Annex 1: Dissemination and audience of the review

Demand for this review has come primarily from the Department for International Development (DFID). However, several other donor agencies and the Donor Committee for Enterprise Development (DCED) are investigating ways of measuring women's economic empowerment. As such, there is likely to be widespread interest in the results of the review.

There has been increasing recognition of the importance of women and girls' participation in development over the past two to three years, not least because of the leadership of the British Government internationally. The post-2015 framework is likely to have some important goals for women and girls. One of the four pillars of the DFID Strategic Vision for Women and Girls is related to economic empowerment. It is therefore likely that private sector and economic growth projects and programmes will need to demonstrate the outcomes they are achieving for women and girls' economic empowerment. Likewise, projects to empower women and girls will need to address economic security, livelihoods and enterprise. So we are expecting a wide audience for this paper.

Audiences

Within DFID: The topic of women and girls' economic empowerment spans two different sectors in DFID – the social sector and the infrastructure and economic growth sector – and this is similar for other agencies. So it can be assumed that one primary audience for the review document is private sector, infrastructure, social development and economic advisers. The other primary audience is DFID's evaluation department, which may provide advice and guidance to the rest of DFID on this topic and will need practical guidance to be able to facilitate this. As such, this report aims to be a tool to inform the design and evaluation of programmes that span different thematic areas and departments, with the aim of advancing women and girls' economic empowerment.

External audience: This will consist mainly of advisers and experts in other donor and international agencies, such as the European Union (EU), the Australian Agency for International Development (AusAID), and the African Union, all of which have gender units. It is also likely to include non-government organisations (NGOs) working in this sector, internationally recognised research institutions and think tanks such as the Economic and Social Research Council (ESRC), the Institute of Development Studies (IDS), and developing country governments. Primarily, some donors will be reached through the following:

- OECD GENDERNET
- Donor Committee for Enterprise Development

Civil society networks will also be reached, for example, through:

- The Gender and Development (GAD) Network (UK)
- Bond

Messages

The recommendations section of this report is designed to be used for developing evaluation approaches for WGEE projects, or general economic empowerment projects involving women. The main messages to be communicated are related to the following topics:

- The importance of using gender expertise to inform the design and analysis of the evaluation
- Analysis of gender context alongside political, economic, social and cultural context

- Development / use of a WGEE theory of change and holistic approaches
- Use of participatory methods for developing methodology and indicators
- Mixed methods as the ideal methodological approach, with recommendations about when to use quantitative and qualitative methods separately, and how to combine these
- Finding ways of involving women and presenting their voices in the research
- Value for money – using evaluation findings to improve results for women and girls.

Annex 2: Theory of change for women and girls economic empowerment

INPUTS — OUTPUTS → SHORT TERM CHANGE → MEDIUM CHANGE → LONG TERM CHANGE

INPUTS

Funds for programmes to build financial services and business support to women owned MSMEs (formal and informal)

Funds and programmes to increase women and girls' fair employment/decent work

Funds and programmes for basic or subsistence livelihoods and minimum economic security for the poorest and most vulnerable women and girls

OUTPUTS

Cross cutting action to address barriers – linked SRHR, justice, education services, action to address gender inequality in social norms and behaviour and in institutions, women's leadership development, time saving technology, child care

SHORT TERM CHANGE

- Individual capability: financial services, business development services, physical asset provision, business and financial literacy skills training
- Communities and institutions: access to markets, ownership rights (social norms), gender equality in organisational development
- Political and legal environment: business regulatory framework (e.g. registration), ownership rights (e.g. land, divorce and inheritance laws that affect women and girls)

.

- Individual capability: employment related financial literacy and skills training and fair pay, conditions, benefits etc
- Communities and institutions: employment unions, cooperatives and collective action, support norms to enable women and girls equality in the workplace, female mobility
- Political and legal environment: employment legal and regulatory frameworks

- Individual: social protection, subsistence income generation (mostly informal such as street trading)
- Communities and institutions: support norms to enable women to control income and expenditure
- Laws and policies that protect women and girls in vulnerable situations

Outputs related to addressing barriers – e.g. links to appropriate SRH services and counselling, networks and technology for time saving and childcare.

MEDIUM CHANGE

Power within: Women are able to access and use services to increase their ability to start and grow businesses; to access markets; to increase employment; and to improve their economic security

Power to: Men and boys and community notables become more aware of gender equality and the benefits of women and girls' enterprise, employment and owning assets

Power over: women and girls gain some control of assets, income and household resources. Women control their fertility and access health services.

Power with: women and girls are part of supportive networks and are more respected in their communities

Formal institutions: action to mainstream gender equality within institutions starts to change ways of doing business and result in products and services more adequately designed for women and girls.

LONG TERM CHANGE

Power within: women and girl's self esteem, knowledge and capability is enhanced. They know more about business, finance, their rights to security and justice and SRHR.

Power to: Women and girls mobility improves; household tasks are shared with men and boys. Women use financial services, employment income and cash transfers and increase their economic security.

Power over: Women and girls have employment, take on leadership roles, own and grow enterprises and control cash transfers

Power with: Community norms and behaviour support women and girls in their efforts to improve income, ownership of assets and

Formal institutions, laws and policies reflect gender equality and are more favourable to women and girls' employment, enterprise and economic security

Impact on women and girls' power and agency and on institutions and social norms

↕

Impact on economic advancement – incomes increase (grants and other); more and better quality employment for women and girls; women owned enterprise start up and growth.

Annex 3: Detailed review methodology

The methodology used for this review has drawn on and adapted systematic review principles in order to ensure a comprehensive search of the literature to identify as many relevant resources as possible. Since the objective of the review is to explore evaluation methodologies and approaches rather than findings (which are generally the focus of reviews), the tools used at every stage were adapted to capture information needed to answer the research questions. The description of the methodology presented below provides details on how the review process was carried out. The scoping and inception reports developed as earlier outputs presented the planned methodology for subsequent stages, which was adapted based on comments received from DFID, the panel of experts and the reference group; it was also based on internal reflections within the team about the best possible way to search, select and review relevant evaluation and research reports that contain useful information on methodological approaches to explore women and girls' economic empowerment.

The review uses a broad definition of 'evaluation' and includes research designs that may not explicitly be described as evaluations but aim to show the impact or effects of a project/programme/policy (an 'intervention') on women's and/or girls' economic empowerment. As such, in addition to evaluations, a number of research reports, appraisals and assessments were included in the database. These used a variety of methodological approaches, including but not limited to: qualitative and participatory evaluations; post-hoc evaluations; quasi-experimental project designs, and randomised control trials (RCTs). This decision to look beyond evaluations that were 'strictly defined' as such was made in consultation with DFID, as part of the aim was to find interesting and innovative methodologies that are being used; research reports and other types of assessments use similar methodological approaches to those used by evaluations, and so they were seen as relevant to include.

1.1 Methodology used for the review

The review process consisted of four stages, which are described below.

1.1.1 Stage 1: Initial search of the literature

The search strategy for this review was designed to identify evidence and evaluations from across disciplines that report outcomes for women and girls' economic empowerment. This includes evaluations that are not explicitly targeting women and girls' economic empowerment, but have wider economic empowerment and wealth creation objectives. The search was designed to take place in both the grey and academic literature and did not privilege any particular discipline. However, the search was influenced by the selection of thematic areas under which the search took place:

1) **Financial services**
2) **Business development services**
3) **Skills training**
4) **Asset provision**
5) **Social protection**
6) **Unions**
7) **Trade and access to markets**
8) **Regulatory and legal**

The systematic search process is based on the principles of transparency and replicability. The strategy was largely based on web searching and as such is dependent on the exclusion of irrelevant sources by the researcher at the point of search. Four research assistants were trained in the search methodology and in the use of Evidence for Policy and Practice Information (EPPI) Reviewer database management software to conduct the search over a period of three months.

