk sharing networks

Outcome: Reliance on informal risk sharing networks

Indicator: Total amount of loans (currently pending) *given to family members and other community members*

Definition of the indicator

This indicator refers to outstanding loans given to family and other community members. The indicator refers to the total monetary amount of the loans as well as to the number of outstanding loans. These loans are, thus, part of the financial assets of a household or individual invested, with or without interest rate attached.

Theory of expected effects

Loans given to family members and other community members are the flip side of the coin of informal risk sharing. The provision of loans to other family or community members is often a deep-rooted characteristic of community and family structures in context, where there is a lack of formal alternatives. Under insurance, the willingness and ability to provide more loans to family and community members might increase because the funds are less needed as precautionary assets for people's own purposes. Or, it could be that insured individuals might be less willing to help those who did not behave cautiously and refused insurance.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

All types of risk coverage listed are applicable in this case as contributions given to family and community members can be intended to cushion a financial shock in many dimensions (health, agriculture, property, life, etc.).

Measurement

- What is the total amount currently pending of loans you provided to family and/or community members? Provide a roster with the beneficiary of the loan, amount of loan, and conditions attached (interest, other obligations, or expectations).

Analysis

It could be interesting to analyse this indicator by different subgroups of recipients. Furthermore, details of the

loan arrangement (repayment arrangements, interest rates, variability regarding income of provider, etc.) are interesting to consider in the analysis.

Combine with

- Other indicators concerning reliance on informal risk sharing networks

Outcome: Variability of costs or profits

Indicator: Variability of costs or profits

Definition of the indicator

The indicator captures the variability of costs or—depending on the insurance type—profits for a household in a certain category. It can, for example, consist of health costs or profits from agricultural activity over time.

Theory of expected effects

The primary expected effect of micro-insurance is to reduce the costs that need to be covered by households in a shock event. At the same time, households have to pay an insurance premium constantly whilst being covered. As a consequence, variability of costs incurred (e.g., for health) should decrease. For insurance covering productivity shocks (such as crop insurance), a similar effect occurs for profits. During good times, whilst profits are high, a household pays an insurance premium. If a shock happens (e.g. a flood) and profits are low, the insurance should pay out. In sum, variability of profits over time is decreased.

Applicable type of risk coverage

- Health insurance
- Life insurance

- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

This indicator refers to all kinds of shocks that can incur costs, thus all types of risk coverage are applicable.

Measurement

This indicator is best measured over time, covering pre- and post-shock periods. It should ask for effective costs incurred (including insurance premiums) or profits made for certain expenditure or profit types related to the risk covered. As an alternative to following clients over time, a survey could also ask for cost or profit histories.

Analysis

Given the theory of expected effects, it is likely that the variability decreases mainly if a shock of considerable size happens. Hence, it may be advisable to focus the analysis on those households with considerable shock events during the time frame covered or to conduct separate analyses for households having experienced different shock exposure.

Combine with

- Total costs in case of shock
- Subindicators involving costs of the shock event

Outcome: Total costs in case of shock

Indicator: Total costs in case of shock

Definition of the indicator

The indicator captures all costs and expenses involved for a household in a shock event. It is, thus, composed of numerous indicators mentioned above and can among others encompass direct out-of-pocket spending, expenses for recovery and reconstruction, costs incurred for taking up loans or borrowing and also opportunity costs if a change in labour supply is necessary.

Theory of expected effects

The primary expected effect of micro-insurance is to reduce the total costs that need to be covered by households in a shock event. For other related indicators, such as out-of-pocket spending (OOPS), taking-up of loans, or costs for reconstruction and recovery, it is expected that a significant amount of these costs will be directly covered by insurance and not strain the economic situation of a household. However, seen from the perspective of insured individuals with potentially increased risk taking behaviour, total costs of the damage after a shock may be even higher as the initial (business) investment was higher.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- Other

This indicator refers to all kinds of shocks that can incur costs, thus all types of risk coverage are applicable.

Measurement

This indicator can be measured by asking for an accumulated estimation of all costs involved in the case of a shock. For the sake of higher accuracy, it is, however, advisable to provide a roster with plausible options/categories of costs and ask the questions separately.

Analysis

As this indicator accumulates all costs involved with a shock event, it provides a good overview of the total economic damage for a household due to the shock. For more detailed analysis, it may be advisable to control for relevant subindicators and subcategories in the analysis.

Combine with

- Variability of costs or profits
- Subindicators involving costs of the shock event

Outcome: Total out-of-pocket spending (OOPS) in case of shock

Indicator: Net OOPS per shock event

Definition of the indicator

This indicator refers to the direct outlay of cash or immediately available liquid assets needed to cover expenses incurred due to a shock event. OOPS can also encompass deductibles paid to access goods and services covered by the insurance. Furthermore, bribes, transport fees, drugs, and medical procedures not covered by the insurance are often covered by OOPS.

Theory of expected effects

Out-of-pocket spending in the case of a shock event is an important risk mitigation tool for low-income households. If it absorbs a considerable amount of household income, OOPS can have critically negative implications for a household's economic situation. Without insurance, costs that are due immediately—a situation particularly relevant in the case of health shocks—are often paid from private cash funds. Also, with insurance, some OOPS is still expected as referred to in the definition of the indicator. Nevertheless, it is expected that OOPS decreases when the costs of the shock event are absorbed by microinsurance and the incidence and depth of monetary outlays diminish. Existing studies on the effect of microinsurance

on OOPS, however, ambiguously depict this expected effect. Whilst Jütting (2004) found a 45-51% decrease in OOPS spending for members of a Senegalese community-based health insurance scheme compared to non-members, Wagstaff et al. (2009) did not depict any statistically significant changes in OOPS for China's New Cooperative Medical Scheme (NCMS).

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Whilst most frequently used in the context of expenses for hospitalisation and other medical treatments, this indicator is relevant for all insurable risks and types of risk coverage that require immediate expenses and/or recovery action.

Measurement

- How much did you spend out-of-pocket for [SHOCK EVENT]?

Analysis

Existing studies on the effect of microinsurance on OOPS show ambiguous results, ranging from a clear reduction in OOPS to no indications of change. This leads to the assumption that the effect of microinsurance on OOPS is strongly bound to the type of insurance policy at stake. Thus, the particularities of the insurance scheme should be taken into account in the analysis as well as other forms of informal and coinsurance held by a household.

Furthermore, the specific use of the OOPS for insured and uninsured should be taken into account.

Combine with

- Risk management strategies (ex-post)
- Net OOPS per full episode of illness
- Net OOPS on varying categories of treatment: hospital stay, deliveries, self-treatment, ambulatory care from formal providers, inpatient care

Outcome: Total out-of-pocket spending (OOPS) in case of shock

Indicator: In case of health: Net OOPS per full episode of illness

Definition of the indicator

This indicator refers to the direct outlay of cash or immediately available liquid assets needed to cover the expenses incurred per full episode of illness. Thus, all OOPS from the first signs of illness until full recovery is captured in this indicator. OOPS can also encompass deductibles paid to access goods and services covered by the insurance. Furthermore, bribes, transport fees, drugs, and medical procedures not covered by the insurance are often covered by OOPS.

Theory of expected effects

Out-of-pocket spending in instances of illness is an important risk mitigation tool for low-income households to cover health expenses. If it absorbs a considerable amount of household income, OOPS can have critically negative implications for a household's economic situation. Without insurance, costs that are due immediately—a situation particularly relevant in the case of health shocks—are often paid from private cash funds. Also, with insurance, some OOPS is still expected as referred to in the definition of the indicator. Nevertheless, it is expected that OOPS decreases when the costs of the shock event are

absorbed by microinsurance and the incidence and depth of monetary outlays diminish. Existing studies on the effect of microinsurance on OOPS, however, ambiguously depict this expected effect. Whilst Jütting (2004) found a 45–51% decrease in OOPS spending for members of a Senegalese community-based health insurance scheme compared to non-members, Wagstaff et al. (2009) did not depict any statistically significant changes in OOPS for China's New Cooperative Medical Scheme. Chankova et al. (2008) found that, whilst inpatient expenses were reduced, out-of-pocket (OOP) expenses for outpatient care were not reduced by the investigated mutuelles and insurance schemes. This finding is attributed by the authors to the coinsurance rates of 25–50% per visit.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

- How much did you pay out-of-pocket for [PARTICULAR EPISODE OF ILLNESS]?

Analysis

Existing studies on the effect of micro-insurance on OOPS in the case of health show ambiguous results ranging from a clear reduction in OOPS to no indications of change. This lead to the assumption that the effect of microinsurance on OOPS is strongly bound to the type of insurance policy at stake. Thus, the particularities of the insurance scheme should be taken into account in the analysis as well as other forms of informal and coinsurance held

by a household. Furthermore, the specific use of the OOPS for insured and uninsured should be taken into account.

Combine with

- Risk management strategies (ex-post)
- Net OOPS per shock event
- Net OOPS on varying categories of treatment: hospital stay, deliveries, self-treatment, ambulatory care from formal providers, inpatient care

Outcome: Total out-of-pocket spending (OOPS) in case of shock

Indicator: Net OOPS on varying categories of treatment: hospital stay, deliveries, self-treatment, ambulatory care from formal providers, inpatient care

Definition of the indicator

This indicator refers to the direct outlay of cash or immediately available liquid assets needed to cover the expenses incurred for varying categories of medical treatment such as hospital stay, deliveries, self-treatment, ambulatory care from formal providers and inpatient care.

Theory of expected effects

Out-of-pocket spending in the case of illness is an important risk mitigation tool for low-income households to cover health expenses. If it absorbs a considerable amount of household income, OOPS can have critically negative implications for a household's economic situation. Without insurance, costs that are due immediately—a situation particularly relevant in the case of health shocks—are often paid from private cash funds. Also, with insurance, some OOPS are still expected as referred to in the definition of the indicator. Nevertheless, it is expected that OOPS decreases when the costs of the shock event are absorbed by microinsurance and the incidence and depth

of monetary outlays diminish. Existing studies about the effect of microinsurance on OOPS, however, ambiguously depict this expected effect. Whilst Jütting (2004) found a 45-51% decrease in OOPS spending for members of a Senegalese community-based health insurance scheme compared to non-members, Wagstaff et al. (2009) did not depict any statistically significant changes in OOPS for China's New Cooperative Medical Scheme. Chankova et al. (2008) found that, whilst inpatient expenses were reduced, OOP expenses for outpatient care were not reduced by the investigated mutuelles and insurance schemes. This finding is attributed by the authors to the coinsurance rates of 25-50% per visit. Thus, the expected effects seems to depend on the particular insurance scheme and the form of treatment, which need to be analysed separately.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

- How much did you pay out-of-pocket for [MEDICAL TREATMENT] during the past month?

Analysis

Existing studies on the effect of micro-insurance on OOPS in the case of health show ambiguous results, ranging from a clear reduction in OOPS to no indications of change. This leads to the assumption that the effect of microinsurance on OOPS is strongly bound to the type of insurance policy

at stake. Thus, the particularities of the insurance scheme should be taken into account in the analysis as well as other forms of informal and coinsurance held by a household. Furthermore, the specific use of the OOPS for insured and uninsured should be taken into account and the different treatments should be analysed separately for each case.

Combine with

- Risk management strategies (ex-post)
- Net OOPS per shock event
- Net OOPS per episode of illness

Outcome: Quality of health-care providers

Indicator: Hospital mortality rate

Definition of the indicator

This indicator for the quality of health-care providers refers to the percentage of patients who die whilst they are in hospital.

