WOMEN IN WAGE LABOUR: A SYSTEMATIC REVIEW OF THE EFFECTIVENESS AND DESIGN FEATURES OF INTERVENTIONS SUPPORTING WOMEN'S PARTICIPATION IN WAGE LABOUR IN HIGHER GROWTH AND/OR MALE-DOMINATED SECTOR IN LOW- AND MIDDLE-INCOME COUNTRIES

A SYSTEMATIC REVIEW PROTOCOL, MAY 2017

UKaid
from the British people

Africa Centre for Evidence
Working together to make evidence-informed decision-making a reality

UNIVERSITY OF JOHANNESBURG

ePPI CENTRE

UCL Institute of Education

The authors of this report are:

Laurenz Langer	(Africa Centre for Evidence, University of Johannesburg)
Janice Tripney	(EPPI-Centre, UCL Institute of Education, University College London)
Yvonne Erasmus	(Africa Centre for Evidence, University of Johannesburg)
Natalie Tannous	(Africa Centre for Evidence, University of Johannesburg)
Charity Chisoro	(Africa Centre for Evidence, University of Johannesburg)
Mary Opondo	(Africa Centre for Evidence, University of Johannesburg)
Luvuyo Zigana	(Africa Centre for Evidence, University of Johannesburg)
Ekwaro Obuku	(African Centre for Systematic Reviews and Knowledge Translation, Makerere University)
Carina van Rooyen	(Africa Centre for Evidence, University of Johannesburg)
Ruth Stewart	(Africa Centre for Evidence, University of Johannesburg; EPPI-Centre, UCL Institute of Education, University College London)

Funding

This is an independent report commissioned and funded by the Research and Evidence Division in the Department for International Development. This material has been funded by UK aid from the UK Government, however, the views expressed do not necessarily reflect the UK Government's official policies.

Acknowledgments

We would like to thank our review advisory group—Harsha Dayal, Thomas De Hoop, Sumana Hussain, Marcel Korth, Bernedette Muthien, and Josephilda Nhlapo-Hlope—for their insightful comments and inputs on the protocol. Sincere appreciation is also expressed to Alison Bullen for leading and conducting the scientific search for academic literature. We are also grateful for the helpful feedback and support received from members of the EPPI-Centre: Mukdarut Bangpan, Sandy Oliver, Claire Stansfield, Dylan Kneale, Jeff Brunton, Sergio Graziosi and James Thomas.

Conflicts of interest

There were no conflicts of interest in the writing of this report.

Contributions

The opinions expressed in this publication are not necessarily those of the EPPI-Centre or the funders. Responsibility for the views expressed remains solely with the authors.

Citation

This report should be cited as: Langer L, Tripney J, Erasmus Y, Tannous N, Chisoro C, Opondo M, Zigana L, Obuku E, van Rooyen C, Stewart R (2017) *Women in wage labour: A systematic review of the effectiveness and design features of interventions supporting women's participation in wage labour in higher growth and/or male-dominated sector in LMICs. A systematic Review Protocol*. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

Picture

WORLD BANK PHOTO COLLECTION. PHOTO ID: LN-SL005 WORLD BANK. CC BY-NC-ND 2.0.

CONTENTS

LIST OF ABBREVIATIONS

3ie	International Initiative for Impact Evaluation
CDG	Center for Global Development
DFID	UK Department for International Development
DoL	South African National Department of Labour
DoW	South African National Department of Women
DPME	South African National Department of Planning Monitoring and Evaluation
GRADE	Grades of Recommendation, Assessment, Development, and Evaluation
GDP	Gross Domestic Product
ICT	Information and Communication Technologies
ICRW	International Centre for Research on Women
IGO	Intergovernmental Organisation
ILO	International Labour Organization
LMICs	Low- and middle-income countries
NGO	Non-governmental Organisation
QCA	Qualitative Comparative Analysis
RR	Response Ratio
SE	Standard Error
SMD	Standardised Mean Difference
UN	United Nations

1 BACKGROUND

1.1 THE IMPORTANCE OF WOMEN'S EMPOWERMENT

Women's economic empowerment is both a means and an end in international development. Following Sen's definition of Development as Freedom (Sen 1999), economic empowerment is an essential capability to live the life one has reason to value and can therefore be seen as an outcome of development in its own right. Gender empowerment is also one of the 17 Sustainable Development Goals (SDGs) and four of the nine targets of this goal relate directly to women's economic empowerment (UN 2015). As a mechanism for development, there is strong longitudinal evidence that increased inclusion and participation of women in the labour market supports economic growth and development (World Bank 2012; IGC 2015; Kabeer 2012). Full participation of women in labour forces for example is estimated to add multiple percentage points to most national economic growth rates, thus increasing aggregate socio-economic development (UN 2015). Similarly, the McKinsey Global Institute estimates that if all women were able to participate fully in the economy, it would contribute up to US$ 28 trillion or 26 percent of annual global gross domestic product (GDP) in 2025 (McKinsey Global Institute 2015). The potential benefits of women's participation in the labour market underlines that the economic empowerment of women is by no means a zero-sum game: society as a whole stands to gain from it. Investing in women's economic empowerment is smart economics (Revenga & Shetty 2012). Increasing women's income leads to greater household spending on education and health (World Bank 2012). It also increases women's self-esteem, aspirations, and domestic bargaining powers while reducing domestic violence and delaying early marriage and pregnancy (UN 2013; World Bank 2012). In sum, as much as women's economic empowerment has intrinsic value and benefits for individual females, so it is equally a virtuous circle that socially and economically benefits society as a whole.

1.2 THE PROBLEM

Despite the well-established positive relationship between women's economic empowerment and socio-economic development, females in all regions of the world face significant barriers to their labour market participation (ILO 2016; UN 2013; World Bank 2012). Globally, only 50 percent of women participate in the labour force as compared to 76 percent of men. This participation gap has been narrowing in most regions of the world over the last decade, albeit at a slow rate: 2.4% increase in Europe; 0.8% increase in South-Eastern Asia and the Pacific; 7.9% increase in Latin America and the Caribbean; 3.2% in sub-Saharan Africa (ILO 2016). While these positive overall trends are encouraging, there is no region that has achieved equal labour market participation for women and men. In addition, women remain at a significantly higher risk of unemployment, constitute a smaller portion of the workforce in wage employment, and face higher barriers in the transition from schooling into the job market (ILO 2016; UN 2015).

What is more, the quality and nature of women's labour market participation differs from men's. Women participants in the labour market earn 24 percent less than men do globally (UN 2013). They present the majority in non-standard, informal, temporary, part-time, and low-paid jobs (ILO 2016). For example, globally 57 percent of all part-time workers are female. The same finding occurs in relation to time-related underemployment which is, in low- and middle-income countries (LMICs) particularly, significantly higher for women.