Inclusion criteria

Researchers included studies based on clear criteria, as shown in Table 1. These criteria are adapted from the PICO (population, intervention, comparator, and outcomes) criteria. They ensure that uploaded sources are relevant to the review. The wide field of search was narrowed by focusing on English-language resources only, which also accommodated the linguistic capacity of the research team.

Table 1: Inclusion criteria

Language	Is the study in English?	If not, exclude – If the paper is in French or Spanish upload and code with section one only
Location	Does the evaluation take place in at least one lower-middle-income country (LMIC)[1]	If not, exclude
Design	Is the paper an evaluation or does it include discussion of evaluations? (Including assessments, analysis of effects, mid- or end-of-year reviews, systematic reviews, etc)	If not, exclude
Outcomes	Does the paper include a measurement of women's empowerment	If not, exclude
Publication date	Was the paper published after 1990?	If not, exclude
Publication type	Is the document a PhD or Master's thesis?	If yes, exclude

Studies which met all the criteria were uploaded into the EPPI Reviewer database and coded by the researcher. During the research process, the team kept detailed research notes and held regular meetings to discuss progress and challenges and to ensure a consistent approach. The project manager (PM) and technical lead (TL) undertook regular spot checks to monitor the quality of the coding and uploading of information. Feedback to researchers was used to improve accuracy of coding and to iron out differences between the approaches of individual researchers.

Search terms

Searches were planned in a methodical way by using a table of terms for each of the thematic areas. In each thematic area, in addition to searching for specific terms relevant to that theme, word searches combined the following words to identify resources relevant to our review:

Table 2: Key search terms used

Women	Economic empowerment	Evaluation
Woman	empower*	"Impact assessment"
Girl	Access	Evidence
Girls	Income	Results
Female	"Income generation"	Impact
Adolescent	Control	Assessment
	Wellbeing	Study

The table below shows the detailed search structure and search terms used

Table 3: Search Combinations Completed

Google and Google scholar

Microcredit and Financial Services – combinations of				
Women	Microcredit	Economic empowerment	Evaluation	
Woman	Microfinance	empower*	"Impact assessment"	
Girl	Microsavings	Access	Evidence	
Girls	Microlending	Income	Results	

[1] LMIC as defined by the World Bank http://data.worldbank.org/about/country-classifications/country-and-lending-groups

Female	microinsurance	"Income generation"	Impact	
Adolescent		Control		
		Wellbeing		
Trade and Access to Markets				
Women	Access to Markets	Economic empowerment	Assessment	
Woman	Market Access	Empower*	Evaluation	
Girl*	Value Chain		Review	
Gender	Fair trade		Study	
Adolescent	Value chain		Mixed method	
	Ethical trade		Case study	
	Trade agreement		Ethnography	
	Trade regulation		Action research	
	Market* for poor		Operational research	
	M4P			
	“Making markets work for the poor”			
Unions and Fair Employment				
Women	Union	Empower*	Evaluation	
Woman	Trade union	Economic empowerment	Review	
Girl*	Labour	Income generation	Study	
Gender	labor			
Adolescent	Collective Action			
Business Development				
Women	Business development	Empower*	Evaluation	
Woman	Market access	Economic empowerment	Review	
Girl*	Value chain	Income generation	Assessment	
Gender	Enterprise		“Performance review”	
Adolescent	Cooperative		“Performance assessment	
	M4P			
	Business Structure			
	Entrepreneurial			
Skill development				
Women	Skill*	Economic empowerment	Evaluation	* In some searches Afghanistan, OR Gambia, OR Myanmar, OR Bangladesh, OR Guinea, OR Nepal, OR Benin, OR Guinea-Bissau, OR Niger, OR Burkina OR Faso, OR Haiti, OR Rwanda OR Burundi, OR Kenya, OR Sierra OR Leone, OR Cambodia, OR Korea,
Woman	Skill training	Employment	Programme	
Girl*	Vocational	Power	Assessment	
Gender	Employment	Capability	Impact	
Adolescent	“Financial literacy”	Income	Evidence	
	Vocational education	Confidence	Qualitative	
	Vocational training	Agency	Quantitative	
	Industry related skills		Meta-evaluation	

	Livelihood		Impact	OR Somalia, OR Kyrgyz OR Tajikistan, OR Chad, OR Liberia, OR Tanzania where used
	Education		Mixed method	
	Literacy		Experiment	
	Train		Quasi experiment	
			Randomised	
			Randomised	
			Participatory	
			Case study	
			Longitudinal study	
			Operational research	
			Action research	
			Review	
			Report	
Assets				
Women	Asset Provision	Economic empowerment	Evaluation	
Woman	Physical asset	Capability	Programme	
Girl*	Asset accumulation	Resilience	Impact assessment	
Gender	Productive Asset	Agency	Mixed method	
Adolescent	Savings	Power	Meta evaluation	
	Self help group	Empower	Mixed method	
		Expenditure/spend	Experiment	
		Capability	Quasi experiment	
		Confidence	Randomised	
			Participatory	
			Action research	
			Case study	
			Longitudinal study	
			Report	
			Ethnography	
			Audit	
			Review	
Social Protection				
Women	Social protection	Economic empowerment	Evaluation	
Woman	Social Assistance	Power	Meta evaluation	
Girl*	Cash Transfer	Empower	Evidence	
Gender	Public works	Expenditure	Impact assessment	
Adolescent	Employment guarantee	Spend	Mixed method	
	Childcare	Capability	Experiment	
	Crèche	Confidence	Quasi-experiment	
	Community care	Awareness	Randomised	
	Nursery	Income	Participatory research	
	Pension	Well-being	Action research	

	Grant	Inclusion	Case study	
	Social security	Opportunity	Longitudinal study	
	Safety net		Operational	
			Review	
			Report	
			Ethnography	
			Audit	
Legal and Regulatory Frameworks				
Women	Law reform	Economic Empowerment	Evaluation	
Woman	Policy reform	Empower	Assessment	
Girl*	Policy change	Income generation	Review	
Gender	"Inheritance law" OR "inheritance right" OR "Intergenerational Transfer	Bargain	Study	
Adolescent	"land ownership" OR "land right*" OR "property right" OR "property ownership	Spend	Research	
	("asset ownership regulation" OR "asset regulation"		Case study	
	("business start-up*" OR "business start up*" OR "enterprise start up*" OR "enterprise start-up*")			
	tax concession*" OR "tax regulation*			
	trade regulation*"			
	policy change*" OR "policy reform*" OR "law reform*			

Table 4: Hand-search Combinations

Women	Economic empowerment	Impact
	Decision making	Intervention
	Economic decision making	Evaluation
	Physical assets	
	Productive assets	
	Commitment savings	
	Social safety net	
	Social transfer	

*Note: The word combinations used on hand searches were different to the google searches as in each institutional website, we were looking across all themes, at evaluations that were looking at economic empowerment outcomes of women and girls.