Theory of expected effects

Microinsurance could lead to a lowering of hospital mortality rates. On the one hand, this is due to the expectation that insured individuals already benefit from more comprehensive and higher quality medical treatment prior to hospitalisation, improving their general health condition and their individual risk factors. On the other hand, empirical evidence shows significant effects of microinsurance lowering hospital mortality rates by offering products tailored to the situation of high risk patients such as pregnant women and/or children. In these cases, insurers often directly contract with hospitals to ensure better services. In Guinea, Centre International de Développement et de Recherche (CIDR) launched a “safe motherhood” health microinsurance product to cover deliveries at hospitals. The product also included emergency evacuation by ambulance and value-added services such as antenatal visits. After the launch of the

product, the maternal mortality rate was about 4% lower than before the launch (non-randomised evaluation). Evidence of obligatory health insurance for school children shows similarly positive results (Radermacher et al. 2012).

Applicable type of risk coverage

- Health insurance

Measurement

For measurement, data should be retrieved directly from hospitals or statistical agencies. For a more accurate analysis, information about the health status of the individuals who died in hospital at the time of their hospitalisation should be obtained as well.

Analysis

In the analysis, the mortality rate should be clustered by different risk factors, prevalent at the time of hospitalisation. Heterogeneous effects are expected here. Thus, the mortality rates should be calculated by dividing the number of deaths amongst hospital patients with a specific medical condition or procedure by the total number of patients admitted for that

same medical condition or procedure. This risk adjustment method accounts for the impact of individual risk factors, such as age, severity of illness, and other medical problems, that can put some patients at greater risk of death than others.

Combine with

- Scoring on quality assessments

Outcome: Quality of health-care providers

Indicator: Scoring on quality assessments

Definition of the indicator

This indicator measures the quality of health-care providers by their scoring on quality assessments. The quality assessment of health-care providers is generally a difficult endeavor since it can be measured both objectively and subjectively using different indicators. The most critical question here is, who conducts the quality assessment? This could be the health-care provider itself, the government, an insurance provider, a non-governmental organisation, etc.

Theory of expected effects

If the health insurance provider has a direct influence on the quality monitoring of contracted health-care providers, it is expected that insurance holders benefit from health services with higher quality and higher scorings in quality assessments. By setting standards for quality of the insured patients, an equal treatment of insured and uninsured patients should be provided. This could occur if the insured patient is not paying directly out-of-pocket and could, thus, be perceived as less solvent. Under the supervision of a health insurance medical advisor, working within the framework of the contractual collaboration between

insurance and health-care provider, quality is generally expected to improve (LeRoy and Holtz 2012).

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

For measurement, data should be retrieved directly from the organisation/agency that conducted the assessment (preferably independent assessments). Furthermore, it could be useful to retrieve additional data directly from hospitals and other relevant health-care providers, if available, in order to reconstruct the assessment and better understand subcategories. If clinical outcome data is not available it can be also useful to further investigate subjective measures of quality based on patients' experiences and perceptions of quality.

Analysis

As there is no universal standard for health-care quality assessments, the

scoring system at stake should be taken into account in the analysis. It may be advisable to not only use the final scoring, but to take a deeper look at the composition of indicators of the assessments and the respective scoring in subcategories.

Combine with

- Hospital mortality rate

Outcome: Quantity of health-care providers**Indicator: Number of modern health-care providers within a defined area/radius****Definition of the indicator**

This indicator measures the number of modern health-care providers within a defined area or radius. In order to use this indicator efficiently, the criteria for a *modern* health-care provider should be predefined.

Theory of expected effects

As modern (allopathic) medicine is widely believed to lead to better health outcomes than many traditional or alternative approaches, this indicator combines quality and quantity aspects of health care. If the health insurance provider has a direct influence on the quality monitoring of contracted health-care providers, it is expected that modern standards are a prerequisite. Under the supervision of a health insurance medical advisor, working within the framework of the contractual collaboration between insurance and health-care provider, quality is generally expected to improve, and, thus, more modern health-care providers are expected to evolve within a certain radius (LeRoy and Holtz 2012). Nevertheless, as the number of modern health-care providers in

certain areas depends on various other criteria— infrastructural conditions, governmental support, availability of skilled human resources, etc.—no direct effect is expected here, at least in the short- and medium-term or for large-scale interventions. Long-term effects can be expected based on the assumptions mentioned, particularly if the introduction of insurance in large-scale interventions displays a form of organised demand, encouraging providers to extend their services. Existing studies focus on the relationship between distance and enrolment rate. It is found that the presence of health facilities within a small radius increases the likeliness of enrolment (Chankova et al. 2008; Wagstaff et al. 2009). Regarding utilisation, distance is perceived as a clear indicator of less utilisation due to the access barriers (Franco et al. 2008; Schneider and Diop 2001).

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

For measurement, data should be retrieved directly from statistical agencies and agencies conducting quality assessments of health facilities, thus providing evidence of the state of their modern equipment. Furthermore, individual accessibility of modern health-care facilities can be inquired by survey questions. For example, in many remote regions with poor infrastructure, distance is not a sufficient indicator of accessibility, but real-time accessibility should be measured (How long does it take to reach the facility by available means of transport? etc.).

Analysis

The analysis could take into account different categories of “modernity” as well as different radiuses or other measures of accessibility. Furthermore, consideration should be given as to whether or not the insurance scheme being analysed covers medical treatment at these health-care providers, and which is the closest one providing treatment for insured/uninsured individuals.

Combine with

- Indicators concerning quality of health-care providers to define criteria of modernity

Outcome: Receiving (appropriate) health care — health-care utilisation (needs-based)

Indicator: Total number of visits to outpatient services per household member within the last month

Definition of the indicator

This indicator measures the number of visits household members make to outpatient services. It refers to health-care services conducted ambulant, in a walk-in manner, and does not include hospitalisation of a patient.

Theory of expected effects

There are robust empirical findings that microinsurance generally increases the use of health-care services.

It is expected that existing barriers to healthcare utilisation, which are particularly its costs and accessibility, will be reduced by microinsurance, leading to a higher utilisation rate. The impact of microinsurance on health-care utilisation rates is one of the most well-researched topics in microinsurance so far. Most studies published found positive or mixed results, in line with the theoretical expectations. For example, Msuya et al. (2004) found that members of the United Republic of Tanzania's Community Health Fund used formal health service with an increased likelihood of 15%. Likewise, Polonsky et al. (2009) found that

members of Oxfam-operated insurance schemes in Armenia had a significantly higher frequency of utilisation at 3.5% compared to non-members.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

- How many times did [NAME] use health-care services within the last month without staying overnight? (ICF 2011)

Analysis

Even though most existing studies show a positive impact of microinsurance on utilisation rates, expected effects depend on the particularities of the insurance policy and its facilitated benefits. Outpatient and inpatient services used should be clearly differentiated in the analysis as well as the suitability of the treatment chosen for the underlying health problem.

Combine with

- Other indicators of health-care utilisation

Outcome: Receiving (appropriate) health care — health-care utilisation (needs-based)

Indicator: Total number of visits to inpatient services (with at least 24 hours hospitalisation) per household member within the last month/year

Definition of the indicator

This indicator measures the number of visits household members make for inpatient services. It refers to health-care services that include inpatient treatment of the patient for at least 24 hours.

Theory of expected effects

There are robust empirical findings that microinsurance generally increases the use of health-care services.

It is expected that existing barriers to healthcare utilisation, which are particularly its costs and accessibility, will be reduced by microinsurance, leading to a higher utilisation rate. The impact of microinsurance on health-care utilisation rates is one of the most well-researched topics in microinsurance so far. Most studies published found positive or mixed results in line with the theoretical expectations. For example, Msuya et al. (2004) found that members of the United Republic of Tanzania's Community Health Fund used formal health service with an increased likelihood of 15%. Likewise, Polonsky et al.

(2009) found that members of Oxfam-operated insurance schemes in Armenia had a significantly higher frequency of utilisation at 3.5 per cent compared to non-members.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

- How many times did [NAME] make use of health-care services, staying overnight within the last month/year?
- How many nights did [NAME] spend at the health-care facility? (ICF 2011)

Analysis

Even though most existing studies show a positive impact of microinsurance on utilisation rates, expected effects depend on the particularities of the insurance policy and its facilitated benefits. Outpatient and inpatient

services utilised should be clearly differentiated in the analysis as well as the suitability of the treatment chosen for the underlying health problem.

Combine with

- Other indicators of healthcare utilisation

Outcome: Receiving (appropriate) health care — health-care utilisation (needs-based)

Indicator: Total number of illness episodes involuntarily self-treated per household member within the last month

Definition of the indicator

This indicator measures the number of illness episodes which were self-treated involuntarily. Thus, the indicator captures cases in which the use of health care would have been the preferred treatment, but was not feasible due to certain barriers.

Theory of expected effects

Self-treatment can cause medical complications such as progression of an untreated or misdiagnosed illness, complications from self-prescribed drugs, or public health problems in the case of infectious disease (Derriennic et al. 2005). There are robust empirical findings, that microinsurance generally increases the use of health-care services.

It is expected that existing barriers to healthcare utilisation, which are particularly its costs and accessibility will be reduced by microinsurance, leading to a higher utilisation rate. If no sufficient funds and/or transport options are available, involuntary self-treatment

can be the consequence, even if the person is aware of the need for formal professional health care. The impact of microinsurance on health-care utilisation rates is one of the best researched topics in microinsurance so far. Most studies published found positive or mixed results in line with the theoretical expectations. For example, Msuya et al. (2004) found that members of the United Republic of Tanzania's Community Health Fund used formal health service with an increased likelihood of 15%. Likewise, Polonsky et al. (2009) found that members of Oxfam-operated insurance schemes in Armenia had a significantly higher frequency of utilisation at 3.5% compared to non-members. Wang et al. (2009) found that membership in China's Rural Mutual Health Care reduced self-treatment by about two-thirds.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

- Within the last month, how many times has [NAME] been ill and unable to see a doctor even though he (or she) wanted to? (ICF 2011)

Analysis

Even though most existing studies show a positive impact of microinsurance on utilisation rates, expected effects depend on the particularities of

the insurance policy and its facilitated benefits. Outpatient and inpatient services, as well as self-treatment used, should be clearly differentiated in the analysis as well as the suitability of the treatment chosen for the underlying health problem.

Combine with

- Other indicators of health-care utilisation

Outcome: Receiving (appropriate) health care — health-care utilisation (needs-based)

Indicator: Regarding children: percentage of children (below age five) seeking diarrhea treatment

Definition of the indicator

This indicator measures the percentage of children below the age of five seeking diarrhea treatment. This indicator, thus, focuses on a particular health-care treatment that is expected to have a positive influence on the health status of the beneficiary.

Theory of expected effects

As health status is difficult to measure objectively, proxies are often used to approach this issue. In medical research there is a widespread opinion that certain treatments and health-promoting behaviour of children have an important long-term effect on their health status, which can, in some cases, even influence their adult health conditions. Diarrhea, in most cases, is an illness that requires only simple outpatient treatment, but it is still one of the highest ranking causes of mortality in children under five. There are robust empirical findings that microinsurance generally increases the use of health-care services. It is expected that existing barriers to healthcare utilisation, which are particularly its costs and accessibility, will be reduced by microinsurance,

leading to a higher utilisation rate. Thus, it is also expected that the rate of treatment of children with diarrhea will increase, especially if the insurance is tied to educational interventions or regular health checkups. Educational interventions, promoting hygienic habits, preventive measures, and Oral rehydration therapy (ORS) can also lead to high rates of self-treatment or efficient direct prevention of diarrhea, leading to a decrease in the usage rate of health-care services.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

- Did [NAME OF CHILD] have treatment the last time they suffered from diarrhea?
- Did you seek advice for the diarrhea treatment or did you treat it at home?
- What kind of treatment was provided? (Provide roster with options,

e.g., ORS liquid, specific diet, drugs) (ICF 2011)

Analysis

As diarrhea treatment is only one of several health-care interventions believed to significantly decrease the mortality rate for children under-five and improve their general health status, other interventions should be considered in the analysis. Furthermore, frequency and quality of

the treatment should also be taken into account as well as the degree of severity of the diarrhea episode. It could also be interesting to control for participation in educational units for child health and diarrhea treatment in particular.