Much of this difference in the quality and nature of women's labour market participation can be explained by sectoral and occupational segregation. Globally women in employment are overrepresented and clustered in particular professions and sectors of the economy that offer lower salaries and less lucrative employment conditions. For instance, 60 percent of women in low-income and lower-middle income countries are employed in the agricultural sector taking up poorly paid but time- and labour-intensive jobs (ILO 2016). Likewise, women are overrepresented in running informal household businesses such as tuck shops and local garment businesses with little potential for growth and high market saturation (Vaessen et al 2014). An analysis of labour market compositions across 142 countries underlines this sectoral segregation with women consistently being overrepresented in the lowest paying professions, which explains the wage difference between women and men (ILO 2016; UN 2013).

In contrast, women are underrepresented in many high growth and well-paying professions. In most countries men dominate the occupation of plant and machine operators and assemblers; law and legislation; business administration and management; finance; and Information and Communication Technologies (ICTs) (ILO 2016). This pattern is particularly acute in LMICs and hinders the use of wage employment as a pathway for economic empowerment for women. The most common cause of this sectoral and occupational segregation is a reflection of gender stereotypes at work, in the family, and society as well as the lack of an effective policy environment for women's empowerment (ILO 2016; UN 2013).

In addition to horizontal segregation in the labour market, women are further experiencing vertical segregation within professions. Horizontal labour market segregation refers to sectors in which female employment is significantly lower compared to men as a share of the total labour force; this has been covered above. On the other hand, vertical labour market segregation refers to sectors in which female employment might be equal to or larger than that of men as a share of the total labour force, but in which women are underrepresented in particular occupations or positions. For example, although females constitute the majority of the labour force in the garment sector they are often underrepresented at a managerial level. Globally only five percent of the world's largest companies are managed by a female chief executive officer (ILO 2016). Vertical segregation can present a large barrier for women to either obtain a sufficient return on their labour or to use wage employment as a means to improve their livelihoods.

1.3 HIGHER GROWTH/MALE-DOMINATED SECTORS

This systematic review is concerned with the participation of women in the labour force in higher growth and/or male-dominated sectors only. Based on a preliminary review of the literature we have identified the sectors listed below to be relevant for inclusion. However, contextual differences negate a universal application of this list as the gender composition of the labour force in economic sectors across states differs, as does the productivity rate of sectors.

Economic sectors with high or growing productivity *and/or* which are male-dominated:

- Commercial agriculture: where productivity is higher than smallholdings or subsistence farming, and agriculture is linked to larger business supply chains and larger markets;
- Energy (mining and quarrying, electricity, gas and water supply);
- Trade;
- Transportation;
- Accommodation and food;
- Business administration services;
- Finance;
- Electronics and ICT;
- Maritime services;
- Wood pulp and forestry;
- Construction;
- Manufacturing;
- Higher education/Science and Technology;

Economic sectors with high or growing productivity that are *not* considered as male-dominated and are therefore excluded from the second stage of the review are:

- Education
- Health
- Social work
- Wholesale retail
- Communication services
- Tourism
- Public administration
- Garment industry
- Micro- and small businesses focussed on self-employment with no employees

However, we will include studies that focus on the issue of vertical segregation in female-dominated sectors. That is, even though the health sector as a whole is not considered a relevant sector, research that focuses on vertical inequalities in employment in the health sector, for example the number of females in senior positions, will be considered in our review.

1.4 BARRIERS TO WOMEN'S LABOUR MARKET PARTICIPATION

A range of barriers impede women's participation in higher growth and/or male-dominated sectors. First, women face discrimination by markets and work institutions. For example, many companies do not cater for women's caregiving responsibilities. The absence of paid parental leave and childcare facilities in LMICs renders many women unable to participate in the labour market. Likewise, in the public sector procurement systems and budget allocation often overlook women's needs and hinder their ability to benefit from public sector investments. For women in work, harassment and discrimination in the workplace presents an equally important barrier to meaningful and valued employment (CDG 2016; ILO 2016; ODI 2016; Peters et al 2016; UN 2013).

Second, women face constraints in access to credit, finance, and assets. In some LMICs women are not able to own property or open a bank account on their own. Economic resources, such as loans, are often controlled by men or – if they are accessible to women – they are often too small to be effective for business development. This structural inequality impedes women's abilities to start or expand their businesses and gain from the opportunities provided by high-growth economic sectors. Third, women experience disadvantages in their employability and entrepreneurship. Technical and business skills required to assume more senior positions or to enter high-skill professions are often inaccessible to women. This lack of access to skills disables them from benefiting from high-return professions such as those in the ICT and engineering industries. Acquisition of labour-related soft skills and access to economic information are also often provided less frequently to women, preventing them from gaining valuable work experience. Female entrepreneurs too face additional constraints to their business success with business networks often dominated by men and market information and opportunities being shared through informal channels rather than public ones (CDG 2016; ILO 2016; Kabeer 2012; ODI 2016; Peters et al 2016; UN 2013).

A fourth key barrier to women's labour market participation in LMICs is restrictive social norms and a subsequent lack of social capital. Women are often expected to refrain from participating in the labour market and when they do, socially acceptable employment opportunities are restricted to a small number of usually low-paying professions such as domestic work. Compared to men, women face greater constraints in their mobility and often lack social support to build economic aspirations. This absence of social capital leaves many females at a structural disadvantage having to overcome restrictive norms as a first step to their labour market participation (CDG 2016; ILO 2016; Kabeer 2012; ODI 2016; Peters et al 2016; UN 2013).

Finally, all human beings are subject to behavioural biases that influence social and economic decision-making (Kahnemann 2011; Thaler & Sunstein 2008). For example, most of us are influenced by the framing of messages or struggle to implement and honour an effective saving schedule. A growing body of research provides insight into how these behavioural and cognitive barriers can be overcome with often simple techniques and small

tweaks to programme designs such as identity cues or commitment devices. The same is true in relation to female economic empowerment. In 'What Works: Gender Equality by Design', Iris Bohnet for example lays out how reframing the wording on job adverts can increase female job applications (Bohnet 2016). The Centre for Global Development refers to similar findings in the context of LMICs: their report 'Women, Economic Empowerment and Smart Design' (CGD 2016) highlights the increased attention that research has paid to overcoming women's behavioural (social and cognitive) barriers to labour market participation.

Each of these five key barriers to women's labour market participation requires the use of a deliberate intervention implemented in the form of public policies and programmes. For example, entrepreneurship training provided to females may overcome their lack of labour market relevant skills; microfinance may provide otherwise inaccessible capital; career guidance could offer access to labour market information; labour laws might include affirmative action clauses to address structural inequality; governments' provision of better childcare facilities and public investments in infrastructure might reduce women's time spent on domestic chores and increase their mobility, etc. The list of possible interventions is much longer and underlines the urgency with which women's social and economic barriers have to be addressed.