Coding

The coding structure designed for this review operated on two levels. A basic level of initial coding was used for all resources identified and which met the inclusion criteria. These included resources that were not empirical evaluations or studies, but were either relevant overviews or systematic reviews that the researchers felt would be useful for snowballing (pursuing references of references). More in-depth coding was applied to papers which were considered as primary empirical research and potentially relevant for answering the review questions. In total, 255 documents were coded with the complete tool, which captured information about the commissioning of the paper, the methodological design, and project outcomes.

Search locations

The search combined keyword searches in Google and Google Scholar, academic database searches, hand-searching of organisational websites, expert recommendations, and snowballing. The tables below provide details about the organisations whose websites were hand-searched, the agencies contacted by ODI during the search process to request evaluations, and names of those who sent evaluations to be considered using the initial inclusion criteria.

Table 5: Organisations Hand searched

Organisations
Care International
UNESCP
Oxfam
Girl Hub
ICRW
UNDP
IDRC
Governance and Social Development Resource Centre (GSDRC)
EPPI Centre
GADE Gender
SOAS Centre for Development Policy and Research
Indian Institute for International Development
Microinsurance Centre
BRAC
UNDP
3ie
Microcredit Summit
Business Fight Poverty Website
Microinsurance Gateway
ICRW
Portals and Databases
Gender equality evaluation portal
Eldis
M4P Hub
Gender equality evaluation portal
R4D
DAC Evaluations Database
International Gateway for Financial Education
Donors
DFID
Danida
Swiss Agency for Development and Cooperation
Global Fund for Women
USAID

Table 6: Organisations contacted by ODI for evaluations

AGENCY	CONTACT PERSON(S)
Donors:	
Swedish Development Cooperation (SIDA)	Cecilia Bisgen Jansson,
	Maria Melbing
AusAID	Wally Cook <Wally.Cook@ausaid.gov.au>
DANIDA	Margrethe Holm Andersen
Multilaterals:	
UN Women and Fund for Gender Equality	Lucia Hanmer
	Claudia Briones
UNICEF	Richard Morgan
	Judith Diers, Development and Participation Section
International Labour Organisation (ILO)	Celine Peyron Bista <bista@ilo.org>
	Joni Simpson
World Bank	Ana Maria Munoz Boudet
	Markus Goldstein
ADB	Sri Wening Handayani
IFAD	Clare Bishop Sambrook
Academia/think tanks:	
Institute of Development Studies (IDS)	Keetie Roelen
Overseas Development Institute (ODI)	Gender workstream
International Centre for Research on Women (ICRW)	Ann Warner, Senior Gender and Youth Specialist
School of Oriental and African Studies (SOAS)	Naila Kabeer
International Food Policy Research Institute (IFPRI)	Agnes Quisumbing
United Nations Research Institute for Social Development (UNRISD)	Sarah Cook
UCL	Maxine Molyneux
Oxford University	Steven Gillenwater
	Jo Boyden, Director of Young Lives
University of Minnesota	Ragui Assaad, Professor of Planning and Public Affairs
Private sector:	
Nike Foundation	Nina Church Adams
	Yohannes Wolday
	Amy Babchek, Pillar Lead
	McDougall, Janna
	Myrum, Jodi
	Maria Eitel (director)
Gates Foundation	Haven Lay
Oak Foundation	Nathalie Chambeyron
	Florence Bruce, Director of the Child Abuse Program
Novo Foundation	Pamela Shifman, Director of Initiatives for Girls and Women

NGOs, Civil Society etc:	
ActionAid	Nicola Curtin and Zohra Moosa
Plan International	Janis Ridsdel
Oxfam	Thalia Kidder
	Kent Glenzer (formerly with CARE US, now OXFAM US)
Clinton Initiative	
Population Council	Andrea Eschen
	Judith Bruce
	Ann Blanc, Vice President & Director of Poverty, Gender and Youth Program
	Kelly Hallman
	Shireen Jejeebhoy,
Womankind	Emil Esplen
Plan UK	Goulds, Sharon <Sharon.Goulds@plan-uk.org>
CARE	Pelham, Larissa <Pelham@careinternational.org>
	Tom Aston <tmtaston@hotmail.co.uk>
	Dan Vexler, CARE UK
HIVOS	Catherine van der Wees
Girl Hub	Ellen Wratten (director)
	Kristen Woolf - girl expert director
IRC	Jeannie, Annan Director of Research & Evaluation
Women's Refugee Commission	Dale Bushcar - senior director of programmes
BRAC	Karishm Hudaa - Former Director of Research

Table 7: Documents were received from the following:

Deutsches Institut für Entwicklungspolitik (DIE) / German Development Institute	Dr. Aimée Hampel-Milagrosa
3IE	Annette Brown
Finmark Trust	Andrea van der Westhuizen
University of Warwick	Christopher Woodruff
Danida	Margrethe Holm
J-PAL	Caitlin Tulloch
Independent Commission for Aid Impact	Tom McDonald
Save the Children UK	Nadja Dolata
World Bank, Africa Region Gender Practice	Katherine Manchester
Swiss Agency for Development and Cooperation **(SDC)**	Anne Bichsel
Ministry of Foreign Affairs, The Netherlands	Paul de Nooijer
Gender Links, South Africa	Colleen Lowe Morna
Plan International, Economic Security Adviser	Karen Moore
Women in Informal Employment: Globalizing and Organizing	Mike Bird
UN Women	Inga Sniukaite
UNRISD	Sarah Cook
UCL	Maxine Molyneaux

OECD	Angelica Salvi Del Pero
Actionaid	Rachel Moussie
ILO, Social Protection	Celine Peyron Bista
AUSAID	Anna Clancy
NIKE	Yohannes Wolday
Spanish Ministry of Foreign Affairs	Paula Caballero Partido
OECD	Somalo Cerise
HIVOS	Marjan van Es
AUSAID	Jacqui Thomson
IRC	Amanda Sim
Freedom from Hunger	Bobbi Gray

A total of 382 papers were uploaded into the database for review, 254 of which were empirical research and were subsequently analysed for quality in the next stage. The team reached 'saturation point' based on the search methodology, drawing on web searches (Google, Google scholar, academic databases, specific journals and selected grey literature sites which included evaluation portals) as well as liaising with experts and relevant contacts who sent us useful, often unpublished evaluations.

After the search was concluded, an inception report was submitted to DFID, which captured trends and statistics of the whole set of evaluations and research documents. Some of these statistics have since been amended as a result of further, more in-depth work with the database, which resulted in adjustments being made to coding. The statistics presented in Section 3 are therefore more accurate.

1.1.2 Stage 2: Narrowing the field of inquiry

In order to narrow down the field of inquiry from the 254 empirical reports included, the research team devised a two-stage process using two different tools. Given that our interest was mainly to analyse methodological approaches in the reports identified to obtain information on different dimensions of women and girls' economic empowerment, the first 'basic scoring tool' captured the availability of information on methodological approaches (see Annex 3). Reports that scored 2 or more points on that tool were then assessed using a purpose-designed 'quality assurance' tool, based on the principles of the Mixed Methods Appraisal Tool (MMAT) developed by Pluye and colleagues (2011), which is designed for the appraisal stage of systematic reviews that include qualitative, quantitative and mixed methods studies.

The 'adapted MMAT tool' used for screening (see Annex 4) was designed to give a rapid assessment of 'quality' based on basic criteria set out for each methodological approach used. Expert panel members commented on the adapted MMAT tool before it was utilised and made suggestions on which variables to include to capture key dimensions of methodological 'quality / content'. This means that it was able to provide a light touch analysis through a rapid review of the documents, enabling further in-depth screening. Reports that scored 2 or more points were considered for the review. In total, 71 documents 'passed' the quality assurance test.