Combine with

- Other indicators of health-care utilisation, particularly regarding children

Outcome: Receiving (appropriate) health care — health-care utilisation (needs-based)

Indicator: Regarding children: percentage of children (below age five) sleeping under a mosquito net

Definition of the indicator

This indicator measures the percentage of children below the age of five sleeping under a mosquito net, a common precautionary measure to prevent infection with malaria.

Theory of expected effects

As health status is difficult to measure objectively, proxies are often used to approach this issue. In medical research there is the widespread opinion that certain treatments and health-promoting behaviour of children have an important long-term effect on their health status, which can, in some cases even influence their adult health conditions. Malaria ranks amongst the most prevalent reasons for mortality in children under five. Several studies show that sleeping under a mosquito net reduces the risk of infection with malaria drastically. There are robust empirical findings that microinsurance generally increases the use of health-care services. For instance, Franco et al. (2008) found that members of four Equity Initiative policies in Mali increased the use of mosquito nets

for children and pregnant women. The effects expected are also dependent on the kind of distribution of mosquito nets. Empirical evidence found that mosquito nets provided as a gift have a rather low utilisation rate, as they are valued less. According to this line of argument, insurance could lead to a lower utilisation rate than expected if the provision of mosquito nets is part of the insurance scheme.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

- Does your household have any mosquito nets that can be used whilst sleeping? Who slept under this mosquito net last night? (ICF 2011)

Note: The surveyor should observe whether mosquito nets exist in the household and if they seem to be in use.

Analysis

As sleeping under a mosquito net is only one of several health-care interventions believed to significantly decrease the mortality rate for children under-five and improve their general health status, other interventions should be considered in the analysis. Furthermore, frequency and quality of the treatment should also be taken

into account (is the net always used, is it in good order etc.). It could also be interesting to control specifically for participation in educational sessions on child health and malaria prevention.

Combine with

- Other indicators of health-care utilisation, particularly regarding children

Outcome: Receiving (appropriate) health care — health-care utilisation (needs-based)

Indicator: Regarding children: percentage of children (below age five) getting vitamin A supplements

Definition of the indicator

This indicator measures the percentage of children below the age of five getting vitamin A supplements. This indicator, thus, focuses on a particular health-care treatment health-care that is expected to have a positive influence on the health status of the beneficiary.

Theory of expected effects

As health status is difficult to measure objectively, proxies are often used to approach this issue. In medical research there is the widespread opinion that certain treatments and health promoting behaviour of children have an important long-term effect on their health status, which can, in some cases even influence their adult health conditions. Randomised medical trials show that vitamin A supplementation is associated with large reductions in mortality, morbidity, and vision problems in a range of setting. Thus, vitamin A supplementation is strongly recommended for children between six-months and five years of age. It is expected that, under microinsurance, pre-existing barriers to this treatment will be removed and access facilitated.

Applicable type of risk coverage

• Health insurance

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

- Within the last six months, was [NAME] given a vitamin A dose like (this/any of these)? Show common types of ampoules/capsules/syrups (ICF 2011).

Analysis

As vitamin A supplementation is only one of several health-care interventions believed to decrease the under-five-mortality rate for children significantly and to improve their general health status, other interventions should be considered in the analysis. Furthermore, frequency and quality of the treatment should also be taken into account. It could also be interesting to control for participation in educational sessions on child health and diet in particular.

Combine with

- Other indicators of health-care utilisation, particularly regarding children

Outcome: Receiving (appropriate) health care — health-care utilisation (needs-based)

Indicator: Regarding children: number of immunisations for children below age one (per child)

Definition of the indicator

This indicator measures the percentage number of immunisations of children from birth to 12 months.

Theory of expected effects

This indicator focuses on a particular treatment of healthcare utilisation that is expected to have a positive influence on the health status of the beneficiary child. As health status is difficult to measure objectively, proxies are often used to approach this issue. In medical research, there is the widespread opinion that certain treatments and health-promoting behaviour of children have an important long-term effect on their health status, which can, in some cases, even influence their adult health conditions. Immunisation is of particular importance during the first months after birth. For many preventable diseases, the first shot of immunisation is recommended by the World Health Organization (WHO) between month 0 and 12. There are robust empirical findings that microinsurance generally increases the utilisation of health-care services. It is expected that existing barriers to healthcare utilisation, which

are particularly its costs and accessibility, will be reduced by microinsurance, leading to a higher utilisation rate. Thus, it is also expected that the number of immunisations for children will increase, particularly if the insurance is bound to educational interventions or regular health checkups.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

Provide a roster with names of children below age one and the most prominent and important immunisations, as recommended by the WHO and adapted to the local context, as immunisation practices may vary from country to country.

Analysis

As immunisation is only one of several health-care interventions believed to decrease the under-five mortality rate

for children significantly and improve their general health status, other interventions should be considered in the analysis. Furthermore, frequency and quality of the treatment should also be taken into account as well as the type of immunisation and whether they are followed up if further shots are needed within a certain time period. In general,

the effect should strongly depend on whether immunisations are paid by the insurance.

Combine with

- Other indicators of health-care utilisation, particularly regarding children

Outcome: Receiving (appropriate) health care — delay in health care seeking

Indicator: Number of days symptoms persisted before treatment was sought

Definition of the indicator

This indicator refers to delay in health care seeking, measured by the number of days the symptoms of the illness persisted before treatment was sought.

Theory of expected effects

There are empirical findings that delays in obtaining health care can lead to increased morbidity and mortality in many cases (Derriennic et al. 2005; Msuya et al. 2004). The number of days that symptoms of an illness persisted before treatment was sought is a convenient indicator of whether appropriate health care is easily accessible. It is expected that existing barriers to health-care utilisation, which are particularly its costs and accessibility will be reduced by microinsurance, leading to a higher utilisation rate. Thus, it is also expected that the number of days symptoms persist before professional consultation decrease, as barriers are lower.

Applicable type of risk coverage

- Health insurance

As this indicator explicitly refers to the case of health, the applicable type of risk coverage is health insurance.

Measurement

- How many days did the illness persist before [NAME] sought treatment?
- What was the reason why treatment was not sought immediately?

Analysis

As delay in health care seeking can also have other reasons than typical barriers, such as cost and infrastructure, it is advisable to find out more about the specific reasons for a delay. Furthermore, whilst a delay can have dramatic consequences for some illnesses, for others it may not be that urgent. Thus, this indicator should be seen in its specific context.

Combine with

- Other indicators concerning health-care utilisation

Outcome: Equity regarding health and health care

Indicator: Use any of the indicators mentioned for subgroups

Definition of the indicator

This indicator refers to equity in the context of health and health care by focusing on subgroups for any indicator specific to health microinsurance. Subgroups can be defined based on numerous categories: socioeconomic, spatial, ethnic or religious, gender based or referring to different household members. In general, equity in health care is established if people who do not have access to and utilise health-care services receive these services at the same level of those who have access already.

Theory of expected effects

The in-depth analysis of subgroups is a suitable indicator to examine the degree of equity in access to and use of health-care services. Equity in access to health care is one hypothesised impact of microinsurance based on the assumption that access to insurance can be provided to members of excluded groups. By focusing on subgroups, detailed analysis can be conducted, addressing the questions of who has access to insurance and who actually receives the benefits. In the ideal case, access to insurance should be egalitarian across subgroups.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

For measurement, the same questions/methods mentioned for the relevant indicators should be applied to the specific subgroups.

Analysis

The analysis should be conducted in the way advised for the respective indicators, focusing on the specific subgroup.

Combine with

- Relevant indicators used, not divided by subgroups
- Use of any of the indicators mentioned for socioeconomic subgroups
- Use of any of the indicators mentioned for subgroups of household members

Outcome: Equity regarding health and health care

Indicator: Use any of the indicators mentioned for socioeconomic subgroups

Definition of the indicator

This indicator refers to equity in the context of health and health care by focusing on any indicator specific to health microinsurance for socioeconomic subgroups. Major characteristics of these socioeconomic subgroups could be their income and other measures of wealth, employment status, etc.

Theory of expected effects

The in-depth analysis of subgroups is a suitable indicator to examine the degree of equity in access to and use of health-care services. Equity in access to health care is one hypothesised impact of microinsurance based on the assumption that access to insurance can be provided to members of excluded groups. By focusing on subgroups, detailed analysis can be conducted, addressing the questions of who has access to insurance and who actually receives the benefits. In the ideal case, access to insurance should be egalitarian across socioeconomic subgroups. In the analysis of 27 Senegalese mutuelles, four Malian Equity Initiative Policies, and Ghana's Nkoranza scheme, Chankova et al. (2008) found that enrolment was significantly higher

for the top quintile, whilst there was no significant difference in the enrolment of the poorest quintile or the remaining four quintiles grouped together. Msuya et al. (2004) showed that a 1% increase in income raised the probability of joining the United Republic of Tanzania's Community Health Fund by 12.5%. Regarding the use of health-care services by subgroups, there is mixed empirical evidence. Whilst Jowett et al. (2004) found in Vietnam that insurance members of the lowest quintile make more use of the accessible health-care services, Schneider and Diop (2001) do not find a significant change in health-care utilisation by income quartile if other factors are taken into account. Wagstaff et al. (2009) did not even find any positive change in behaviour with regards to inpatient and outpatient health-care services for the lowest 10% of the income distribution, compared to the other income groups.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

For measurement, the same questions/methods mentioned for the relevant indicators should be applied to the specific subgroups.

Analysis

The analysis should be conducted in the way advised for the respective indicators, focusing on the specific subgroup.

Combine with

- Relevant indicators used, not divided by subgroups
- Use of any of the indicators mentioned for subgroups (general)
- Use of any of the indicators mentioned for subgroups of household members

Outcome: Equity regarding health and health care

Indicator: Use any of the indicators mentioned for subgroups of household members

Definition of the indicator

This indicator refers to equity in the context of health and health care by focusing on specific indicators for health microinsurance of the subgroups within households. These subgroups could be clustered by age, gender, status within the household, educational level, etc.

Theory of expected effects

The in-depth analysis of subgroups is a suitable indicator to examine the degree of equity in access to and use of health-care services. Equity in access to health care is one hypothesised impact of microinsurance based on the assumption that access to insurance can be provided to members of excluded groups. This hypothesis can also be adopted for household subgroup structures. By focusing on subgroups detailed analysis of the questions of who has access to insurance and who actually receives the benefits can be conducted. Regarding subgroups of household members, existing studies predominantly focused on gender differences and female headed households. In this context, Chankova et al. (2008) found that female headed

households in Ghana, Mali, and Senegal were more likely to enrol in insurance schemes than maleheaded households. The authors associated these findings with traditional roles for women as caregivers in households. Wagstaff et al. (2009), in contrast, did not find an increased likeliness of joining China's NCMS in female headed households. Schneider and Diop (2001) disaggregated the utilisation rate by gender and found that the probability for utilisation did not vary by gender in Rwanda.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

For measurement, the same questions/methods mentioned for the relevant indicators should be applied to the specific subgroups.

Analysis

The analysis should be conducted in the way advised for the respective

indicators, focusing on the specific subgroup.

Combine with

- Relevant indicators used, not divided by subgroups
- Use of any of the indicators mentioned for subgroups (general)
- Use of any of the indicators mentioned for socioeconomic subgroups

Impact: Risk of poverty/financial protection/financial vulnerability

Indicator: Percentage of households living below the poverty line: percentage of households living on less than USD 1.25 Purchasing Power Parity (PPP) per person per day

Definition of the indicator

This indicator captures the percentage of households in the sample living on less than USD 1.25 a day. The international line of \$1.25 a day is the average of the national poverty lines in the poorest 10-20 countries and defines those living under this estimate as extremely poor. It is the recalculated measure of the initially \$1 a day per person line, the World Bank published in 1999, based on 2005 PPP prices.