1.5 THE INTERVENTION

This review is not limited to a particular kind of intervention. Its conceptual framework (Figure 1) is defined by the barriers to women's labour market participation in higher growth and/or male-dominated sectors. These barriers set the scope for what is considered a relevant intervention in our review: the review extends to any intervention likely to support women in LMICs to overcome any of the five barriers to their labour market participation introduced above. Broad groupings for potentially relevant interventions – which may take the form of a policy, programme, strategy, or other type of action – are presented in Table 1. This table is not intended as an exhaustive list of intervention types, and is likely to be refined as the review progresses. Eligible interventions may vary on several different dimensions. For example, relevant interventions may be:

- Complex, specialised, multi-dimensional programmes or much simpler interventions based on a single strategy;
- Implemented in different settings;
- Either routine and/or structured/tailored interventions;
- Varied by type and intensity;
- Delivered at various stages of the employment process (pre-employment, transition to employment, and post-employment); and
- Focussed primarily on something other than the objective of overcoming women's barriers to employment.

Table 1: Intervention categories

Categories	Examples of relevant programmes
I. Overcoming discrimination by markets and work institutions	
Interventions to balance work and family responsibilities	• Flexible working-time arrangements • Maternity and parental leave • Sick leave • Social security • Provision of care facilities (child and elderly)
Increase women's financial returns	• Salary/wage increases • Salary/wage alignments • Tax incentives
Changing business culture/practice	• Quotas and reservation approaches • Workplace gender equity programmes • Business leadership and management • Public sector practices in employment and procurement • Setting and enforcing effective laws to protect women from violence and exploitation at work • Anti-discrimination policies
Macroeconomic changes	• Gender mainstreaming • Gender-responsive budgeting • Trade liberalisation/export orientated production • Public works programmes • Local production systems • Active labour market policies • Industrialisation policies • Investment policies (domestic & or foreign) • Fiscal policies • Monetary policies
Provision of infrastructure	• Roads, public transport, street lights • Water and sanitation • Electricity and energy • Other time-saving consumer goods
II. Overcoming constraints in access to credit, finance, and assets	
Microfinance	• Micro-credit • Micro-savings • Micro-insurance • Financial inclusion
Cash transfers	• Conditional • Unconditional
Economic assets	• Access to formal loans and capital • Provision of capital in-kind

	• Business technology
Changes to land titles, business ownership, and inheritance	• Land reform • Property rights • Inheritance laws
Bundled services/combined structural interventions	
III. Overcoming constraints in employability & entrepreneurship	
Interventions to provide education/skills	• Technical skills training • Business skills training • Literacy/numeracy skills training • Soft skills training • Financial literacy training
Interventions to provide access to economic opportunities	• Provision of economic information (e.g. business networks, peer-support) • Job search assistance • Business advisory and mentoring • Career guidance and counselling
Interventions to provide work experience	• Internships • Apprenticeship schemes • Vouchers and subsidies (demand-side) • Vouchers and subsidies (supply-side) • Job placements
Interventions to provide support to businesses and entrepreneurs	• Value chain services and market access • Formalisation/certification of businesses (including fair trade) • Matching grants • Innovation support • Business advisory/mentoring • Micro-franchising
IV. Overcoming a lack of social capital & norms	
Social organisation	• Self-help groups • Collaboratives and collectives • Organised labour
Changes in norms and attitudes	• Mentoring/role modelling • Gender empowerment (e.g. ambition/confidence building and autonomy) • Outreach and awareness-raising including 'policy advocacy' • Women's political participation • Mass media and public education campaigns • Youth empowerment
V. Overcoming behavioural (social and cognitive) barriers	
Gender sensitive design	• Designing for time and mobility constraints of women • Designing for restrictive social norms (e.g. use of female implementers)

	• Designing for risk preference (e.g. risk aversion of women due to larger care responsibilities)
Behavioural nudges	• Commitment devices • Framing & identity cues • Micro-incentives • Reminders

Figure 1 below outlines the conceptual relationship between the interventions and the final impact on women's economic empowerment[1]. The left hand column lists the categories of interventions grouped by the barriers to female labour market participation that they aim to address. The second column from the left indicates the intermediate changes that are assumed to result from the application of these interventions. For example, a business skills training programme might increase a female entrepreneur's managerial skills or increase the employability of a female graduate. However, these intermediate changes cannot be regarded as a proxy for women's labour market participation as they only enhance the probability of participation. Empirical measures of labour market participation are provided in the final outcome column. Broadly our review is concerned with two types of female labour market participation in higher growth and/or male-dominated sectors: (i) wage employment in such sectors and (ii) the success of female entrepreneurs in such sectors. More detailed outcome measures of these two final outcomes are provided in section 3.2.4 though neither type of labour market participation can be equated with female economic empowerment. Employment in a higher growth sector such as construction, for instance, might not translate into meaningful improvements of women's economic empowerment if harassment at work is experienced and/or women have no control over the income they receive. As a last step in our conceptual framework, our review will therefore investigate to what extent increased labour market participation translates into meaningful and valued changes in women's economic empowerment.

[1] Figure 1 presents our initial conceptual framework used for the evidence map. Following the full review, we will revise the framework using the results of the QCA on design features of interventions.