1.1.3 Stage 3: Review

Out of the 70 documents shortlisted, there was at least one in each of the eight thematic areas, although some were better represented than others, as Table 3 indicates.

Table 8: Distribution of shortlisted evaluations across the eight thematic areas

Microfinance	Social protection	Business development	Trade	Asset transfer	Skill development	Regulatory frameworks	Unions
36	8	8	8	2	7	1	1

There was a distribution of shortlisted reports and papers across the different methodological approaches, but a majority used quantitative (32) and mixed methods (26), with only 13 using qualitative methods alone. This lower representation of qualitative methods (despite a more balanced representation in the overall database) might result from the fact that qualitative evaluations and studies tend to have much less detail about their research methodology than quantitative reports, and thus might have been excluded in stage 2 on this basis. Still, there was a good representation of interventions using qualitative methods through the mixed methods evaluations and studies.

The review stage started with the development of a 'review template'. The questions in the template were based on a document review of several evaluation best practice documents, as well as on key resources on women's economic empowerment – for example, the Development Assistance Committee (DAC) criteria for evaluating development assistance[2], the Bond quality of evaluation checklist, DFID's evaluation policy (DFID, 2013), and Kabeer's (1999) paper on Reflections on Women's Economic Empowerment. The purpose of the review template was to enable the expert panel to assess the full documents to make a robust analysis of the quality of methodologies used, particularly in relation to their ability to capture relevant information about the effects of the intervention on women and girls' economic empowerment.

Before starting the review process, there was a 'pilot' in which experts reviewed the same three reports to jointly discuss and exchange ideas about how women and girls' economic empowerment was understood, what aspects of the methodology were more relevant to capture in the review, and to ensure some consistency in the review approach. After this pilot, DFID and the experts commented on the review template, and on this basis, a final version of the template (see Annex 5) was circulated to reviewers. The template required experts to 'score' the report's methodology using a scale of 1 (highest) to 5 (lowest), and provide a separate score for the methodology used to assess economic empowerment, if different. It also required them to provide an analysis of the intervention's strengths, weaknesses, appropriateness, understanding of economic empowerment of women and/or girls, and the contribution it makes to the literature.

Each reviewer was allocated between 15 and 20 documents to look at, based on the days they had available to work on the project, and the number of reports shortlisted in their area of thematic/methodological expertise. For example, one of the panel members has strong expertise in microfinance, so she was allocated only microfinance papers using a variety of methodologies. Another panel member is a quantitative evaluations specialist, so he was assigned the majority of the quantitative evaluations across the different thematic areas. At the end of the exercise, we asked two of the reviewers to look at two other reports that were assessed by another expert initially in order for us to identify how their different backgrounds could influence their analysis in the review. In addition to completing the reviews and submitting the templates, each expert sent a synthesis of ideas based on the documents reviewed highlighting particularly good (or bad) evaluations and identifying patterns or issues that arose and could be generalised from the review. The individual reviews and the synthesis provided by each expert form the basis for the analysis and recommendations presented in the following sections of this report.

In addition to the review templates, a research assistant prepared factsheets with basic information on the 15 highest scoring reports. This was so that the reader could get a good sense of the evaluation (particularly its methodology and findings) by looking at the factsheet and review template together, without needing to read the full report.

1.1.4 Reflections on the expert panel review process

The four experts in the panel each have significant experience of issues of gender equity and women's empowerment, as well as collectively having robust knowledge of quantitative and qualitative research methodologies. By allocating each expert a set of documents within their area of expertise, and with the help of the review templates, it was possible to consolidate valuable information regarding strengths and weaknesses of the selected reports.

Nevertheless, when assessing the quality of a report, panel members were influenced by their individual expertise, knowledge base and methodological preference. While the pilot at the start of stage three was designed to maximise consistency in the review approach, each expert provided insights that were informed by their own understanding of WGEE and their perceptions of the strengths of different evaluation methodologies.

[2] http://www.oecd.org/dac/evaluation/daccriteriaforevaluatingdevelopmentassistance.htm

Section 4 presents a balanced analysis based on careful reading of all the individual reviews and the analytical syntheses prepared by each expert. For example, while three of the experts highlighted the importance of baseline information to be able to assess changes as a result of a specific intervention on a selected population group, another expert noted that baselines were not necessary to analyse causality, emphasising the value of well-designed historical tools and triangulation, which could enable credible recall of the factors contributing to change. The report explains when each approach might have value, and while one of the recommendations is for interventions to have baseline data, it is also important to acknowledge that other tools can help determine causality. Still, despite the different perspectives, we found a consistent 'scoring' of quality across all four experts, indicating that their individual expertise, knowledge base and methodological preference did not have an impact on the assessment scores they gave.

The divergence in the way the 'quality' of a methodological approach is assessed is a useful indicator of the challenges faced by organisations commissioning evaluations to explore WEE in terms of prioritising specific research designs. Notwithstanding these discrepancies, cross-checking a small number of reviews with more than one panel member resulted in similar scores for the WGEE methodological approach, even if specific answers to questions in the template were different. This indicates that the rigour and robustness of evaluation quality can be assessed objectively, even if there are nuances in the interpretation of the definition of WGEE with respect to specific context and when the perception of the usefulness of specific methodologies can diverge.

1.2 Lessons learned and suggestions for future review methodologies

The review process was based on prior experiences within the team conducting adapted systematic reviews. The search phase was designed on the basis of prior reviews. However, this exercise was distinct in that the focus is on methodological approaches rather than findings, which should have resulted in more substantive adjustments during stage one. For example, when the team got to the second stage, there was a sense that the work to be done would have been simplified if an indicator for availability of methodological information had been included as part of the inclusion criteria in stage one. While this would not have taken much additional time in stage one, it was a time-consuming process during stage two, as it required a specific screening tool. This diverted time away from the quality assurance process in stage two. If there had been more time, this process could have been more analytical.

A useful adjustment during stage two to respond to the nature of the search would have been to limit the inclusion of journal articles to those that had detailed description of methodologies, or where possible, looking for original reports to which the journal articles were linked. Though some very interesting papers reviewed were journal articles, reviewers noted that in general these did not have enough details about the methodology on which to make a robust assessment, and they would have liked to have seen the original reports. Since most reports linked to articles were not available through the web searches conducted, additional time would have been needed to contact authors to request these reports.

Another lesson learned from this process, if the same methodology were to be followed for subsequent reviews, would be to build in more senior researcher time (and resources) for spot checks, and more days for experts to review selected reports, particularly to enable more than one expert to review all the reports. In total we were able to have 4 out of the 71 papers reviewed by two experts, which brought interesting analytical insights. This would have been more insightful if done with all the papers. In particular, having quantitative experts look at qualitative reports and vice versa gives entirely different perspectives about how evaluations or studies are designed and commissioned.

It is important to note that while this rapid review process was useful and identified very valuable resources, the experts on the panel observed that a few of the 'included' documents had poor quality in their use of methodology. Thus, the adapted MMAT tool was able to capture whether reports had certain information related to their methodology, but not whether the methodology was used robustly enough or in a technically accurate way. It was only through the more comprehensive assessment of each review by the experts who read the reports in greater detail that it was possible to make a solid assessment of quality. Similarly, by excluding reports from stage two based on the detail of the methodology, it is possible that the review might have excluded relevant papers with useful findings because they might not have reported their methodology well (even if the methodology used was robust).