Theory of expected effects

Microinsurance is mostly promoted as a tool aimed at the prevention of unanticipated and undesirable events that could exacerbate or deepen the poverty level of affected individuals. People living with an income around the poverty line are considered particularly financially vulnerable and need efficient financial protection. Financial protection aims at providing a viable alternative to inefficient coping mechanisms often applied by low-income households in shock and stress situations. These informal coping mechanisms can include numerous sorts of

actions such as depletion of savings, selling of goods (including livestock and food), consumption-smoothing, change in labor supply, withdrawing children from school, and engaging in mutual self-help arrangements. Under microinsurance, it is expected that financial vulnerability and the risk of falling (back) into poverty decreases, particularly after a shock event.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

This indicator is applicable to all insurable risks and types of risk coverage.

Measurement

For measurement, all sources of income have to be considered in order to calculate a daily average and

determine whether it lies below or above the poverty line.

Analysis

This general indicator of household wealth and income and its degree of vulnerability is essential to analysing the impact of microinsurance on poverty reduction. It is, however, unlikely that a direct short-term effect will be found here. The indicator is rather interesting to analyse in the context of income variability as it is expected that, under insurance, individuals are less likely to fall under the poverty

threshold. It may take a longer time period until significant effects can be identified since shocks may only occur occasionally. Furthermore, it could be interesting to study whether income level is significantly correlated with the access to insurance and/or the take-up rate.

Combine with

- Percentage of households living between USD 1.25 and USD 2.00 PPP per person
- Indicators for general outcomes of microinsurance

Impact: Risk of poverty/financial protection/financial vulnerability

Indicator: Percentage of households living below the poverty line: percentage of households living on between USD 1.25 and USD 2.00 purchasing power parity (PPP) a day per person

Definition of the indicator

This indicator captures the percentage of households in the sample living between USD 1.25 and US D2.00 (normalised as PPP) a day per person. The international line of USD1.25 a day is the average of the national poverty lines in the poorest 10-20 countries and defines those living under this estimate as extremely poor. It is the recalculated measure of the initially USD1.00 a day per person line, the World Bank published in 1999, based on 2005 PPP prices. The range between USD 1.25 and USD 2.00 is a commonly-used measure for people who are particularly vulnerable to falling (back) into extreme poverty.

Theory of expected effects

Microinsurance is mostly promoted as a tool aimed at the prevention of unanticipated and undesirable events that could exacerbate or deepen the poverty level of affected individuals. People living with an income around the poverty line are considered particularly financially vulnerable and need efficient financial protection. Financial

protection aims at providing a viable alternative to inefficient coping mechanisms often applied by low-income households in shock and stress situations. These informal coping mechanisms can include numerous sorts of actions such as depletion of savings, selling of goods (including livestock and food), consumption-smoothing, change in labor supply, withdrawing children from school, and engaging in mutual self-help arrangements. Under microinsurance, it is expected that financial vulnerability and the risk of falling (back) into poverty decreases, particularly after a shock event.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

This indicator is applicable to all insurable risks and types of risk coverage.

Measurement

For measurement, all sources of income have to be considered in order to calculate a daily average and determine if it lies below or above the poverty line.

Analysis

This general indicator of household wealth and income and its degree of vulnerability is essential for the analysis of the impact of microinsurance on poverty reduction. It is, however,

unlikely to find a direct effect here. Consequently, the analysis should include various related wealth and income indicators. Furthermore, it could be interesting to study whether the income level is significantly correlated with the access to insurance and/or the take-up rate.

Combine with

- Percentage of households living on less than USD 1.25 PPP per person
- Indicators for general outcomes of microinsurance

Impact: Economic situation of the household

Indicator: Level of assets—housing conditions

Definition of the indicator

Housing conditions are an important indicator for measuring the living standard of a household. Living in satisfactory housing conditions is a highly important factor determining people's lives and their satisfaction level. Housing is essential for meeting basic needs, such as shelter, but, in addition, it should offer a place to sleep and rest where people feel safe and have privacy and personal space; it should be somewhere they can raise a family. All of these elements help make a house a home (Organisation for Economic Co-operation and Development [OECD] 2013)¹. In economic analysis, housing conditions are often used as a proxy for the economic situation of a household and its wealth status.

Theory of expected effects

Under insurance, more resources can be invested for improving the housing conditions ex-ante in the occurrence of a household shock. This could be particularly valid in a regional context, which is highly affected by devastating weather and natural events such as typhoons, floods, earthquakes, fire, etc.

In such a context, insured households could be more willing to put resources into their housing conditions, as they fear destruction less. In the ex-post situation, it is expected that the housing conditions remain more or less stable under insurance as housing assets do not need to be used as a tool to mitigate risk.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- **Property insurance**
- Other

Measurement

Measurement of housing conditions is usually conducted on the basis of the interviewer's observations and/or available data on the community level. Additionally, questions about the housing conditions can be directly included in the survey. This is also a way to provide valuable information about people's satisfaction with their housing conditions, which is an important factor of living standards.

¹ See the OECD Better Life Index at <http://www.oecdbetterlifeindex.org/topics/housing/>.

Examples for subindicators:²

- Number of rooms in the dwelling
- Main material of the floor of the dwelling
- Main material of the exterior walls in the dwelling
- Main material of the roof of the dwelling
- Percentage of households with access to tapped water within 30 minutes (walking)
- Percentage of households with tapped water at their home
- Percentage of households with a toilet in their home
- Percentage of households with electricity in their home

Analysis

Housing conditions should be analysed with consideration of general

community/neighbour housing conditions. This will help exclude local external effects (there may be no running water, electricity supply, etc., for the whole community, irrespective of a household's willingness to improve their housing conditions in this regard). The information on housing conditions can also serve as the basis for the construction of an index.

Combine with

- Total value of sold assets in case of shocks
- Percentage of assets recovered/replaced after being sold in case of shock six months after shock event
- Amount of liquid assets
- Other indicators measuring the economic situation of the household

² For more subindicators and sample questions consult the World Bank Living Standards Measurement Study surveys at http://econ.worldbank.org/wbsite/external/extdec/extresearch/extlsms/0_contentmdk:21610833~pagepk:64168427~pipk:64168435~thesitepk:3358997,00.html

Impact: Economic situation of the household

Indicator: Level of household assets / consumer appliances

Definition of the indicator

The level of household assets and consumer appliances is a suitable indicator to measure the living standard of a household. Living in satisfactory housing conditions is a highly important factor determining people's lives and their satisfaction level. In economic analysis, housing conditions and the level of household assets are often used as a proxy for the economic situation of a household and its wealth status. For this purpose, household assets and consumer appliances are used for the construction of a comparable index.

- Crop insurance (indemnity or index)
- **Property insurance**
- Other

Measurement

Measurement of household assets can be conducted on the basis of the interviewer's observations and/or questions about the housing conditions, directly included in the survey. This is also a way to provide valuable information about peoples' satisfaction with their housing conditions and level of household assets, which is an important factor of living standards.

Theory of expected effects

This indicator is particularly relevant for the analysis of an ex-post shock situation. Under insurance, it is expected that the economic situation of a household, displayed by the level of assets and consumer appliances, remains more or less stable as assets do not need to be used or sold as a tool to mitigate risk.

Examples for subindicators:

- Number of pots and pans in dwelling
- Existence of bicycle, motor bike, car (and number)
- Existence of TV, refrigerator, washing machine, AC, Hifi
- Existence of mobile phone(s)

(For more subindicators and sample questions consult the LSMS World Bank surveys.)

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)

Analysis

For the analysis of the level of assets and consumer appliances, local and cultural particularities should be

taken into account as some assets may have a higher or lower (subjective) value. Furthermore, general community/neighbour conditions should be considered in order to exclude local external effects (there may be no running water, electricity supply, etc. for the whole community, irrespective of a household's willingness to invest in certain assets/appliances such as TV, AC, washing machine, etc.). The information gathered for this indicator can be used for the construction of an asset index, functioning as a proxy for

household wealth and the economic situation of a household.

Combine with

- Total value of sold assets in case of shocks
- Percentage of assets recovered/replaced after being sold in cases of shock six months after a shock event
- Amount of liquid assets
- Other indicators measuring the economic situation of the household

Impact: Economic situation of the household

Indicator: Level of household assets—savings and other working capital

Definition of the indicator

Savings refer to the amount of household income not spent on consumption or investment, but put aside for future use to cover recurring costs, unexpected expenditures, or planned consumption (wedding, dowry, funeral, health costs, education expenses, farming inputs, etc.). The indicator includes liquid savings (bank accounts, cash-on-hand, money in savings groups, gold, jewellery, etc.) as well as illiquid savings (land ownership irrigated or non-irrigated—livestock, housing, etc). Working capital refers to operating liquidity available for business and other income-generating activities (e.g., machinery, fishing boats, stock of goods).

Theory of expected effects

Liquid and illiquid savings, as well as working capital, are important risk mitigation tools for low-income households to protect themselves against shocks and stabilise cash flow. With insurance, precautionary savings are expected to decrease, whilst funds and working capital are expected to be preserved

by insurance. If a shock occurs, assets do not need to be sold or used for other purposes such as liability for a loan.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

Measurement

Besides measurement via survey questions, direct observations of the surveyor can be a valuable source of information, particularly with regards to illiquid assets and working capital.

Analysis

Due to the expected change in risk taking behaviour, it might be interesting to analyse heterogeneous effects by risk aversion. Furthermore, different purposes and intentions behind the saving behaviour should be taken into consideration in the analysis.

Combine with

- Total value of sold assets in a shock event
- Percentage of assets recovered/ replaced after being sold in cases of shock six months after shock event
- Amount of liquid assets
- Other indicators measuring the economic situation of a household

Impact: Nutrition (in quantity and quality)

Indicator: Total expenditures on food per person per week

Definition of the indicator

This indicator captures each household's per person expenditure on food in the time frame of a week. A differentiation should be made between food bought for immediate versus future consumption. Quantity and quality of food are not taken into consideration in this indicator (see following indicators on quality and quantity of nutritional intake).

Theory of expected effects

As consumption-smoothing is a common mechanism of household risk mitigation in cases of income or health shocks, food security is a closely connected issue. As a consequence of a household shock, individuals have to cope with the expenses related to the respective event *and* continue to meet ongoing household needs, of which expenditure for food ranks on top. This double burden can easily lead to the decision to eat fewer meals or eat less nutritious food. This effect is particularly relevant for poor households, who would need to reduce their consumption and their food expenditures in the case of a shock. As microinsurance aims at stabilising,

expenditures on food are expected to stay stable or even increase if it comes to a shift in consumption preferences due to the change in risk mitigation. In this context, Wagstaff and Pradhan (2005) found that the Vietnam Health Insurance increased households' non-health related expenditures, i.e., non-medical goods such as food and education.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Health insurance can be an applicable type of risk management here, particularly if the insurance product encompasses additional interventions related to nutrition. Furthermore, in the case of self-subsistence of the household, agricultural insurance schemes, such as crop or livestock insurance, could be applicable as well. They mitigate the risk of crop failure and/or livestock death, both of which could have a negative impact on nutrition.

Measurement

Because self-reported and aggregated information of quantity and quality of food can be inaccurate, it is common practice to measure food consumption in the form of an extensive roster, covering each food item separately by amount/quantity and price for each person of a household.

Example question: I want to ask about all food consumed per person in this household. Have you consumed [FOOD] during the past 7 days? Please exclude from your answer any [FOOD] purchased for processing or resale (National Statics Directorate Caicoli, Dili, Timor Leste and World Bank 2001).

Analysis

Smoothing of food expenses is a consumption-smoothing mechanism typically employed only when other alternatives are lacking, it mostly affects poor households. It could be interesting to analyse its use across different levels of poverty. Analysing the heterogeneous effects on different types of household members could also be interesting, assuming that, in the case of a shock, a poor household would smooth food expenditures differently for its members, depending on their particular needs (especially for those household members with increased health risks such as pregnant women, children, elderly people, or sick individuals).