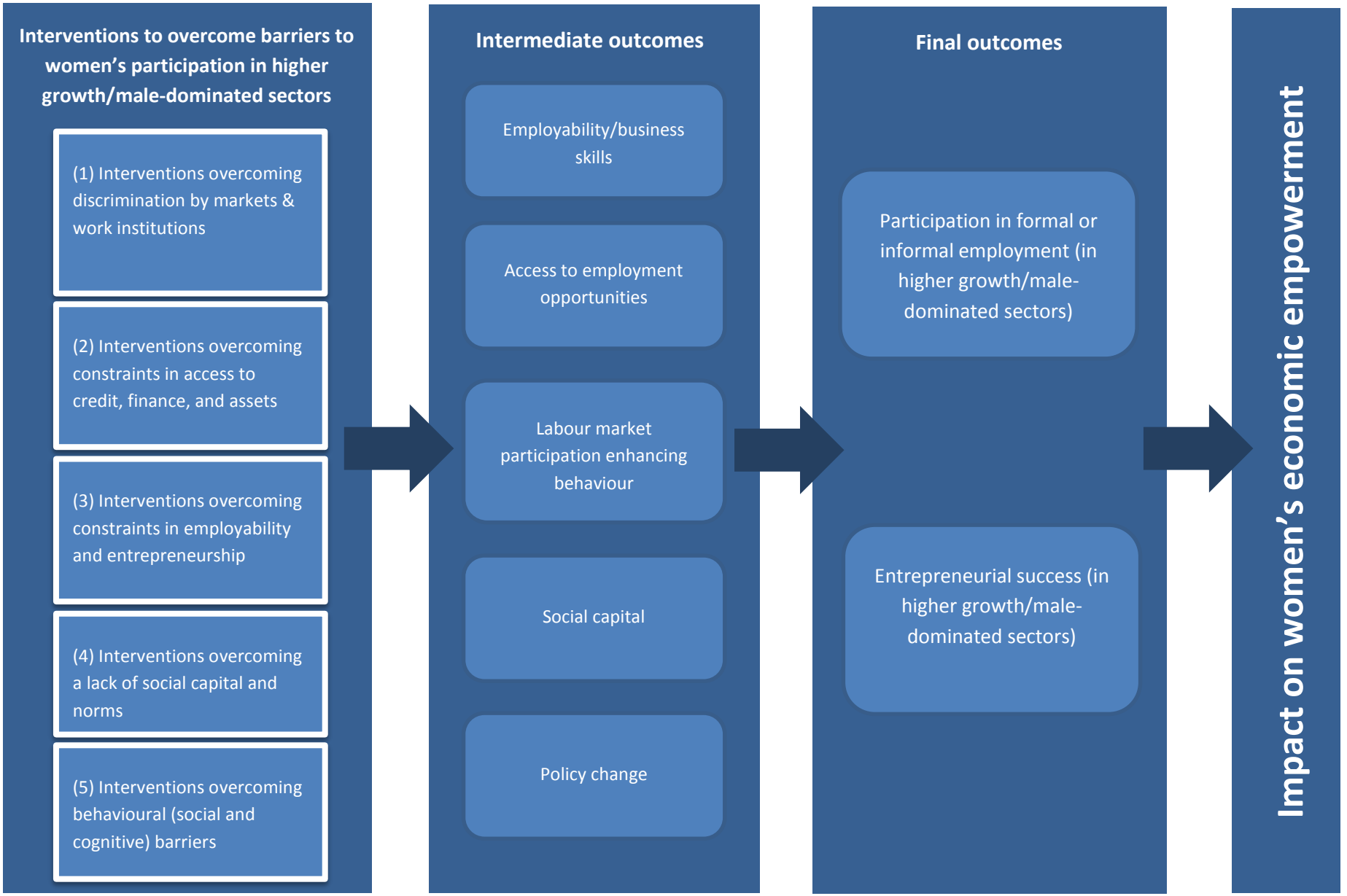

Figure 1: Conceptual framework

1.6 WHY THE REVIEW IS NEEDED

There are currently no systematic reviews that address the question of which interventions work to improve women's participation in the labour market in higher growth and/or male-dominated sectors. While there is ample research evidence attesting to women's underrepresentation in such sectors, there is so far no rigorous synthesis as to which interventions can change the economic empowerment of women in such sectors. There are a number of systematic reviews on the effects of microfinance and cash transfers that also conduct subgroup analyses for impacts on women's empowerment (e.g. Gibbs et al 2012; Yoong et al 2012; Stewart et al 2012; Vaessen et al 2014; Gopalaswamy et al 2016). However, none of these reviews explicitly defines women's empowerment as participation in labour markets, let alone participation in higher growth and/or male-dominated sectors. The same applies to the systematic reviews on entrepreneurial training in LMICs (e.g. Cho & Honorati 2014; Tripney et al 2013; 2015; Kluve et al 2016) which are not exclusively focussed on women or higher growth sectors. Brody and peers' (2015) review of economic self-help group programmes to improve women's empowerment as well as Gibbs and colleagues' (2012) synthesis of combined structural interventions for gender equality and livelihood security each include relevant interventions, but do not focus exclusively on labour market participation in higher growth and/or male-dominated sectors. Two recent reviews on the effects on business support (Piza et al 2016) and employment services, and subsidised employment (Kluve et al 2016) are of high relevance to women's labour market participation too, but neither focus on women exclusively or on sectors that are high growth and/or dominated by males.

At a policy level in international development, women's economic empowerment is a high priority. The International Labour Organization (ILO) has declared women's full participation in the labour market one of its centenary goals. Addressing women's economic empowerment is also directly mandated by SDG 5, and the UN Foundation published a major report on the state of research on women's economic empowerment in 2013 (UN 2013). This Roadmap report to women's economic empowerment is based on a review and synthesis of 136 empirical evaluations of women empowerment programmes and policies. The report was updated by the Center for Global Development (CDG) in 2016, which added 96 new evaluations that have been published since the launch of the 2013 Roadmap report. There is thus a rich and growing body of research evidence evaluating the effect of policies and programmes on women's economic empowerment.

However, the focus of this work is not on women's economic empowerment through labour market participation *per se*. Rather it investigates how women can be empowered to have greater control over their economic situation and ambitions. For example, interventions to increase women's decision-making power within the household, regardless of whether this increase is caused by participation in the labour market, are of high relevance in the UN's Roadmap. This scope is slightly different from our review as it does not focus specifically on wage employment as a means to empowerment, let alone employment in higher growth and/or male-dominated sectors. In addition, aside from the CGD and UN reports, little attention has been paid overall to the design features of women's economic empowerment interventions, which limits our understanding of why some interventions work (or do not work) and under what circumstances.

A systematic review of research evidence evaluating the effects and design features of interventions aiming to overcome barriers to women's labour market participation in higher growth and/or male-dominated sectors is therefore a timely contribution to the ongoing policy and practice debates on what works to support women's economic empowerment. The review will rigorously synthesise a sub-set of economic empowerment interventions focussed on labour market participation as a pathway to women's economic empowerment. The application of systematic review methodology, statistical meta-analysis, and qualitative comparative analysis (QCA) will further enhance the contribution of this review compared to existing synthesis work on women's economic empowerment. In conducting our systematic review, we will draw on the abovementioned existing reviews to ensure that we build upon, and do not duplicate, previous or ongoing efforts (these also refer to ongoing systematic reviews which overlap with some of the included interventions of our review, e.g. Chinen and peers (2016) and Ibanez and colleagues (2016)).

2 OBJECTIVES OF THE REVIEW

Our review aims to answer the following review question:

What are the effectiveness and design features of interventions that aim to overcome the barriers to women's participation in the labour market in higher growth and/or male-dominated sectors in low- and middle-income countries?

In doing so, we will address the following review objectives:

1. To produce an interactive evidence map of research evaluating interventions aiming to overcome barriers to women's economic empowerment in LMICs.
2. To provide a rigorous synthesis of impact evaluation evidence to identify the effects of interventions aiming to overcome barriers to women's economic empowerment in LMICs.
3. To identify design features that influence the effects of interventions aiming to overcome barriers to women's economic empowerment in LMICs.

3 METHODS

3.1 TYPE OF REVIEW

We will use a two-stage review approach that begins with a first stage initial evidence map before a second stage analytical review and synthesis.

EVIDENCE MAP

The evidence map will be broader in scope than the full systematic review. It will map evidence from impact evaluations and systematic reviews on interventions that aim to overcome barriers to women's labour market participation. It will use an intervention-outcome matrix to highlight the size and nature of the evidence for different configurations of interventions and outcomes. The framework used for this matrix is available online (Supplement 1[2]). The evidence map will be visualised on an interactive online interface using software similar to the International Initiative for Impact Evaluation's (3ie) evidence gap maps.[3] Stakeholders will be able to create their own custom evidence maps by filtering the evidence-base according to region, study design, and economic sector.