The work of well-trained research assistants, with knowledge of the review methodology as well as gender empowerment, was critical to be able to distil good evaluations in a well-populated field of research. In addition to this, throughout the process, spot checks by the senior team were useful and relevant.

Despite these minor limitations on the search process, given the systematic method used during the search, the documents reviewed are a thorough and good representation of the work that has been done in this field.

Annex 4: Basic scoring tool

Title				
Author	**Year**			
Document content quality criteria	**Responses**			
	Plenty	**Some**	**None**	**Comments**
	1	**0.5**	**0**	
SCORING				
• Are there clear qualitative and quantitative research questions (or objectives) or a clear mixed methods question (objective)				
• Are there details of interview guides, questionnaires or tools?				
• Is there information on sampling (sample size for quant, sample design for qual (what types of respondents, selection process)				
• Is there information on data collection methods? (e.g, research methods used– survey, Focus groups, interviews etc)				
• Is data analysis described?				
TOTAL SCORE				
INFORMATION ONLY – Please score and provide notes				
• Is research communication, reporting, dissemination and/or use described?				
• Is this an evaluation (based on DAC definition)? •				

Annex 5: Quality assurance tool (adapted 'MMAT' tool)

Types of mixed methods study components or primary studies	Methodological quality criteria	Responses			
		Yes	No	Can't tell	Comments
Screening questions	• Are there clear qualitative and quantitative research questions (or objectives) or a clear mixed methods question (objective)				
	• Are there details of lower level research questions, interview guides, questionnaires or tools?				
	• Do the collected data address the research question (objective)				
	• Is there any disaggregation by sex or analysis of results for women and girls? (either for economic advancement or agency/empowerment)				
Further appraisal may not be feasible if the answer is no or can't tell to one or all of the screening questions					
Qualitative	Are the sources of qualitative data (archives, documents, informants, interviewees / research participants, observations) relevant to address the research question or objective of the research??				
	Is the process for collecting and analyzing qualitative data relevant to address the research (objective)?				
	Is appropriate consideration given to how findings relate to the context – where the research was carried out? Or the socio-cultural groups with which the research took place?				
	Is appropriate consideration given to how researchers' profiles might influence their interactions with participants? (e.g. gender, nationality, ethnicity, language, etc.)				
	Is there a description of the sampling design and sample selection process?				
Quantitative					
Randomised	Is there a clear description of the randomisation?				

controlled trials					
	Is there a clear rationale for doing a RCT in relation to the research question?				
	Does the study provide detailed description of the methodology used to control for confounding and selection bias? (including to minimise contamination between groups)				
	Is there low non-response rate? (below 20%)				
	Is there a test/table comparing the control and treatment groups?				
	Are fewer than 10% of the observables statistically significantly different from each other?				
Quantitative non-randomised	Are participants (organisations) recruited in a way that minimises selection bias?				
	Are measurements appropriate (clear origin, standard instrument, strategy to avoid contamination between groups) regarding the intervention and outcomes?				
	In the groups being compared are the participants comparable, or do the researchers take into account (control for?) the differences between these groups?				
	Are there complete outcome data (80% or above) and either an acceptable response rate (60%) or above, or an acceptable follow-up rate for cohort studies (depending on the duration of follow up)				
	Are the assumptions and methodology clear? If the method is not standard do they discuss the identification technique?				
Quantitative Descriptive	Is the sampling strategy relevant to address the quantitative research question (quantitative aspects of a mixed methods study)				
	Is the sample representative of the population under study				
	Are measurements appropriate (clear origin/ validity known/standard instrument				

	Is there an acceptable response rate (60% or above?)				
Mixed Methods	Is the rationale for integrating qualitative and quantitative methods to answer the research question explained				
	Is the integration of qualitative and quantitative data (or results) done effectively (triangulation or informing methodology) to address the research question (objective)				
	Is appropriate consideration given to the limitations associated with this integration e.g. the divergence of qualitative and quantitative data (or results) in a triangulation design?				

Annex 6: Expert panel members' profiles

Nicola Jones, ODI Research Fellow

Nicola has a PhD in Political Science and is a Senior Research Fellow in the Social Development Programme at ODI and the institute's Gender Theme coordinator. She has carried out policy research, mixed methods evaluations, advisory and capacity building work for a range of funders (including ADB, AusAID, DFID, EU, FAO, Gates Foundation, GAVI Alliance, IDRC, Netherlands Organisation for Scientific Research, Oak Foundation, Oxfam, Plan International, Save the Children, UNICEF, UNDP, UNFPA, UN Women) across diverse developing country contexts. She has published extensively for academic, policy and practitioner audiences, including six books and twenty peer reviewed journal articles, reports, briefing papers, blogs, toolkits and podcasts. In the past five years, Nicola has led a number of multi-country evaluations on a range of gender-related issues; including on gender-responsive social protection, women's reproductive and sexual health, girls' and women's economic empowerment in the context of global crises, and the role of social norms in shaping girls' and young women's experience of chronic poverty. Her evaluation expertise lies in qualitative and participatory evaluation approaches, as well as the application of mixed quantitative and qualitative data.

Patrick Nolen, Professor (tenured), university of Essex

Dr Patrick Nolen has nearly ten years of experience in gender, development, education, and impact evaluations. He is an expert on quantitative methods and has worked on evaluations of school meal programs, educational achievement, and the rollout of cellular networks. As an advisor to the Vice-President's Office of Ghana he led on the creation of monitoring and evaluation framework for the Savannah Accelerate Development Authority (SADA). In working with DFID he has been responsible for reviewing the quality of impact evaluations supported by DFID and has served as an advisor on the monitoring and evaluation strategy for the DFID funded Millennium Village. He has reviewed grant applications of funding for systematic reviews. He has also served as external project advisor for 3ie where he has reviewed the quality of the impact evaluations funded by 3ie and as a panel member of the grant review boards for 3ie's open window and policy window applications. He has experience in data collection, analysis and results presentation. He has published in highly ranked peer-reviewed journals such as the *Economic Journal* and *Journal of Economics and Behavior and Organization* and served as a referee for such journals as *Quarterly Journal of Economics* and *American Economic Review*.

Linda Mayoux, Independent Researcher

Linda is a member of the Social Development Direct consultant network and women's economic empowerment specialist with 29 years' experience working with DFID, the World Bank, the United Nations and international NGOs such as Oxfam. She has extensive experience in social, enterprise and institutional development, particularly in sub-Saharan Africa and Southern Asia. She has a sound background in management of community development programmes, micro-finance, participation and participatory development including community led learning, social and economic inclusion and exclusion, social capital, social movements and interlinkages between specific forms of participation and macro-level change. Linda also has extensive experience in gender policy and a sound background in training methodologies and study skills for development. She is a skilled researcher who brings experience of qualitative and quantitative data and evaluation techniques from both academic and consultancy roles.

Caroline Pinder, Independent Researcher

Caroline is a member of the Social Development Direct consultant network who brings extensive experience in social, enterprise and institutional development, particularly in sub-Saharan Africa and Southern Asia, and specialises in gender equality and women's empowerment. She has a sound background in management of community development programmes and commercial and social enterprises in the UK and internationally. She also has extensive experience in policy analysis, strategic planning and monitoring, evaluation and impact assessment of UK-based and international public, private and not-for-profit programmes. In 2000, Caroline founded WISE Development (Women in Sustainable Enterprise Development), a global network of women consultants specialising in gender equality and women's empowerment.

Annex 7: Review of evaluation approaches and methods for economic empowerment of women and girls interventions

Expert panel reporting template

Write into the space under each question or by replacing the questions with your text.