Combine with

- Indicators of quantity and quality of nutritional intake

Impact: **Nutrition (in quantity and quality) — quality of nutritional intake**

Indicator: **Frequency of eating vegetables or fruit (number per week)**

Definition of the indicator

This indicator captures the quality of weekly nutritional intake by the frequency of vegetables or fruit consumed.

Theory of expected effects

As consumption-smoothing is a common mechanism of household risk mitigation, in cases of income or health shock, food security is a closely connected issue. As a consequence of a household shock, individuals have to cope with the expenses related to the respective event *and* continue to meet ongoing household needs, for which expenditure for food ranks on top. This double burden can easily lead to the decision to eat fewer meals or eat less nutritious food. This effect is particularly relevant for poor households, who would need to reduce their consumption and their food expenditures less in a shock event. As microinsurance aims at stabilising, expenditures on food are expected to stay stable or even increase if it comes to a shift in consumption preferences due to the change in risk mitigation. In this context, Wagstaff and Pradhan (2005) found that the Vietnam Health Insurance increased households'

non-health related expenditures, i.e., non-medical goods such as food and education.

Applicable type of risk coverage

- **Health insurance**
- Life insurance
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- Property insurance
- Other

Health insurance can be an applicable type of risk management here, particularly if the insurance product encompasses additional interventions related to nutrition. Furthermore, in the case of self-subsistence of a household, agricultural insurance schemes, such as crop or livestock insurance, could be applicable as well. These schemes mitigate the risk of crop failure and/or livestock death, both of which could have a negative impact on nutrition.

Measurement

Because self-reported and aggregated information of quantity and quality of food can be inaccurate, it is common

practice to measure food consumption in the form of an extensive roster, covering each food item (fruit, vegetables) separately by amount/quantity and price.

Example question: I want to ask about fruit and vegetables consumed by your household, regardless of which person ate it. Has your household consumed [FOOD] during the past 7 days? Please exclude from your answer any [FOOD] purchased for processing or resale (National Statics Directorate Caicoli, Dili, Timor Leste and World Bank 2001).

Analysis

Smoothing of food expenses is a consumption-smoothing mechanism typi-

cally employed only when other alternatives are lacking, it mostly affects poor households. It could be interesting to analyse its use across different levels of poverty. Analysing the heterogeneous effects on different types of household members could also be interesting, assuming that, in the case of a shock, a poor household would smooth food expenditures differently for its members, depending on their particular needs (especially for those household members with increased health risks such as pregnant women, children, elderly people, or sick individuals).

Combine with

- Indicators of quantity and quality of nutritional intake

Impact: Nutrition (in quantity and quality) — quality of nutritional intake

Indicator: Frequency of eating meat (if people eat meat at all) (number per week)

Definition of the indicator

This indicator captures a household's weekly consumption of meat for non-vegetarian household members.

The measure is frequency, not the actual amount or type of meat.

Theory of expected effects

As consumption-smoothing is a common mechanism of household risk mitigation, in cases of income or health shock, food security is a closely connected issue. As a consequence of a household shock, individuals have to cope with the expenses related to the respective event *and* continue to meet ongoing household needs of which expenditure for food ranks on top. This double burden can easily lead to the decision to eat fewer meals or eat less nutritious food and/or less expensive food. As meat ranks amongst the most expensive foods, whilst its nutritional value can be substituted by other types of food, its consumption is often cut down first during financial constraints to food consumption. This effect is particularly relevant for poor households, who would need to reduce their

consumption and their food expenditures less in the case of a shock. As microinsurance aims at stabilising households, expenditures on food are expected to stay stable or even increase if it comes to a shift in consumption preferences due to the change in risk mitigation. In this context Wagstaff and Pradhan (2005) found that the Vietnam Health Insurance increased households' non-health related expenditures, i.e., non-medical goods such as food and education.

Applicable type of risk coverage

- **Health insurance**
- Life insurance
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- Property insurance
- Other

Health insurance can be an applicable type of risk management here particularly if the insurance product encompasses additional interventions related to nutrition. Furthermore, in the case of self-subsistence of the household agricultural insurance schemes such as crop or livestock insurance could

be applicable as well as they mitigate the risk of crop failure and/or livestock death, which could both have a negative impact on nutrition.

Measurement

Because self-reported and aggregated information of quantity and quality of food can be inaccurate, it is common practice to measure food consumption in the form of an extensive roster, covering each food item (different forms of meat) separately by amount/quantity and price. In the case of this indicator, interview questions should be preceded by the question, “Are there any people in the household who do not eat meat?”

Example question: I want to ask about meat consumed by your household, regardless of which person ate it. Has your household consumed [FOOD] during the past 7 days? Please exclude from your answer any [FOOD] purchased for processing or resale (National Statics

Directorate Caicoli, Dili, Timor Leste and World Bank 2001).

Analysis

Smoothing of food expenses is a consumption-smoothing mechanism typically employed only when other alternatives are lacking, thus, it mostly affects poor households. It could be interesting to analyse its use across different levels of poverty. Analysing the heterogeneous effects on different types of household members could be also interesting, assuming that, in a shock event, a poor household would smooth food expenditures differently for its members, depending on their particular needs (especially for those household members with higher health risks such as pregnant women, children, elderly people, or sick individuals).

Combine with

- Indicators of quantity and quality of nutritional intake

Impact: Nutrition (in quantity and quality) — quantity of nutritional intake/extent of hunger

Indicator: Average number of meals eaten per day in last month

Definition of the indicator

This indicator captures the quantity of nutritional intake by the average number of meals eaten per day in the last month.

Theory of expected effects

As consumption-smoothing is a common mechanism of household risk mitigation in cases of income or health shock, food security is a closely connected issue. As a consequence of a household shock, individuals have to cope with the expenses related to the respective event *and* continue to meet ongoing household needs, of which expenditure for food ranks on top. This double burden can easily lead to the decision to eat fewer meals or eat less nutritious food. This effect is particularly relevant for poor households, who would need to reduce their consumption and their food expenditures less in a shock event. As microinsurance aims at stabilising, expenditures on food are expected to stay stable or even increase if it comes to a shift in consumption preferences due to the change in risk mitigation. In this context, Wagstaff and Pradhan (2005) found that the Vietnam

Health Insurance increased households' non-health related expenditures, i.e., non-medical goods such as food and education. It is, thus, expected that the number of meals taken remains stable after a shock under microinsurance since the reduction of food quantity is not needed as a consumption-smoothing mechanism.

Applicable type of risk coverage

- **Health insurance**
- Life insurance
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- Property insurance
- Other

Health insurance can be an applicable type of risk management here, particularly if the insurance product encompasses additional interventions related to nutrition. Furthermore, in the case of self-subsistence of the household agricultural insurance schemes such as crop or livestock insurance could be applicable as well as they mitigate the risk of crop failure and/or livestock death, which could both have a negative impact on nutrition.

Measurement

- What was the average number of meals eaten per day on average during the past month?

Analysis

Smoothing of food expenses is a consumption-smoothing mechanism typically employed only when other alternatives are lacking, thus, it mostly affects poor households. It could be interesting to analyse its use across different levels of poverty. Analysing the heterogeneous

effects on different types of household members could also be interesting, assuming that, in a shock event a poor household, would smooth food expenditures differently for its members, depending on their particular needs (especially for those household members with increased health risks such as pregnant women, children, elderly people, or sick individuals).

Combine with

- Indicators of quantity and quality of nutritional intake

Impact: Nutrition (in quantity and quality) — quantity of nutritional intake/extent of hunger

Indicator: Number of days when food was insufficient for the household in last month

Definition of the indicator

This indicator refers to the number of days when food was insufficient for a household in the last month. This is a subjective measure, capturing the extent of hunger that is suffered amongst household members.

Theory of expected effects

As consumption-smoothing is a common mechanism of household risk mitigation in cases of income or health shock, food security is a closely connected issue. As a consequence of a household shock, individuals have to cope with the expenses related to the respective event *and* continue to meet ongoing household needs, of which expenditure for food ranks on top. This double burden can easily lead to the decision to eat fewer meals or eat less nutritious food. This effect is particularly relevant for poor households, who would need to reduce their consumption and their food expenditures in the case of a shock event. As microinsurance aims at stabilising households, expenditures on food are expected to stay stable or even increase if it comes to a shift in consumption preferences

due to the change in risk mitigation. In this context Wagstaff and Pradhan (2005) found that the Vietnam Health Insurance increased households' non-health related expenditures, i.e., non-medical goods such as food and education. Thus, it is expected that the number of days with insufficient supply of food will decrease under microinsurance, since consumption-smoothing is no longer necessary to mitigate the consequences of the shock. This is particularly true for an ex-post shock context.

Applicable type of risk coverage

- **Health insurance**
- Life insurance
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- Property insurance
- Other

Health insurance can be an applicable type of risk management here particularly if the insurance product encompasses additional interventions related to nutrition. Furthermore, in the case of subsistence farming of the household, agricultural insurance schemes

such as crop or livestock insurance could be applicable as well as they mitigate the risk of crop failure and/or livestock death, which could both have a negative impact on nutrition.

Measurement

The insufficiency of food in a household can be treated as a fully subjective indicator, based on the self-assessment of survey respondents and their feeling of hunger. A further option of measurement is to take the caloric intake, officially recommended by the WHO as a reference for sufficient/insufficient consumption of food, and use the collected information about food consumption.

Analysis

Smoothing of food expenses is a consumption-smoothing mechanism typically employed only when other alternatives are lacking, thus, it mostly affects poor households. It could be interesting to analyse its use across different levels of poverty. Analysing the heterogeneous effects on different types of household members could also be interesting, assuming that, in a shock event, a poor household would smooth food expenditures differently for its members, depending on their particular needs (especially for those household members with increased health risks such as pregnant women, children, elderly people, or sick individuals).

Combine with

- Indicators of quantity and quality of nutritional intake

Impact: Health/Morbidity

Indicator: Number of sick days in the household per person within last month (per subgroups such as children, women, elderly, etc.)

Definition of the indicator

This indicator measures the number of illness episodes per subgrouped individuals in a household within the last month.

Theory of expected effects

Microinsurance is intended to provide a reliable, adequate level of affordable health care (Leatherman et al. 2012). Nevertheless, the direct impact of a microinsurance scheme on health outcomes is a particularly difficult target to measure and most valid outcome data existing use proxies for health outcomes and focus on specific aspects of health care such as health-care utilisation or access. The number of sick days in a household per person could be such a proxy for health status/morbidity. If microinsurance provides access to health-care services that are inaccessible without insurance, one would expect that, in the medium- to long-term, the number of sick days would decrease. At a minimum, this decrease would be expected for illnesses that are easily treatable in most cases, such

as diarrhea. Due to the difficulties of measuring direct health outcomes, empirical studies focused instead on subjective self-reports of health status. In this context, Lei and Lin (2009) estimated that members of China's NCMS, were 2.8% less likely to report that they were feeling ill.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

Regarding the measurement of illness, DHS distinguishes between different degrees of severity of illness. Thus, it can be useful to cover both general illnesses and illnesses considered as rather severe, including fever and/or coughing as indicators.

- How many days has [NAME] been ill in the last month?
- How many days has [NAME] been ill with a fever/cough in the last month? (ICF 2011)

Analysis

Because it is expected that the number of sick days in a household per person is primarily influenced by general living conditions and numerous external factors, such as epidemics and local infection rates, number of sick people in close surroundings, water and sanitation, food quality and accessibility,

etc., it is very difficult to analyse a direct impact of microinsurance on health status and morbidity. The analysis should, thus, include as many other factors and indicators as possible.