Although a product in its own right, the evidence map will primarily be used to support stakeholder engagement with the evidence-base and subsequent decision-making on the most effective review approach and scope. The bodies of evidence presented on the map will aid the identification of the interventions and economic sectors of most interest. The evidence map will also guide our initial discussions about the best way to synthesise the evidence in answer to the review questions. We envisage four possible scenarios on how the evidence map will inform the approach and scope of the review:

(1) A lack of evidence: The evidence map identifies a lack of evidence for relevant interventions, leading to the decision to exclude these interventions from the review scope.
(2) Existing systematic reviews: The evidence map identifies existing systematic reviews evaluating the effects of relevant interventions. Two options exist in this case. First, conducting a review of reviews, which synthesises the findings of the existing systematic reviews. Second, a re-analysis of the impact evaluations included in the existing systematic reviews. This second approach would be appropriate if existing reviews do not disaggregate intervention effects by gender.
(3) Existing impact evaluations: The evidence map identifies a sufficient number of existing impact evaluations of relevant interventions, justifying the production of a full systematic review.

[2] http://www.africaevidencenetwork.org/supported-by-ace/diary-of-systematic-reviews/

[3] http://gapmaps.3ieimpact.org/evidence-maps/land-use-change-and-forestry

(4) Deciding on relevant economic sectors: The evidence map will allow users to filter the evidence-base according to different economic sectors to explore patterns in women's labour market participation. Understanding the respective size of the evidence for different economic sectors will guide stakeholder engagement that informs decision-making regarding the scope of the review.

FULL SYSTEMATIC REVIEW

We will conduct an effectiveness systematic review (Snilstveit 2012). We will focus on (i) the effectiveness of interventions supporting women's participation in the labour market in higher growth and/or male-dominated sectors, and (ii) the design features of such interventions. The review will therefore include studies that measure the effect of interventions and that reliably attribute observed effects to the applied interventions. Individual effects will first be synthesised into an overall estimate of treatment effects, and then disaggregated according to identified design features of interventions.

As indicated above, the available evidence-base might warrant the production of a review of reviews for certain relevant interventions. This decision is dependent on stakeholder engagement and would require the drafting of a separate set of inclusion criteria. The remainder of this protocol assumes that this systematic review will synthesise existing primary studies.

The systematic review will be published as a full technical report and summary report. Both review reports as well as the evidence map will be publicly available.

3.2 CRITERIA FOR INCLUSION AND EXCLUSION OF STUDIES IN THE REVIEW

Studies must meet the following selection criteria to be included in the evidence map. These criteria for the evidence map are broader in terms of eligible interventions and economic sectors than are the inclusion criteria for the full systematic review. The latter will be further refined during our meeting with stakeholders to determine the scope for the full systematic review. A summary of the inclusion criteria for both the evidence map and the full systematic review is provided in Appendix 2.

POPULATION

Eligible studies have to meet all of the following population criteria to be included in the evidence map *and* the full systematic review.

Women: The study sample must include women aged 15 years and older. The study sample must either be majority[4] female or the study results must be disaggregated by gender. Where the study sample includes women below the age threshold, the majority of the study sample must either meet our age threshold or the study results must be disaggregated by age.

[4] Majority refers to a portion of 51% or higher of the sample.

Geographical location: Low- or middle-income country (as classified by the World Bank at time of data collection for the study, see Appendix 3). Middle-income countries refer to both lower-middle and upper-middle income countries.

EMPLOYMENT SETTING

The study must evaluate the effects of relevant interventions (as listed in Table 1) applied in economic sectors with high or growing productivity *and/or* which are male-dominated.

Productivity here is used as a proxy for a range of indicators such as growth in revenue, profits, income, employment, etc. Sectors that have already achieved high productivity in LMICs (e.g. transportation) will be included as well as sectors that are likely to experience productivity growth in LMICs (e.g. finance, ICT).

A preliminary list of sectors was presented in section 1.3 of this report and will be applied for the inclusion of studies in the evidence map. For the full systematic review, we will investigate the relevance of each economic sector individually in the context of the included studies. The above list will be further refined following stakeholder engagement after the mapping stage of the review.

We will include economic sectors (see section 1.3) with either horizontal or vertical male-dominance as defined in section 1.2. Employment status or experience does not present a criterion for exclusion or inclusion in our review.

INTERVENTION

We will include *any* intervention that aims to overcome the barriers to women's labour market participation in LMICs. It is therefore not possible to exhaustively pre-specify a list of relevant interventions. Section 1.5 outlined the main categories of interventions that we expect to encounter in this review as well as provided examples of interventions for each category. The term intervention in this context can refer to a policy, programme, strategy, technology, device, or any other type of deliberate action.

We will include multi-component (also known as bundled or combined) interventions only if (i) all intervention components aim to overcome a relevant barrier and/or (ii) studies can attribute the observed outcome to an intervention component that aims to overcome a relevant barrier.

EXCLUDED INTERVENTIONS

We will exclude macro-level interventions such as investment in basic to tertiary education, health care, citizenship, social welfare, and economic growth which are known to benefit labour market participation rates of the general population.

OUTCOMES

To be included in the evidence map, studies must evaluate the impact of interventions on one of the below three final outcomes:

(1) Participation in formal or informal employment (in higher growth and/or male-dominated sectors)
(2) Entrepreneurial success (following outcome 1: participation in higher growth and/or male-dominated sectors)
(3) Economic empowerment (following outcome 1: participation in higher growth and/or male-dominated sectors)

Relevant indicators for each final outcome are specified below in Table 2.

Table 2: Examples of relevant indicators for each outcome

Outcome	Relevant indicators
(1) Participation in formal or informal employment	• Employment status[5]; • Under-employment; • Nature of the employment (e.g. security of contracts, working conditions; salary and wage levels); • Progression and career prospect (e.g. promotion); • Changing employment from traditional to untraditional sectors for women's employment.
(2) Entrepreneurial success	• Business creation/survival; • Business income levels; • Business profits; • Revenue; • Firm size.
(3) Economic empowerment	• Household income and any other poverty related measure; • Control over household income/spending; • Access to economic assets; • Investment in economic assets/productivity; • Individual savings; • Women's well-being; • Women's empowerment.

Following stakeholder engagement, the scope of the full systematic review will be limited to participation in formal or informal wage labour only. This refers to outcome construct (1) in table 2. Note that therefore creation of and self-employment in micro- and small-businesses is not a relevant outcome.

[5] Excluding self-employment.

In cases where the majority of intervention participants are not female, studies will be eligible for inclusion only if impacts on women are assessed separately from those on men (i.e. in sub-group analysis) or in comparison to men.