Please answer all questions on the 'compulsory questions' section (1 to 4)

The supplementary questions also provide useful information for the review. Please answer the ones which are relevant for the evaluation reports you are reviewing. You can skip the response or note that the question cannot be answered from the report, as necessary.

Note: The users of this review are people commissioning, managing and/or designing evaluations of EEWG projects/project components, so please bear this in mind when responding to the questions, expanding as much as possible on your answers.

Basic information		
Name of evaluation / report	Authors	Would you describe this as an evaluation or a research report?
Compulsory Questions		
1. Overarching questions (probably easiest to answer these after the other questions) • What is the purpose / objective of the study (in 2 or 3 sentences)? • Based on your review of the text does this evaluation/research use/present robust, **high-quality research methodology relating to the economic empowerment of women and/or girls?** (please answer with regard to the methodology used to assess WG economic empowerment, rather than the methodology in general) • Please grade the overall methodology using the following scale, and describe why you have assigned that grade: 1=Excellent, 2=Good, 3=Satisfactory, 4= Significant weaknesses, 5= Very weak. • Please grade the methodology used particularly to assess Women's and Girls' economic empowerment (only respond in cases where it is different from the general methodology, that is, if this is not the main focus of the evaluation), using the following scale, and describe why you have assigned that grade: 1=Excellent, 2=Good, 3=Satisfactory, 4= Significant weaknesses, 5= Very weak. • What are the main contributions of this report to the methodological toolkit for evaluating the economic empowerment of women and girls? E.g. any aspects that are particularly interesting or i**nnovative** approaches to the evaluation of women and girls' economic empowerment (please develop your answer fully, highlighting in particular if the report contributes to the evaluation literature with any important innovations or where it ignores relevant literature/ knowledge please specify) • Please describe if you find any methodological problems / weaknesses in the evaluation methodology		

used.

- Does the evaluation indicate that there was a relevant analysis / understanding of the context? (household, community, etc)
- Does the evaluation show that there is a relevant understanding of the intervention?
- Does the evaluation provide a hypothesis about how change happens or a theoretical framework?
- Is there an implicit / explicit theory of change? If so, does it contribute in a relevant way to the literature on women's and girls' economic empowerment?
- Does this meet the OECD DAC definition for an evaluation? *(The systematic and objective assessment of an on-going or completed project, programme or policy, its design, implementation and results.)*

2. Methodology

- Do you think this was the best method to use given the type of intervention, context and **objectives** of the evaluation? Why / why not?
- What might have been better or what complementary additional approaches might have been used? (this can mean a different approach within the same methodology (e.g. a different quant model) or a complementary approach).
- Do the authors mention (directly or indirectly) that they would have preferred to use a different methodology, but were limited by resource constraints?
- Has the data been rigorously collected, analysed and reported? How credible are the findings given the data collection, triangulation and analysis?
- Reflect on resources that would have been required (funding, local researchers with certain skills set, gender, etc) to have expanded or improved the methodology (where information is available/ can be gleaned)

3. Gender and empowerment in evaluation design and methodology

- Do you consider the indicators used to measure changes for women and girls at impact and outcome level to be useful, appropriate, and/or innovative? (e.g. are they comprehensive enough to measure empowerment?)
- Have women and girls voices been captured through the methodology used?
- Is transformation in women and girls' agency measured, if so, how?
- If the population being analysed are adolescent girls what, if any, specific considerations are taken in the methodology?
- Have men and/or boys been included in the research and analysis? If so, in what ways? (e.g. male-female relations informing the theory of change; men/boys as participants in the evaluation; assessment of differential impacts of the intervention on men and women, etc)

4. Findings – outcomes and impact

- What do you think of the paper in terms of the broader evaluation literature in that thematic area? Does it add value? If so, how?

- Is the evaluation able to explain how and/or why an intervention contributes to intended and/or unintended, positive and negative effects for beneficiaries? (e.g. causality)

- What are the main findings of the evaluation? Does the evaluation methodology used help uncover interesting findings with respect to women's and girls' economic empowerment

Supplementary questions to consider

(i). Evaluation design

- Is the evaluation design explained in the context of the methodology and possible data collection limitations?

- Is the timing of the evaluation appropriate with respect to the intervention cycle?

(ii). Evaluation methodology

- Does the evaluation measure long term changes that indicate sustainable impacts / effects of the intervention? If so, how?

(iii) Findings and their use (there is very little information on the latter in reports)

- Are results adequately contextualised (in terms of social, cultural, economic, political, programming context)?

- Do conclusions and recommendations clearly follow findings?

- What are the gender and policy dimensions of use, that is, do authors explain how findings are or can be used to inform better programming from a gender and policy perspective? Is there something about the methodology that makes the evaluation more usable (e.g. including key decision makers involved in the development of methodology implementation of the evaluation or a stakeholder group)

(iv). Gender and empowerment understanding and mainstreaming

- Are men and boys considered as actors in the process of economic empowerment of women and girls? Are attitudes, practices and behaviours related to social norms considered? Does the evaluation look at change in these norms?

- Is there an assessment of mainstreaming of gender equality in formal institutions within the project (e.g. in microfinance institutions or cash transfer intermediaries)?

Additional comments (optional)

Please add anything specific you might want to highlight from this evaluation, for example, whether it is particularly good in its understanding of economic empowerment; or whether it is very innovative / added value to the literature in the sector; etc.

Annex 9: Examples of good theories of change exploring WGEE

The Saving for Change project (Mali) impact assessment (Bureau of Applied Research, 2013) presents the following theory of change picture (Figure 1). This project supports women to set up savings groups that can then lend to members of the group. The appeal of this theory of change is that it shows potential causal relationships and the multiple relationships that link up different processes of change in a simple and easy to understand format. It also shows change across several different time frames – which helps to develop a deeper understanding of how change may differ in the short term, medium term and long term. The evaluation text refers to the causal links as 'predictions' of what will happen, but does not explicitly identify assumptions that are underlying the predictions. Nor does it explicitly refer to information or existing evidence that may form the basis of the 'predictions'. And finally, it should be noted that this theory of change has no reference to gender relations or changes that might happen in men's attitudes or behaviour (to support the changes shown here). There are some unidentified underlying assumptions about gender relations that may not have been explored fully (for example, that women owning assets means that they make decisions about them without men's influence).

Figure 1: The Saving for Change project in Mali impact assessment – theory of change

Source: Bureau of Applied Research, 2013

A more extreme example of a theory of change that is gender blind can be found in Blattman's 2013 study 'Credit Constraints, Occupational Choice, and the Process of Development: Long Run Evidence From Cash Transfers In Uganda'. This is a useful and fairly thorough evaluation using randomised control trial (RCT) methodology and it has a rare focus on young people. However, the theory of change looks almost exclusively at economic advancement – which, though of value in itself, does not give any information about transformation in girls' agency or about overall economic empowerment or transformation in gender relations, attitudes and behaviours. This gender blind approach is pervasive throughout the evaluation and analysis.