Combine with

- Other indicators of health/morbidity
- Indicators of living standards

Impact: Health/Morbidity

Indicator: Number of days household members were unable to perform usual activities because of poor health per person within last month

Definition of the indicator

This indicator measures the number of days household members could not carry out their usual activities due to poor health.

Theory of expected effects

Microinsurance is intended to provide a reliable, adequate level of affordable health care (Leatherman et al. 2012). Nevertheless, the direct impact of a microinsurance scheme on health outcomes is a particularly difficult target to measure. Most of the valid, existing outcome data are based on proxies for health outcomes and focus on specific aspects of health care such as health-care utilisation or access. The number of days household members could not carry out their usual activities due to poor health conditions could be such a proxy for health status/morbidity. The discontinuation of usual activities is an appropriate indicator for health status, as low-income households cannot afford to forgo any regular income source and discontinuation is often a decision of last resort. If microinsurance provided access to health-care services that are inaccessible

without insurance, one would expect that, in the medium- to long-term, the number of sick days would decrease and, thus, usual activities of household members could continue. Due to the difficulties of measuring direct health outcomes, empirical studies focused instead on subjective self-reports of health status. In this context, Lei and Lin (2009) estimated that members of China's NCMS, were 2.8% less likely to report that they were feeling ill.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

- How many days has [NAME] been unable to perform her/his usual activities because of poor health within the last month?

Analysis

As it is expected that the number of sick days in a household per person

is primarily influenced by general living conditions and numerous external factors, such as epidemics and local infection rates, number of sick people in close surroundings, water and sanitation, food quality and accessibility, etc., it is very difficult to analyse a direct impact of microinsurance on this health status and morbidity. The analysis should, thus, include as many other factors and indicators as possible. Furthermore, the indicator should be seen in the context of the

prevalent illness and its severity. This could provide important information about the long-term impact of an illness, as for instance in the case of a chronic disease where usual activities are likely to be stopped for a long period of time.

Combine with

- Other indicators of health/morbidity
- Indicators of household income of the relevant activities

Impact: Health/Morbidity

Indicator: Body Mass Index (BMI)

Definition of the indicator

This indicator measures the Body Mass Index (BMI) for all household members. The BMI is a simple index of weight-for-height that is commonly used to classify underweight, overweight and obesity in adults. It is defined as the weight in kilograms divided by the square of the height in metres (kg/m²). BMI values are age-independent and the same for both sexes. However, BMI may not correspond to the same degree of obesity in different populations due, in part, to different body proportions. The health risks associated with increasing or very low BMI are continuous and the interpretation of BMI gradings in relation to risk may differ for different populations. In recent years, there has been a growing debate on whether there are possible needs for developing different BMI cut-off points for different ethnic groups due to the increasing evidence that the associations between BMI, percentage of body fat, and body fat distribution differ across populations (World Health Organization 2013).

Theory of expected effects

Wagstaff and Pradhan (2005) conducted one of the rare studies, examining direct health outcomes of

microinsurance. Based on data from the Living Standards Measurement Survey, they found that the Vietnamese Health Insurance significantly influences the BMI of adults and height and weight of young children.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

The BMI can be calculated by the formula:

$$\text{BMI} = \text{weight (kg)} / (\text{height (m)})^2$$

If possible, measurement of weight and height is directly conducted by the surveyor, and not based on self-reporting, in order to retrieve the most exact information possible.

Analysis

As pointed out for the previous indicators for health outcomes, it is very difficult to analyse a direct impact of microinsurance on this health status

and morbidity. The analysis should, thus, include as many other factors and indicators as possible. The ongoing discussion about the BMIs generalisability across ethnic groups also points to the importance of seeing this indicator in the local context.

Combine with

- Other indicators of health/morbidity
- Nutrition-related variables

Impact: Health/Morbidity

Indicator: Percentage of children with anemia

Definition of the indicator

This indicator measures the percentage of children within a household who are anemic. Anemia is considered as a public health problem in both rich and poor countries by the WHO. However, its prevalence in developing countries is particularly high, with nearly half of all women and children being anemic. Iron deficiency has been identified as a widespread cause of anemia. Nevertheless, there are numerous other factors associated with anemia such as malaria, parasitic infections, nutritional deficiencies, and hemoglobinopathies. Overall, anemia is perceived as an indicator of poor health and poor nutrition. Anemia can occur at all stages of the life cycle but is more prevalent in pregnant women and young children, on whom it has the most devastating health effects. Severe forms of anemia are directly related to an increased maternal and child mortality. The most prevalent form of anemia—Iron Deficiency Anemia (IDA)—is associated with negative consequences for the cognitive and physical development of children and on the physical performance of adults (Benoist et al. 2005). Particularly in children, anemia has been associated with impaired cognitive performance,

motor development, coordination, language development, and scholastic achievement (ICF 2011).

Theory of expected effects

Anemia is a very good example of a disease that is relatively easy to prevent and treat but still very prevalent and associated with acutely negative health outcomes, particularly in young children and pregnant women. Increased access to and use of health-care services, including educational advice on nutrition, is expected to decrease the percentage of children with anemia.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

Data on anemia could be retrieved from local health agencies or, if possible, directly by a blood test. The DHS contains a large section on anemia, for which women (15–49 years) and children (usually six-months up to five years) are tested for the disease

through finger prick or, in the case of young children, heel prick blood testing using the HemoCue blood hemoglobin testing system. In the case of DHS, testing is voluntary and respondents receive the results of their anemia test immediately as well as information about how to prevent anemia (ICF 2011).

Analysis

The analysis should take into account whether the microinsurance scheme in focus encompasses any complementary educational or active interventions with regards to appropriate

nutrition of children, particularly regarding the supplementation of iron. or whether the individuals are subject to such an intervention independent of their membership in a microinsurance scheme. Due to the correlation between maternal nutrition and health status and the prevalence of anemia for young children, it could also be interesting to control for the prevalence of anemia in children by anemia status of the mother.

Combine with

- Other indicators of health/morbidity
- Nutrition-related variables

Impact: Health/Morbidity

Indicator: Percentage of women with anemia

Definition of the indicator

This indicator measures the percentage of women within a household, who are anemic. Anemia is considered as a public health problem in both rich and poor countries by the WHO. Iron deficiency has been identified as a widespread cause of anemia. Nevertheless, there are numerous other factors associated with anemia such as malaria, parasitic infections, nutritional deficiencies, and hemoglobinopathies. Overall, anemia is perceived as an indicator of poor health and poor nutrition. Anemia can occur at all stages of the life cycle but is more common in pregnant women and young children, on whom it has the most devastating health effects. Severe forms of anemia are directly related to an increased maternal and child mortality. The most prevalent form of anemia—Iron Deficiency Anemia (IDA)—is associated with negative consequences for the cognitive and physical development of children and on the physical performance of adults (Benoist et al. 2005). For women, anemia reduces their work productivity and places them at risk for poor pregnancy outcomes, including increased risk of maternal mortality, prenatal mortality, premature births, and low birth weight (ICF 2011).

Theory of expected effects

Anemia is a very good example of a disease that is relatively easy to prevent and treat but still very prevalent and associated with very negative health outcomes, particularly in young children and pregnant women. Increased access to and use of services, including educational advice on nutrition, is expected to decrease the percentage of women with anemia.

Applicable type of risk coverage

- **Health insurance**

As this indicator explicitly refers to the case of health, the applicable type of risk coverage here is health insurance.

Measurement

Data on anemia could be retrieved from local health agencies or, if possible, directly by a blood test. The DHS contains a large section on anemia, for which women (15-49 years) and children (usually six-months up to five-years) are tested for the disease through finger prick or, in the case of young children, heel prick blood testing using the HemoCue blood hemoglobin testing

system. In the case of DHS, testing is voluntary and respondents receive the results of their anemia test immediately, as well as information about how to prevent anemia (ICF 2011).

Analysis

The analysis should take into account whether the microinsurance scheme at stake encompasses any complementary educational or active interventions with regards to appropriate nutrition, particularly regarding the supplementation of iron, or

whether the individuals are subject to such an intervention independent of their membership in a microinsurance scheme. Due to the correlation between maternal nutrition and health status and the prevalence of anemia for young children, it could also be interesting to control for the prevalence of anemia in children by anemia status of the mother.

Combine with

- Other indicators of health/morbidity
- Nutrition-related variables

Impact: Educational attainment of children

Indicator: Percentage of children age 6–16 (or other age) attending school

Definition of the indicator

This indicator measures the percentage of children between the ages of six and 16 attending school. The indicator, thus, focuses on primary, middle and high school (until approximately 10th grade) education.

Theory of expected effects

When a household lacks alternatives for coping with a shock event, they may be forced to not only use their financial assets, but also their physical ones to handle the corresponding consequences. Taking children out of school can be based on two intentions. One is to save money, if school fees and other expenses are needed to keep the children in school; the other intention is to take children out of school in order to put them to work in order to cope with the consequences of the shock. Seen from a long term perspective, this coping mechanism is inefficient. It impedes the educational and skills development of children, which is essential for the future socioeconomic situation of a household. Under insurance, it is expected that children will remain in school after a shock event as educational expenses can still be covered

and no additional man power or support are needed in their household.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

Numerous types of insurances are applicable in this case since taking children out of school can be a reaction to all sorts of shock events, placing households in a difficult economic situations.

Measurement

The indicator can be measured both on a household level as well as on a community/school level by using official statistics of school attendance rates.

Analysis

Taking children out of school is a typical coping mechanism of poor to very poor households who lack of alternative coping mechanisms. However, as

this indicator implies not only financial but also physical coping in the form of putting the children to work, the focus should be taken on households involved in agricultural production or small business, as additional man power is most likely to be needed here. Furthermore, it could be interesting to control for the age and sex of the children taken out of school. Taking children out of primary school for example can be more devastating for the educational development than at a later stage of school. For some country contexts, studies show that girls are more likely to be taken out

of school to support a household than boys, as their educational prospects are valued less. In this context, it could be interesting to combine the indicator with a question about how much a household spent for education during the past 12 months.

Combine with

- Other indicators of educational attainment of children
- Other indicators of risk management strategies (ex-post)/coping strategies in case of shock
- Child labour measures

Impact: Educational attainment of children

Indicator: Days missed in school per child within last month

Definition of the indicator

This indicator measures the days of absence from school per child within the last month.

Theory of expected effects

In the case of a lack of alternatives to cope with a shock event, households may be forced to not only use their financial but also physical assets to cope with the corresponding consequences. Taking children out of school can be based on two intentions. One is to save money, if school fees and other expenses are needed to keep the children in school; the other intention is to take children out of school in order to put them to work in order to cope with the consequences of the shock. Under insurance, it is expected that children will remain in school after a shock event as a result of the financial protection of the insurance scheme, making additional income less necessary for a household. In addition to these economic aspects, school days may be missed if a child is too weak or sick to go to attend. In this context, experimental evidence on deworming interventions demonstrated that a better health status significantly increases

the ability to attend school (J-PAL 2012). Seen from a long term perspective this coping mechanism is inefficient, as it impedes the educational and skills development of the children, which are essential for the future socio-economic situation of the household. Under insurance it is expected that children remain in school after a shock event, due to the financial protection of the insurance scheme, making additional income of the household less needed. Furthermore, a potentially enhanced health status may increase a child's physical and mental ability to attend school.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

Numerous types of insurances are applicable in this case since taking children out of school can be a reaction to all sorts of shock events, placing households in difficult economic situations.

Measurement

- How many days did [CHILD] miss school during the past month?

The indicator can be measured both on the household level as well as on the community/school level by using official statistics of school attendance rates.

Analysis

Taking children out of school is a typical coping mechanism of poor to very poor households who lack alternative coping mechanisms. However, as this indicator implies not only financial but also physical coping in the form of putting the children to work, the focus should be taken on households involved in agricultural production or small business, as additional manpower is most likely to be needed here. Furthermore, it could be interesting to control for the age and sex of the children taken out of school. Taking children out of primary school, for example, can be more devastating for the educational development than at a later stage of school.