INTERMEDIATE OUTCOMES

Studies will only be included in the evidence map if they assess any of the three final outcomes. For the full systematic review, studies have to focus on the single final outcome of participation in formal or informal wage labour. Studies investigating intermediate outcomes only will be excluded. The following intermediate outcomes are eligible for inclusion if they are reported as part of a study assessing final outcomes in the evidence map.

- Employability/business skills (e.g. technical knowledge, adoption of technology, change in business practices)
- Access to employment opportunities (e.g. awareness about job openings, changes in employment policies)
- Employment enhancing behaviours (e.g. attending job interviews)
- Social capital (e.g. networks, self-esteem)
- Policy change (e.g. change in labour laws, anti-discrimination policies)

For the full systematic review, only those intermediate outcomes referring to wage labour participation will be included. For example, this can refer to employability enhancing intermediate outcomes such as technical skills and access to wage labour opportunities such as information about opportunities.

We will also record information on unintended outcomes, e.g. an increase in violence against females and effect capture, as well as information on intervention costs or cost-effectiveness where reported.

STUDY DESIGN

We will include studies using either of the following quantitative experimental or quasi-experimental study designs:

(a) Designs using a random or quasi-random method of group assignment in which one of the following is true:

- Units (individuals or clusters of individuals) are randomly assigned to treatment and control groups by the investigator using a fully random procedure, such as computerised random number generation;[6]
- A quasi-random procedure presumed to produce comparable groups has been used by the investigators. For example, allocation by date of birth or next person to walk in the door (i.e. the method of allocation falls short of full randomisation);

[6] With or without matching.

- Regression discontinuity designs in which participants are assigned by the investigator to intervention or control groups solely on the basis of a cut-off score on a pre-programme measure.

(b) Designs employing non-random methods of assignment, in which one of the following is true:

- The investigator controls group exposure and assigns participants using a non-random procedure (e.g. alphabetically by surname);
- The investigator constructs the comparison group after the start of the intervention (e.g. by exploiting existing survey data);
- A natural experiment in which units exposed to the treatment and control conditions are determined by nature (e.g. change in policy or divergence in practice between regions) or by other factors outside the control of the investigators);[7]
- Assignment to conditions (treatment versus comparison) is by means of self-selection by participants or by administrator selection (e.g. by welfare officials).

Studies employing non-random methods of assignment must use appropriate methods that take account of selection bias and confounding at the design and/or analysis stage in order to be included in this review. In the context of this review, these include statistical matching (e.g. propensity score matching), difference-in-differences estimation, interrupted time series analysis, regression discontinuity analysis, instrumental variables (IV) regression, and certain forms of multivariate regression analysis such as the Heckman sample selection (two-step) model. If a study uses matching and/or covariate adjustment it may be done individually or by groupings (clusters) of individuals, and it may be based on participant characteristics observed before or after the start of the intervention. If there is no matching or statistical adjustments, then pre-treatment information on equivalence must be available and groups shown to be comparable.

Finally, control or comparison conditions in eligible studies refer to the population receiving no treatment, treatment as usual, an alternative treatment, or pipeline treatment.

DESCRIPTION OF METHODS USED IN THE COMPONENT STUDIES

Studies included in this review will employ experimental or quasi-experimental research designs that compare outcomes for an intervention group to those for a control or comparison group. The following studies in Table 3 exemplify the methods likely to meet the eligibility criteria for the review.

[7] Such studies depend on the premise that the 'assignment' of subjects to the treatment and control groups is equivalent to random assignment (though not in a controlled way). Natural experiments are often associated with IV, RDD, and difference-in-differences.

Table 3: Examples of eligible studies

Berlinski et al. (2011) Pre-school and Maternal Labor Market Outcomes: Evidence from a Regression Discontinuity Design	Intervention: pre-school attendance Design and estimation strategy: regression-discontinuity design (RDD) using pooled household surveys and exploiting sharp differences in the probability of attendance for young children born on either June 30 or July 1.
Field et al. (2010) Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India	Intervention: business training Design and estimation strategy: exploits the random assignment to treatment (i.e. to being invited to a training session) using instrumental variables (IV) specification, and a survey conducted on a rolling basis four months after training.
Gustavo et al. (2014) The Impact of Day Care on Maternal Labor Supply and Child Development in Mexico	Intervention: day care programme Design and estimation strategy: considers the natural process of enrolment of children in day care settings, where beneficiary children make up the treatment group and those on the waiting list make up the comparison group (with researchers verifying that groups were similar in observable variables).
Nopo et al. (2008) Occupational Training to Reduce Gender Segregation: The Impacts of ProJoven	Intervention: labour training program Design and estimation strategy: two-stage matching procedure that includes propensity scores (on the first stage), and gender and labour income (on the second one).

3.3 SEARCH STRATEGY FOR IDENTIFICATION OF RELEVANT STUDIES

METHOD TO IDENTIFY RELEVANT STUDIES

A comprehensive search strategy will be used to search the international research literature for qualifying studies. The aim is to identify all available evidence relevant to the review question. Our aim is to be sensitive rather than specific. We will deliberately formulate search strings and identify search sources that are over-inclusive. While this will increase the amount of citations to be screened, it will reduce the risk of missing relevant studies. We will search different types of sources, including sources with a particular focus on LMICs (see Appendix 3 for the current World Bank classification of low- and middle-income countries, grouped by region). We will use a wide range of

sources to capture both academic and 'grey' literature and reduce the omission of relevant studies to ensure that our search is as unbiased as possible.

DATA SOURCES AND SEARCH STRATEGIES

Our search strategy will rest on three pillars: (1) a formal search of academic databases using explicit search strings based on Boolean operators; (2) a formal search of grey literature using mainly key word searches, but applying full search strings where organisational databases allow for the application of Boolean operators; and (3) an informal search using different snowballing techniques.

ACADEMIC DATABASES

A wide range of social science bibliographic databases will be searched covering international development, economics, sociology, psychology, education, and health care. These are listed below:

- Web of Science
- Econlit (EBSCO)
- ERIC (EBSCO)
- Business Source Complete (EBSCO)
- PRISMA database
- Sociological Abstracts (CSA)

The search of academic sources will be led by an information scientist, who will develop a detailed search query (or string) based on the inclusion criteria, relying on the database's index terms where available and/or free-text terms. Synonyms and wildcards will be applied as appropriate. Database thesauri will be consulted to ensure that all appropriate synonyms have been included. The search query will be tailored for each bibliographic database, and these will be peer-reviewed and piloted. There will be no language restrictions to the search. A publication year filter to identify studies published since 1990 will be used. This cut-off date was chosen as structural inequalities in women's economic empowerment only started to receive increasing attention in the 1990s (Kabeer 2012). A master search query for the Web of Science database is presented in Appendix 4.