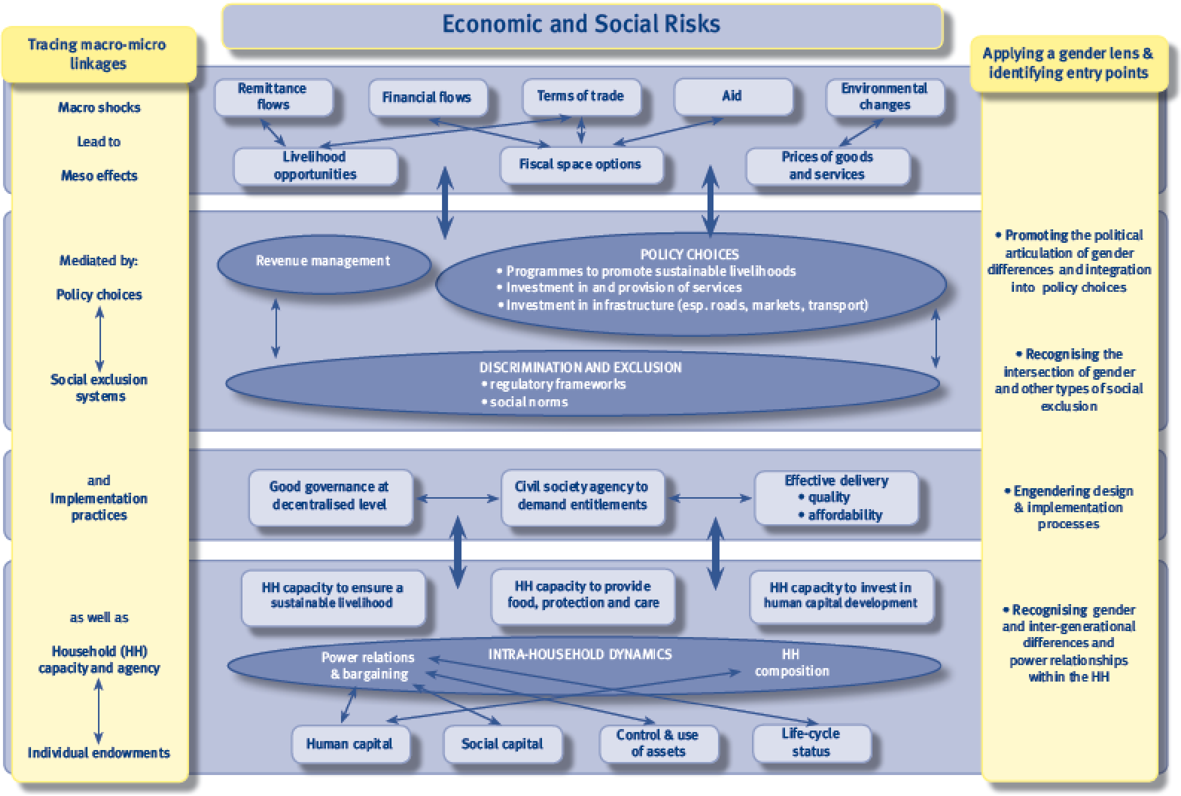
In their 'Evaluation of Concern Worldwide's Dowa Emergency Cash Transfer Project (DECT) in Malawi', Devereux and Sabates' (2004) present a good example of a theory of change.

Their framework of social protection, the objectives of the full range of social protection interventions are fourfold:
- *Protective*: providing relief from deprivation (e.g. disability benefits or non-contributory pensions)
- *Preventive*: averting deprivation (e.g. through savings clubs, insurance or risk diversification)
- *Promotive*: enhancing real incomes and capabilities (e.g. through inputs transfers)
- *Transformative*: addressing concerns of social equity and exclusion by expanding social protection to arenas such as equity, empowerment, and economic, social and cultural rights, rather than confining the scope of social protection to respond to economic risks alone through targeted income and consumption transfers.

Social protection refers to a set of instruments (formal and informal) that provide:
- social assistance (e.g. regular and predictable cash or in-kind transfers, including fee waivers, public works schemes, food aid)
- social services targeted to marginalised groups (e.g. family counselling, juvenile justice services, family violence prevention and protection)
- social insurance to protect people against risks of shocks (typically health, employment and environmental)
- social equity measures (e.g. rights awareness campaigns, skills training) to protect against social risks such as discrimination and abuse.

Additionally, Figure 2 (below) presents an analysis of the gendered economic and social risks which ensure that cash transfer programmes are fully integrating an understanding of gender differences into project design in order to maximise the benefits for women and men. This forms the basis of thinking around a holistic approach to women's economic resilience and empowerment.

Figure 2: Impact pathways of vulnerability to economic and social risks



Source: Holmes and Jones, 2009

Annex 10: Examples of empowerment indicators from reports included in the database

Empowerment area.	Type of outcome indicators measured.
Power within – agency and capability[1]	• Awareness of rights • Self assurance • Self confidence • Skills and capability • Self confidence in managing the business • Have future life goals/life planning • Feel more economically independent • Behaviour and communication with parents/guardians on financial issues (girls) • Respect from spouse • Attitudes to work • Life satisfaction
Power to – agency and institutional environment	• Business related knowledge and skills • Having to ask husband for money • Freedom of movement • Decision making along with husband • Status inside the family • Decision making about education, health and food. • Save and budget • Financial literacy • Gender norms / social norms / institutional gender norms • Confidence and ability to refuse sexual advances (girls) • Time poverty
Power over - Economic advancement	• Income of clients of services (micro finance) • Employment and self employment / labour force participation • Job promotion • Average weekly earnings / wages • Earning from different activities • Consumption of goods • Paying for healthcare and housing • Ownership of physical assets – productive assets and household assets • Control over cash • Saving • Informal borrowing • Managing production and marketing • Owning land • Access to credit • Access to market information • Increased productivity • Price advantage (for sales)

1 We have used some of the terminology of the DCED literature review on "Measuring change in women's economic empowerment at household level" that was just presented at the recent DCED meeting. This is combined with the concepts in VeneKlasen, L. and Miller, V., 2002, 'A New Weave of Power, People and Politics: The Action Guide for Advocacy and Citizen Participation', Practical Action Publishing.

	• Business revenue
Power with - Social relations, networks	• Social prestige, being recognised and respected • Recognition as productive contributor to the household and community economy • Self help groups formation • Networks • Social capital and social networking • Networks of friends • Collective voice and collective identity (cooperative) • Participation in village politics • Awareness of seats in local institutions

Annex 11: Examples of cost benefit analysis

The example below is taken from BARA/IPA (2013) and presents a useful example of cost-benefit analysis:

In this section, we report cost-benefit calculations using different assumptions and outcomes to assess costs and benefits. We first look at the cost-benefit ratio for the program in general and then consider structured vs. organic replication.

Overall SfC Program

On the cost side, a *first* cost component is the program's implementation costs. Using financial and administrative data from Oxfam America/FFH and our data on take-up, we estimate the implementation cost per household to be $16.72. This takes into account only costs incurred by the NGOs that were subcontracted by Oxfam America/FFH and not the costs for supervision and management by Oxfam America/FFH staff. As our impact estimates are ITT estimates, the implementation costs too are averaged over the number of (small) households in the treatment villages – independent of whether a member of the household participates in SfC or not. Saving more money also entails a cost for the household in the form of reduced consumption. As a proxy for this *second* cost component, we calculated the total cumulative contributions to SfC groups since the start of the program. This amounts to $17.50 per household. We use different measures to assess program benefits. A *first* measure is the total assets of the (small) household. Total assets include livestock, household and agricultural assets, and financial assets. The increase in total assets is $149.38 (se=60.1). Usually assets are not used directly in a cost-benefit calculation since we would rather use a measure which captures household's flow benefits from assets. Consumption is often viewed a good proxy for household wellbeing. Therefore, as a *second* measure, we use the yearly consumption of the (small) household – i.e. the sum of food consumption and non-food expenditures over 12 months – complemented by an assumed 5% return on livestock. We include a return on livestock to consumption because we are assuming that at the time of our endline survey the newly accumulated livestock has not yet increased income (and accordingly consumption) streams. This is consistent with the report's overall emphasis on the short-term nature of this evaluation, and how we are getting an incomplete picture of the program's long-term impacts.