For some country contexts, studies show that girls are more likely to be taken out of school to support a household than boys, as their educational prospects are valued less. Following the argument that health problems are a potential impediment to regular school attendance, the analysis should take into account whether children benefit from any health interventions, such as deworming programmes, that would make absences due to preventable diseases and health problems less likely. As the indicator refers to the days missed in school during the preceding month, the analysis should take into account whether there was a shock event during this month or recently or a particular occasion, such as harvest season, which often requires increased labor supply within the family.

Combine with

- Other indicators of educational attainment of children
- Other indicators of risk management strategies (ex-post)/coping strategies in shock event
- Child labour measures

Impact: Educational attainment of children — child labour

Indicator: Percentage of children below age 15 (or other age) engaged in income generation, housework, farming, or other economic activity as their main occupation

Definition of the indicator

This indicator measures the percentage of children below the age of 15 (or other age), who are engaged in income generation for a household.

Theory of expected effects

In the case of a lack of alternatives to cope with a shock event, households may be forced to not only use their financial but also physical assets to cope with the corresponding consequences. Taking children out of school and engaging them in household income generation can be based on two intentions. One is to save money, if school fees and other expenses are needed to keep the children in school; the other intention is to take children out of school in order to put them to work in order to cope with the consequences of the shock. Seen from a long term perspective this coping mechanism is inefficient, as it impedes the educational and skills development of the children, which are essential for the future socioeconomic situation of the household. Under insurance it is expected that children remain in

school after a shock event as educational expenses can still be covered and no additional manpower and support is needed in the household.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Numerous types of insurances are applicable in this case since taking children out of school can be a reaction to all sorts of shock events, placing households in difficult economic situations.

Measurement

For the measurement of this indicator, official statistics of school attendance rates below the age of 15 could be used. Directly asked to a household, this question could be considered as very sensitive, particularly in contexts where school attendance is obligatory

up until a certain level/age and child labour prohibited. Thus, consider obtaining the information via indirect questions about the engagement of children in work for the household.

Analysis

Taking children out of school is a typical coping mechanism of poor to very poor households who lack alternative coping mechanisms. However, as this indicator implies not only financial but also physical coping in the form of putting the children to work, the focus should be taken on households involved in agricultural production or small business, as additional man power is most likely to be needed here. The indicator should be seen in the light of seasonal or other context particularities and analysed regarding

its duration, i.e., whether the income generated by a child is of short-term or long-term duration. Furthermore, it could be interesting to control for the age and sex of the children taken out of school. Taking children out of primary school, for example, can be more devastating for the educational development than at a later stage of school. For some country contexts, studies show that girls are more likely to be taken out of school to support a household than boys, as their educational prospects are valued less.

Combine with

- Other indicators of educational attainment of children
- Other indicators of risk management strategies (ex-post)/coping strategies in case of shock

Impact: Educational attainment of children—child labour

Indicator: Average number of working hours per week for children below age 15

Definition of the indicator

This indicator measures the working hours of children below the age of 15.

Theory of expected effects

In the case of a lack of alternatives to cope with a shock event, households may be forced to not only use their financial but also physical assets to cope with the corresponding consequences. Taking children out of school and engaging them in household income generation can be based on two intentions. One is to save money, if school fees and other expenses are needed to keep the children in school; the other intention is to take children out of school in order to put them to work in order to cope with the consequences of the shock. Seen from a long term perspective this coping mechanism is inefficient, as it impedes the educational and skills development of the children, which are essential for the future socioeconomic situation of the household. Under insurance it is expected that children remain in school after a shock event as educational expenses can still be covered and no additional manpower and support is needed in the household.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Numerous types of insurances are applicable in this case since taking children out of school can be a reaction to all sorts of shock events, placing households in difficult economic situations.

Measurement

- How many hours a week does [CHILD] work per week for the household?

Analysis

Taking children out of school is a typical coping mechanism of poor to very poor households who lack alternative coping mechanisms. However, as this indicator implies not only financial but also physical coping in the form of putting the children to work, the focus should be taken on households involved in

agricultural production or small business, as additional man power is most likely to be needed here. The indicator should be seen in the light of seasonal or other context particularities and analysed regarding its duration, i.e., whether the income generated by a child is of short-term or long-term duration. Furthermore, it could be interesting to control for the age and sex of the children taken out of school. Taking children out of primary school, for example, can be more devastating for the educational development than

at a later stage of school. For some country contexts, studies show that girls are more likely to be taken out of school to support a household than boys, as their educational prospects are less valued.

Combine with

- Other indicators of educational attainment of children
- Other indicators of risk management strategies (ex-post)/coping strategies in case of shock

Impact: Educational attainment of children — child labour

Indicator: Percentage of children age 6–16 (or other age) attending school

Definition of the indicator

This indicator measures the percentage of children at the age of 6 to 16 attending school. The indicator, thus, focuses on primary as well as middle or high school (until approximately 10th grade) education.

Theory of expected effects

In the case of a lack of alternatives to cope with a shock event, households may be forced to not only use their financial but also physical assets to cope with the corresponding consequences. Taking children out of school can be based on two intentions. One is to save money, if school fees and other expenses are needed to keep the children in school; the other intention is to take children out of school in order to put them to work in order to cope with the consequences of the shock. Seen from a long term perspective this coping mechanism is inefficient, as it impedes the educational and skills development of the children, which are essential for the future socio-economic situation of the household. Under insurance it is expected that children remain in school after a shock event as educational expenses can

still be covered and no additional manpower and support is needed in the household.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

Numerous types of insurances are applicable in this case since taking children out of school can be a reaction to all sorts of shock events, placing households in difficult economic situations.

Measurement

The indicator can be measured both on the household level as well as on the community/school level by using official statistics of school attendance rates.

Analysis

Taking children out of school is a typical coping mechanism of poor to very poor households, who lack of

alternative coping mechanisms. However, as this indicator implies not only financial but also physical coping in the form of putting the children to work, the focus should be taken on households involved in agricultural production or small business, as additional manpower is most likely to be needed here. Furthermore, it could be interesting to control for the age and sex of the children taken out of school. Taking children out of primary school, for example, can be more devastating for the educational development than

at a later stage of school. For some country contexts, studies show that girls are more likely to be taken out of school to support a household than boys, as their educational prospects are valued less.

Combine with

- Other indicators of educational attainment of children
- Other indicators of risk management strategies (ex-post)/coping strategies in case of shock

Impact: Educational attainment of children — child labour

Indicator: Days missed in school per child within last month due to labour

Definition of the indicator

This indicator measures the days of absence from school per child within the last month, which were particularly due to labour of the child.

Theory of expected effects

In the case of a lack of alternatives to cope with a shock event, households may be forced to not only use their financial but also physical assets to cope with the corresponding consequences. Taking children out of school can be based on two intentions. One is to save money, if school fees and other expenses are needed to keep the children in school; the other intention is to take children out of school in order to put them to work in order to cope with the consequences of the shock. Seen from a long term perspective this coping mechanism is inefficient, as it impedes the educational and skills development of the children, which are essential for the future socioeconomic situation of the household. Under insurance it is expected that children remain in school after a shock event as educational expenses can still be covered and no additional man power

and support is needed in the household and, thus, days of absence will decrease for these periods.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

Numerous types of insurances are applicable in this case since taking children out of school can be a reaction to all sorts of shock events, placing households in difficult economic situations.

Measurement

- How many days did [CHILD] miss school during the past month to support the family with work?

Note: This question may be sensitive in contexts where school attendance is obligatory by law. In this case, the question may be restricted to the days missed at school without mentioning the particular reason.

Analysis

Taking children out of school is a typical coping mechanism of poor to very poor households, who lack alternative coping mechanisms. However, as this indicator implies not only financial but also physical coping in the form of putting the children to work, the focus should be taken on households involved in agricultural production or small business, as additional man power is most likely to be needed here. Furthermore, it could be interesting to control for the age and sex of the children taken out of school. Taking children out of primary school, for example, can be more devastating for the educational development than at a later stage of school. For

some country contexts, studies show that girls are more likely to be taken out of school to support a household than boys, as their educational prospects are valued less. As the indicator refers to the days missed in school during the preceding month, the analysis should take into account whether there was a shock event during this month, recently, or during a particular occasion, such as harvest season, which often requires increased labor supply within the family.

Combine with

- Other indicators of educational attainment of children
- Other indicators of risk management strategies (ex-post)/coping strategies in case of shock

Outcome: Educational attainment of children — child labour

Indicator: Highest education levels attained by children and young adults in household (below age 20)

Definition of the indicator

This indicator measures the highest education levels attained by household members below the age of 20.

Theory of expected effects

In the case of a lack of alternatives to cope with a shock event, households may be forced to not only use their financial but also physical assets to cope with the corresponding consequences. Taking children out of school can be based on two intentions. One is to save money, if school fees and other expenses are needed to keep the children in school; the other intention is to take children out of school in order to put them to work in order to cope with the consequences of the shock. Seen from a long term perspective this coping mechanism is inefficient, as it impedes the educational and skills development of the children, which are essential for the future socioeconomic situation of the household. Under insurance it is expected that children remain in school after a shock event as educational expenses can still be covered and no additional man power and support is needed in the household

and, thus, the likeliness of higher educational levels attained by household members below the age of 20 is expected to increase in the long run.

Applicable type of risk coverage

- **Health insurance**
- **Life insurance**
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- **Property insurance**
- **Other**

Numerous types of insurances are applicable in this case since taking children out of school can be a reaction to all sorts of shock events, placing households in difficult economic situations.

Measurement

The measurement of the highest educational levels attained can be focused on the highest level completed or the highest degree obtained.

- What is the highest educational level you completed?
- What is the highest degree you attained? (National Bureau of Statistics, Nigeria 2010)

Analysis

This indicator has to be analysed in a long-term perspective as short-term changes are unlikely to occur after the take-up of insurance.

Combine with

- Other indicators concerning the educational attainment of children
- Other indicators about risk management strategies (ex-post)/coping strategies in shock event

Impact: Peace of mind/perception of financial vulnerability

Indicator: Changes in risk prioritisation through risk ranking exercise

Definition of the indicator

This indicator measures changes in risk prioritisation through a risk ranking exercise. Risk ranking exercises can be conducted in various manners, with the basic principle that participants are asked to rank all risks that are applicable to their situation.

Theory of expected effects

Under the assumption that risk prioritisation is based on both recent actual shock events and subjective perceptions about prevalent risk, it is expected that a change in risk prioritisation can be found for insured individuals, down-ranking insured risks. This is primarily due to a reduction of fear and worries about consequences associated with the insured risks.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

All types of risk coverage are relevant for this indicator as the risk ranking exercise includes all risks applicable for the studied context.

Measurement

For measurement, risk ranking exercises should be conducted with treatment and control groups. The exercise can be carried out in many different ways. Morsink and Geurts (2012) conducted a classical risk ranking exercise with three steps. In the first step, participants viewed all cards indicating risks and were asked for any amendments necessary. In the second step, they were asked to collect all cards with risks that are applicable to them. In the third step, they were asked to select the risk they worried about most. This card was put away and the procedure continued until a complete rank-order was reached.

Analysis

Because risk prioritisation is not only influenced by risk coverage available, but also—and most importantly—by the actual risks people are exposed to and have been exposed to in the past, these incidents should be considered in

the analysis. It might also be interesting to analyse heterogeneous effects by risk aversion because the decisions of risk averse individuals should be more affected by insurance.

Combine with

- Perception about confidence in the future
- Perception about financial vulnerability

Impact: Peace of mind/perception of financial vulnerability

Indicator: Level of confidence about the future

Definition of the indicator

This indicator measures the level of confidence an individual has about the future. It refers to the strong belief that the socioeconomic situation will develop in a positive way in the future. Faith and trust in the future can be used synonymously in this context.