The general key concepts that we will use for the search query are presented below. These are directly informed by our inclusion criteria and the interventions and outcome framework presented in section 1.5. We will combine search terms related to the four key concepts using the following Boolean combination: 1 AND 2 AND 3 AND 4. We deliberately opted not to search for outcome terms as women's economic empowerment and labour market participation are not well defined. As such we did not apply terms related to an outcome concept in our search string to avoid missing relevant citations.

Key concepts

1. Developing Countries
2. Women
3. Type of study
4. Intervention

Search 1, 2, 3, and 4 in Title, Abstract, Keyword, Subject Heading.

GREY LITERATURE SEARCH

The grey literature search will be informed by our content experts and advisory group members and will aim to cover a large variety of organisational repositories and websites. This selection will include umbrella organisations with general knowledge repositories relevant to the review question such as the World Bank's knowledge hub or the R4D database of the UK Department for International Development (DFID), as well as organisations specifically focussed on a sub-area of knowledge relevant to the review question such as the International Centre for Research on Women (ICRW) and the ILO. A full list of sources is presented in Appendix 4.

USE OF SNOWBALLING TECHNIQUES

The snowballing search will include hand-searching the content pages of key journals of particular relevance to the review question; backward citation searches (i.e. searching the reference list of included studies and seminal papers); forward citation searches (i.e. using Google Scholar to search for papers that cited included studies); requests to key authors and organisations to share studies with the review team; and a Twitter call for the wider community of practice to contribute relevant studies. Appendix 4 lists sources for the snowballing search.

MANAGING AND DOCUMENTING THE SEARCH AND SELECTION PROCESS

Review management software (EPPI-Reviewer 4) will be used to manage the entire review process (Thomas et al 2010). All potentially relevant items identified through the academic database search will be exported to EPPI-Reviewer and then manually screened for eligibility, with EPPI-Reviewer used to keep track of decisions made about each citation. Search hits from organisational repositories and snowballing will be stored in MS Word, and only the details of studies deemed relevant for the map, plus those over which there is any doubt, will be transferred to EPPI-Reviewer. In such cases, it will be necessary to check whether the item is already in EPPI-Reviewer before proceeding to manually enter details. Upon screening against the selection criteria, a record of all decisions taken (include/exclude/unsure) will be kept in EPPI-Reviewer and MS Word, as appropriate.

All information retrieval and selection activities in the review will be documented and described in sufficient detail in the final report so that the processes can be replicated by other researchers. Summary flowcharts will be used to convey information, where relevant. Based on the Cochrane PRISMA checklist for reporting results of searching and screening, the following information will be recorded: databases, database platforms, search strategy for at least one database, dates of search and time-frame.

STUDY INCLUSION

Selection of primary studies will be based on the pre-developed selection criteria described in section 3.2. These criteria will be piloted by two researchers who will screen a sample of search hits independently and then compare and discuss their assessments. Discrepancies will be resolved by further examination of the respective titles and abstracts. If a final decision cannot be reached, a third reviewer will be asked to reconcile differences. This process will be repeated until consistency in application of the selection criteria is achieved.

Following pilot testing, the remaining literature will be screened for eligibility by individual reviewers (i.e. single screening) who will work through the selection criteria hierarchically. The study selection process will be undertaken in main two phases – screening for the map and screening for the in-depth review – and will commence once all the hits from academic databases have been exported into the EPPI-Reviewer 4 database and all potentially relevant items from the grey literature and snowballing searches have been saved in MS Word. The selection criteria for the map are those listed in section 3.2. The criteria for the full review will be finalised after engaging with our stakeholders, and are likely to be slightly narrower than those used for the evidence map.

PHASE ONE: TITLE AND ABSTRACT SCREENING FOR THE MAPPING EXERCISE

We will screen studies for inclusion in the evidence map on title and abstract only. During this phase, we will screen two sets of studies: those that have been imported to the reviewing software and those that have been saved in MS Word. To this end, we will start by manually examining the titles and abstracts of all records entered into the reviewing software after removing duplicates. The relevance of each item will be assessed and decisions regarding each study recorded in EPPI-Reviewer. Items will be included in the map if they appear to meet the criteria outlined in section 3.2 on the basis of the information in the title and abstract, and excluded if they are clearly ineligible. Where we are in any doubt as to a study's eligibility (e.g. because no abstract is available, or it does not provide enough information), we will classify it as 'unsure'. We will give studies the benefit of doubt if there is reference within the wording of the title to (women of working age) AND (a relevant outcome OR a term suggesting the study was an evaluation).

For sources that do not allow us to export bibliographic information, we will screen the items that have been saved in MS Word, transferring the details of studies deemed eligible for the map, and any for which we are unsure, to EPPI-Reviewer 4. Items classified as 'unsure' during phase one screening will not be included in the map.

In cases where the title and/or abstract are not in English, the translation service offered by Google (http://translate.google.com/) will be used to translate the information into English; screening against the selection criteria will then proceed as normal.

PHASE TWO: FULL-TEXT SCREENING TO IDENTIFY STUDIES FOR IN-DEPTH REVIEW

We will screen studies for inclusion in the in-depth review using full texts. These will be obtained for studies included in the map and for the items marked as 'unsure', and detailed manual examination

of the full reports will be undertaken independently by pairs of reviewers. Disagreements between the reviewers' decisions will be resolved by identification of the source of the disagreement, re-reading of the text, and discussion. If a final decision cannot be reached, a third reviewer will be asked to reconcile differences.

In the event that we cannot determine the eligibility of a study (e.g. because the full text is unavailable, insufficient information is provided, or the only version we have is in a language other than English), the study authors will be contacted to request additional information about the study or access to English-language versions. If this is not successful, the references for these 'possibly relevant' studies will be listed as such in an appendix of our technical report, but the studies will not be included in the systematic review.

CRITERIA FOR DETERMINATION OF INDEPENDENT FINDINGS

Sometimes one evaluation leads to several study reports (e.g. working papers and journal articles). Efforts will be made to identify all affiliations between study reports before coding commences, using information on study sample sizes, intervention details, grant numbers, and so on. In cases where multiple reports are found to relate to a single study, reviewers will choose one as the main report (e.g. the publication containing the most complete dataset). When extracting data, the full set of available reports will be used to code each study.

It may also be the case that a single report describes more than one study (e.g. a single publication could describe a series of evaluations conducted in different countries, using different datasets). In this event, each study will be coded individually as if they had come from separate reports.

Sometimes an intervention will have been evaluated several times. If we find that several evaluations are based on the same data, these will be treated as part of the same study, even if the reports are written by different authors. If the intervention has been evaluated on multiple occasions using different datasets, then we will treat the different reports as separate studies whilst noting their relationship. Should it be unclear whether multiple reports provide independent findings, authors of primary studies will be contacted for clarification.

We intend to use meta-analysis to synthesise results across primary studies. In a single meta-analysis, it is important to include only one effect size measure per study. Estimated treatment effects cannot be regarded as independent of each other if the underlying data are derived from the same sample populations (i.e. some participants contribute information to more than one effect estimate).