Since we do not have a reliable estimate of the returns to livestock, as a sensitivity check, we constructed a *third* outcome measure, similar to the second, but assuming a 10% return on livestock. Our ITT estimates of the treatment effect on the second and third outcome measures are $34.57 (se=25.9) and $40.62 (se=26.7) respectively. Note that these impact estimates correspond to program benefits over a 1 year period. If the program benefits are sustained over several years, these benefit estimates are very conservative (i.e. they are lower bounds).

Putting the cost and benefit sides together yields the return on investment (ROI) rates presented in Table A6.1. When using the assets measure to assess program benefits, the point estimate of the ROI is very high (794%). Although statistically different from zero, the precision of the estimate is rather low. Including the savings contributions as a cost factor lowers the estimate of the ROI to 243%. The estimate is still high though and statistically significant at the 10% level. The high ROI rates are primarily driven by the large increase in livestock holdings described above. When using the consumption-based measures to capture the economic benefits of the program, the point estimates of the ROI are 107 and 143%, depending on whether we assume the rate of return on livestock holdings to be 5 or 10%. When taking into account the SfC contributions as a cost factor, the estimates of the ROI are basically zero. None of the consumption-based ROI estimates are statistically different from zero. In conclusion, the quantitative impact evaluation has shown that SfC brought modest impacts on household's wellbeing.

The program led to a large increase in livestock holdings in particular, but small impacts on food consumption – an outcome often used to capture household's wellbeing. We have not yet observed how those large increases in assets will translate to consumption down the road. Given that consumption is most often used in a cost-benefit analysis, and not assets, this presents a methodological challenge in determining *statistically* whether

SfC's benefits outweigh the costs. What this cost-benefit analysis highlights clearly is that the modest impacts stemming from SfC were achieved through a very inexpensive program.

Source: BARA/IPA, 2013

Another interesting example can be found in Ahmed (2009: 112) who calculated the cost effectiveness of cash transfers by using the findings of impact and effectiveness in terms of the economic results accruing to participants. The evaluators calculated the cost of transferring one taka worth of food to participants and the cost of increasing calorific intake by 100 calories. They were able to then compare these costs across four different programmes. With these figures they were also able to calculate the cost of reducing extreme poverty by 1%. The evaluation has a section with thorough analysis of the impact on women's empowerment of the four transfer projects. This section ends with an addition to the cost effectiveness discussion (Ahmed, 2009: 156). Here, they calculate the cost of increasing women's participation in decision-making on food and the cost of increasing the percentage of women taking NGO loans by 1%.

Another example can be found in page 25 of Bandeira 2012 – Empowering adolescent girls – Evidence from an RCT in Uganda.

Annex 12: Recommendations for Microfinance evaluations (Linda Mayoux)

- Include detailed consideration of context and not to assume that any signs of 'empowerment' are due to microfinance. In particular women should not be viewed as 'passive victims of subordination', but active agents who are constantly negotiating and addressing the challenges they face.
- Have a theory of change linked to MF interventions – whether it has a narrow or broad indicator/programme focus
- Carefully differentiate the range of different types of microfinance – different products, social intermediation services – different types of group structure and different types of training
- Be contextually grounded not only in gender norms, but also possibilities for change and the wider financial context
- Include empowerment priorities of women themselves
- Differentiate between women from different backgrounds, including implications for very poor women – even if the conclusion is that they need something apart from microfinance
- Include consideration of implications of MF for men – both men's empowerment, and effects of men's access to MF on gender relations with women. Compare impacts on women with impacts on men, and look at the reasons for any differences and the implications for designing more gender equitable products and services.
- Be participatory with multistakeholder perspectives and linked to the implementation/planning cycle/market research
- Disaggregate information to look at different impacts on women from different social and economic backgrounds, in particular possible different patterns of impact on very poor women and women from marginalized communities.

Microfinance gender impact list – for designing impact assessment or evaluation of microfinance services.

1. HOW FAR AND IN WHAT WAYS HAS WOMEN'S ACCESS TO RURAL FINANCIAL SERVICES INCREASED?
1.1. What informal and formal financial services (credit, savings, insurance, remittance transfer etc) exist in the area? Which did women normally use before the intervention? What financial services did men normally use? What were the differences and reasons for any differences? 1.2. Has access to these sources changed since the intervention? If so what and why? Do women now have equal access to all types of financial services compared with men? 1.3. Are the institutions gender-friendly? Are the products suitable? 1.4. Does the institution/intervention track gender-disaggregated data? What gender differences appear in the data with respect to access to different financial services? 1.5. If there are any continuing differences in numbers of women and men using different types of service, what are the reasons for this? Is this because of differences in aspirations and motivation? Explicit or implicit institutional gender discrimination (See also organizational checklist above)
2. HOW FAR AND IN WHAT WAYS HAVE RURAL FINANCIAL SERVICES INCREASED WOMEN'S ECONOMIC EMPOWERMENT?
2.1 What economic activities were already women engaged in? What economic activities were men involved in? 2.2 How were assets, income and resources distributed within households? Did women and men have different degrees of access? Different degrees of control? 2.3 Do women control the decisions about use of financial services? Loans, savings income, proceeds from

insurance claims, remittances? Independently, jointly with men, or marginally/not at all?

2.4 Have financial services enabled women to increase incomes from their economic activities? To enter new and more profitable/productive activities? To increase assets? To decrease economic vulnerability? Significantly or only a little?

2.5 How far do women control this income and assets? Is there more equal control over household income and assets? Significantly or only a little?

2.6 What do women use the loans/savings income/income from economic activities for? Investment in economic livelihoods? Or consumption?

2.7 Has women's market access increased? In existing markets only? In new markets? Has vulnerability to market fluctuations decreased?

2.8 Even if women do not use the income for their own economic activities, has their role in household decision making and their control over household income and/or assets increased?

2.9 What have been the impacts of financial services for men on all the above?

3. HOW HAVE RURAL FINANCIAL SERVICES CONTRIBUTED TO INCREASED WELLBEING FOR WOMEN AND THEIR FAMILIES?

3.1 What was the extent of gender inequality in well-being before the intervention? Food security? Health? Literacy and education? Freedom from violence? Other poverty indicators?

3.2 Even if women do not use the income for their own economic activities, has their role in household decision-making and control over household income increased?

3.3 What have been the impacts on women's nutrition, health, education, freedom from violence, happiness?

3.4 What have been the impacts on the nutrition, health, education, freedom from violence, happiness of other female household members: girls and the elderly?

3.5 What have been the impacts on the wellbeing of boys and men?

3.6 What have been the impacts of financial services for men on all the above?

4. HOW HAVE RURAL FINANCE PROGRAMMES CONTRIBUTED TO WOMEN'S SOCIAL AND POLITICAL EMPOWERMENT

4.1 Did women have personal autonomy and self-confidence before the intervention? freedom of mobility? of social and political activity? If not in what ways were they limited compared with men?

4.2 How have rural finance programmes increased women's self-confidence and personal autonomy?

4.3 How have they extended and strengthened women's networks and mobility?

4.4 How have they enabled women to challenge and change unequal gender relations? In property rights? Sexual violence? Political participation? Other dimensions of inequality?