Theory of expected effects

Microinsurance is intended to increase individual empowerment and mental peace of mind by reducing fears and worries about the future through an increased feeling of security. A confident perception about the future is essential for risk taking behaviour and decisions about productive investment. Karlan et al. (2012) show this “peace of mind effect”, implying that insurance can have both a protective and productive effect for clients. They found that farmers provided with insurance against weather-related risk increase productive expenditures on their farms, and their demand for insurance increases as they, or an acquaintance, experience an insurance payout.

Applicable type of risk coverage

- **Health insurance**
- Life insurance
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- Property insurance
- Other

Because levels of confidence about the future are primarily based on the perception and assessment of the present situation, the types of risk coverage most applicable are those related to productive investment (livestock and crop insurance) as well as health insurance.

Measurement

- How confident are you about the future? (rank on scale from not confident to very confident)

Analysis

It could be interesting to analyse whether increased confidence in the future has a noticeable impact on productive and

investment decisions. It might also be interesting to analyse heterogeneous effects by risk aversion because the decisions of risk averse individuals should be more affected by insurance.

Combine with

- Changes in risk prioritisation through risk ranking exercise
- Perception of financial vulnerability

Impact: Peace of mind/perception of financial vulnerability

Indicator: Perception of financial vulnerability

Definition of the indicator

This indicator measures the individual's perception of financial vulnerability: the subjective assessment by individuals of their economic situation and its vulnerability.

Theory of expected effects

The ultimate proclaimed role of microinsurance is to reduce the vulnerability of people living on low incomes by enabling them to manage their risks more efficiently. Consequently, it is expected that microinsurance also evokes a subjective change in perception about financial vulnerability.

Applicable type of risk coverage

- **Health insurance**
- Life insurance
- **Livestock insurance (indemnity or index)**
- **Crop insurance (indemnity or index)**
- Property insurance
- Other

Because levels of confidence about the future are primarily based on the perception and assessment of the present situation, the types of risk coverage most applicable are those related to productive investment (livestock and crop insurance) as well as health insurance.

Measurement

- How much do you worry that a severe financial shock could occur to your household?

Analysis

It could be interesting to compare the perception of financial vulnerability with the actual financial vulnerability retrieved from indicators of wealth and income.

Combine with

- Changes in risk prioritisation through risk ranking exercise
- Perception about confidence in the future

Impact: Reliance on informal risk sharing networks

Indicator: Total amount of money currently borrowed from others without interest

Definition of the indicator

The indicator captures the total amount of money borrowed from others (family, friends, informal networks, etc.) without interest.

Numerous types of insurances are applicable in this case as borrowing from informal networks can be a reaction to all sorts of shock events placing the household in a difficult economic situation or other non-shock related occasions.

Theory of expected effects

Borrowing from others in the community without interest is a positive sign for strong social capital within a community. However, informal borrowing from others in the community can involve putting pressure on the borrower other than interest such as social pressure, expectations of reciprocity, etc. Under insurance, a direct effect expected is that, in the context of ex-post risk mitigation, less borrowing from informal networks is conducted to cushion the shock (Dercon et al. 2012).

Measurement

- How much do you currently borrow from others without paying interest?

Analysis

Here, it could be interesting to control for different informal sources (could be more than one) of borrowing. Furthermore, it could be interesting to investigate whether the borrowing is bound to any other kind of obligation than interest (expectations of reciprocity, etc). Furthermore, controlling for formal lending could be also of interest in order to find out more about the lending practices of a household (rather informal or formal) and whether the choice to borrow informally is due to independent preferences or lack of access to formal mechanisms.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Combine with

- Total amount of money currently lent to others without interest

Impact: Reliance on informal risk sharing networks

Indicator: Total amount of money currently lent to others without interest

Definition of the indicator

The indicator captures social capital by the total amount of money lent to others (family, friends, informal networks, etc.) without interest.

Theory of expected effects

Borrowing from and lending to others in the community without interest is a positive sign for strong social capital within a community and intrinsically desirable. However, informal lending to others in the community can involve features that put pressure on the borrower other than interest such as social pressure and expectations of reciprocity. When microinsurance is available and taken-up, a direct effect expected is that, in the context of ex-post risk mitigation, less lending needs to be conducted to cushion the shock (Dercon et al. 2012). Also, the willingness and ability to provide more loans to family and community members might increase, as the funds are less needed as precautionary assets for people's own purposes. Otherwise, insured individuals might be less willing to help those who did not behave cautiously and refused insurance.

Applicable type of risk coverage

- Health insurance
- Life insurance
- Livestock insurance (indemnity or index)
- Crop insurance (indemnity or index)
- Property insurance
- Other

Numerous types of insurances are applicable in this case since lending to informal networks can be a reaction to all sorts of shock events or other non-shock related occasions, placing households in difficult economic situations.

Measurement

- How much do you currently lend to others without taking interest?

Analysis

It could be interesting to investigate whether the lending is bound to any other kind of obligation than interest (expectations of reciprocity, in-kind compensations, or services, etc.). Furthermore, details of the loan arrangement, such as repayment

policies, are interesting to consider in the analysis.

Combine with

- Total amount of money currently borrowed from others without interest
- Total amount of loans (currently pending) *given* to family members and other community members

References

- Anderson, S. and J.M. Baland. 2002.** The economics of roscas and intra-household resource allocation. *Quarterly Journal of Economics* 117(3): 963-995.
- Benoist, B., E. McLean, I. Egli, and M. Cogswell, eds. 2005.** Worldwide prevalence of anaemia 1993-2005. WHO Global Database on Anaemia. Geneva: World Health Organization (WHO).
- Besley, T. and S. Coate. 1995.** Group lending, repayment incentives and social collateral. *Journal of Development Economics* 46(1): 1-18.
- Cai, H., Y. Chen, H. Fang, and L. Zhou. 2010.** Microinsurance, trust and economic development: Evidence from a randomized natural field experiment. National Bureau of Economic Research (NBER) Working Paper no. 15396. Cambridge, MA. <http://www.nber.org/papers/w15396>.
- Chankova, S., S. Sulzbach, and F. Diop. 2008.** Impact of mutual health organizations: Evidence from West Africa. *Health Policy and Planning* 23(4): 264-276.
- Dercon, S., J. W. Gunning, A. Zeitlin, and S. Lombardini. 2012.** The impact of a health insurance programme: Evidence from a randomized controlled trial in Kenya. Research Paper no. 24. International Labour Office, Microinsurance Innovations Facility, Geneva.
- Dercon, S. and L. Christiaensen. 2011.** Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. *Journal of Development Economics* 96(2): 159-173.
- Derriennic, Y., K. Wolf, and P. Kiwanuka-Mukiibi. 2005.** An assessment of community-based health financing activities in Uganda. The Partners for Health Reformplus Project. Bethesda: Abt Associates Inc.
- Franco, L. M., F. Diop, C. R. Burgert, A. G. Kelley, M. Makinen, and C. H. T. Simpara. 2008.** Effects of mutual health organisations on use of priority health-care services in urban and rural Mali: A case-control study. *Bulletin of the World Health Organization* 86(11): 830-838.
- Giné, X. and D. Yang. 2009.** Insurance, credit, and technology adoption: Field experimental evidence from Malawi. *Journal of Development Economics* 89(2009): 1-11.
- Hintz, M. 2010.** Understanding the context is understanding the impact: Evidence from a qualitative microinsurance impact survey in Indonesia. In: *Microinsurance: An innovative tool for risk and disaster management*, ed. E. Morelli, G. A. Onnis, W. J. Ammann, and C. Sutter, 109-135. Davos: Global Risk Forum (GRF).
- ICF International (2011).** Demographic and Health Surveys Methodology - Questionnaires: Household, Woman's, and Man's. MEASURE DHS Phase III: Calverton, Maryland, USA. (<http://www.measuredhs.com/publications/publication-DHSQ6-DHS-Questionnaires-and-Manuals.cfm>)
- Jowett, M., A. Deolalikar, and P. Martinsson. 2004.** Health insurance and treatment seeking behavior: Evidence from a low-income country. *Health Economics* 13:845-857.
- Abdul Latif Jameel Poverty Action Lab (J-PAL). 2012.** Deworming: A best buy for development. *J-PAL Policy Bulletin* (March).
- Jütting, J. P. 2004.** Do community-based health insurance schemes improve poor people's access to health care? Evidence from rural Senegal. *World Development* 32(2): 273-288.
- Karlan, D., R. D. Osei, I. Osei-Akoto, and C. Udry. 2012.** Agricultural decisions after relaxing credit and risk constraints. Facility Research Paper No. 23. International Labour Office, Microinsurance Innovation Facility, Geneva.
- Leatherman, S., L. Jones Christensen, and J. Holtz. 2012.** Innovations and barriers in health microinsurance. In: *Protecting the poor: A microinsurance compendium*. Vol. II, ed. C. Churchill and M. Matul, 132-155. Geneva: International Labour Organization.
- Lei, X. and W. Lin. 2009.** The new cooperative medical scheme in rural China: Does more coverage mean more service and better health? *Health Economics* 18:S25-S46.
- LeRoy, P. and J. Holtz. 2012.** Third-party payment mechanisms in health microinsurance. In *Protecting the Poor: A Microinsurance Compendium*. Vol. II, ed. Churchill, C. and M. Matul, 112-131. Geneva: International Labour Organization.
- Morsink, K. and P. Geurts. 2012.** The trusted neighbour effect: Local experience and demand for microinsurance. Institute of Governance Studies (IGS) Working Paper. The Netherlands: University of Twente.

Msuya, J. M., J. P. Jütting, and A. Asfaw. 2004. Impacts of community health insurance programs on health care provision in rural Tanzania. ZEF—Discussion Papers on Development Policy No. 82. Centre for Development Research (ZEF), University of Bonn, Germany.

National Bureau of Statistics, Federal Republic of Nigeria. Nigeria General Household Survey (GHS), Panel 2010, Ref. NGA_2010_GHS_v02_M. Dataset downloaded from (<http://go.worldbank.org/PPN1H3RM10> on 08/10/2013).

National Statistics Directorate Caicoli, Dili, Timor Leste and World Bank 2001: LSMS Timor-Leste, Household Survey - Individual Survey. Dataset downloaded from (<http://go.worldbank.org/N4HTRNUWS0> on 08/10/2013).

Polonsky, J., D. Balabanova, B. McPake, T. Poletti, S. Vyas, O. Ghazaryan, and M. K. Yanni. 2009. Equity in community health insurance schemes: Evidence and lessons from Armenia. *Health Policy and Planning* 24(3): 209-216.

Radermacher, R., H. McGowan, and S. Dercon. 2012. What is the impact of microinsurance? In *Protecting the poor: A microinsurance compendium*. Vol. II, ed. C. Churchill and M. Matul, 59-82. Geneva: ILO. 59-82

Rosenzweig, M.R. and H.P. Binswanger. 1993. Wealth, weather risk and the composition and profitability of agricultural investments. *Economic Journal* 103(416): 56-78.

Schneider, P. and F. Diop. 2001. Synopsis of results on the impact of community-based health insurance on financial accessibility to health care in Rwanda. HNP discussion paper series. Washington, D.C.: World Bank.

Wagstaff, A. and M. Pradhan. 2005. Health insurance impacts on health and nonmedical consumption in a developing country. World Bank Policy Research Working Paper 3563, Washington, D.C.

Wagstaff, A., M. Lindelow, G. Jun, X. Ling, and Q. Juncheng. 2009. Extending health insurance to the rural population: An impact evaluation of China's New Cooperative Medical Scheme. *Journal of Health Economics* 28(1): 1-19.

Wang, H. W. Yip, L. Zhang, and W. C. Hsiao. 2009. The impact of rural mutual health care on access to care: Evaluation of a social experiment in rural China. Working paper. Cambridge, MA: Harvard School of Public Health Program in Health Care Financing.