Individual effect size estimates may be correlated if, for example, the study analyses:

- Different sub-groups of the treatment population (e.g. young women; highly educated women);
- Outcomes at different times but on the same units (e.g. midline and endline findings);
- Multiple outcome constructs for the same type of outcome (e.g. effects on hourly wages and earnings);
- Multiple interventions with the same sample of participants;

- Multiple treatment groups and the same control group;
- Effects using different methods;
- Several types of treatment effect estimates (e.g. Average Treatment Effect (ATE), Intention to Treat Analysis (ITT));
- Multiple time-points for the same individual (e.g. repeated observations for several follow-up periods); and
- Variations of the above.

In such cases, one of the following approaches will be pursued, as appropriate. Where studies report multiple effect sizes by subgroup we will report data in separate analyses. In the event that we have more than one effect size per outcome construct and study, we will combine different estimates within each study into one effect size per subgroup. Estimating a single (within-study) composite effect size will involve computing a sample-weighted average effect size for each study that accounts for differences in sample size, using appropriate formulae (Borenstein et al 2009). If a study included in the review has used several different techniques to estimate treatment effects for the same outcome (e.g. both statistical matching and regression analysis), the estimate with the lowest risk of bias will be used in the meta-analysis.[8] In the event that risk of bias assessments are similar, we will choose between different estimation approaches using the approach outlined by Tripney and colleagues (2013) (e.g. comparing the efficiency of the estimator in studies that use both matching and covariate adjustment).

3.4 DATA EXTRACTION AND CRITICAL APPRAISAL

DATA EXTRACTION

We will use a pre-defined data extraction tool in order to systematically and transparently extract data from the included primary studies. The tool, a draft version of which is presented in Appendix 5, will be translated into a coding set on EPPI-Reviewer to extract information required for both the mapping exercise and the in-depth review. Data will be entered directly into the EPPI-Reviewer database.

For the map, data collection will focus on key variables of interest to position studies on the evidence interface. Key variables refer to the intervention, measured outcomes, applied study design, the region in which the research was conducted, and the economic sector targeted. Coding will be based on information detailed in the study titles and abstracts only.

For the in-depth analysis and synthesis of study results, full text reports will be examined and studies coded on variables related to the study context, the characteristics of the study samples, details of the intervention design and its implementation, study methods, and the outcome variables and

[8] A study may have used the same set of data but different estimation methods and published the results in a single report or in separate multiple reports (e.g. a different report on each of the estimation methods used).

data. An explicit breakdown of each intervention according to the taxonomy of design features will be crucial in allowing us to allocate effect sizes to different configurations of intervention designs.

Two members of the review team will pilot the data extraction tool, working independently on a purposive sample of eligible studies selected to test the tool on the full range of included impact evaluation designs and methods. This process will be repeated until a very high level of consistency in reviewers' application of the codes is achieved, at which point the tool will be finalised. From thereon, the remaining studies will be coded by individual reviewers with a sub-set of full texts being coded by different combinations of two reviewers independently extracting information from each study report and then coming together to compare their decisions. Any uncertainties and discrepancies will be resolved by discussion, further review of the respective study reports, and where necessary consultations with a third reviewer.

CRITICAL APPRAISAL

A critical appraisal tool will be applied to assess the impact of bias on the trustworthiness of primary study results, where trustworthiness refers to the confidence of the review team that the findings reported in the included studies used for the synthesis are rigorous and credible. In order to assess the risk of bias of primary studies, we will adapt the Cochrane risk of bias tool for non-randomised studies (Sterne et al 2016), which we have previously used and adapted in international development reviews (Stewart et al 2015). Sterne and colleagues use a domain-based risk of bias assessment covering the following six indications of trustworthiness: (i) selection bias; (ii) confounding bias; (iii) bias due to departures from applied interventions; (iv) bias due to missing data; (v) bias due to measurement of outcomes; and (vi) bias due to selection of the reported result. Each domain of bias will receive a low, moderate, high, or critical risk of bias rating allowing for a transparent calculation of an overall risk of bias score for each study. Studies with a critical risk of bias will be included in the review but excluded from the synthesis.

A draft critical appraisal tool to assess studies for the full review is presented in Appendix 6. It will be piloted using a similar approach to that used for the data extraction tool. Two reviewers will independently assess each study and then come together to compare their decisions. In the event that these reviewers cannot come to an agreement about the risk of bias rating for a particular study, a third reviewer will be consulted.

3.5 EVIDENCE SYNTHESIS

Where possible, this systematic reviews intends to conduct an aggregative synthesis in order to (a) distinguish between effective and non-effective interventions, and (b) identify those configurations of intervention design features, participants, and contextual characteristics that may be associated with a given outcome. The analysis will focus particularly on identifying design features of interventions that are critical to their success. It is our intention to use meta-analysis to identify overall effects and qualitative comparative analysis (QCA) for the disaggregated analysis according to intervention design features (Thomas et al 2014).

STANDARDISED EFFECT SIZES

In order to standardise outcomes, the calculation of effect sizes for each intervention outcome is required for both of the quantitative synthesis methods mentioned above. Extracted statistical information for effect size calculation will be exported into MS Excel and reviewers will document the different computations and formulae used for the effect size estimates derived from each study. Other web-based resources (e.g. the Campbell Collaboration's effect size calculator) will be utilised for the less common statistical representations.[9] As noted above, it may be necessary to combine effect sizes at the study level, allowing each composite effect size to be considered independent before analysis takes place (see section 3.3.5).

Standardised mean differences (SMD) or Response Ratio (RR) effect sizes present the most relevant metrics in the context of this review. Where possible, SMD and RR will be calculated for continuous outcome variables, while only RR will be calculated for dichotomous outcome variables. By using SMD and RR calculated treatment effects, we will indicate the ratio of, or difference between, treatment and control groups. Therefore, an increase in the outcome attributed to the intervention as compared to the control—an 'effective intervention'—is indicated by a SMD greater than zero or an RR greater than 1. An SMD less than zero or an RR between zero and one indicates a reduction of outcomes for the population exposed to the intervention as compared to the control group—i.e. a harmful intervention. An SMD equal to, or not significantly different from zero, and an RR equal to one indicates that there were no observable changes in outcomes between the treatment and control group. All effect sizes will be coded such that positive effect sizes represent positive outcomes (e.g. the intervention lowers duration of unemployment, leads to higher wages).

In the event that we do not have consistency across our data (i.e. effect sizes based on either all raw data or all log-transformed data), Higgins and colleagues (2008) will be consulted for guidance on data transformation. We will correct for potential sample bias in the effect sizes (due to small sample sizes) by converting them into Hedges *g* using the correction procedure developed by Hedges and Olkin (1985). The detailed approach to effect size calculations is provided in Appendix 7.

UNIT OF ANALYSIS ISSUES

We will attempt to identify and address any unit of analysis errors, which typically arise if the study conducts programme assignment and analysis at different levels and the analysis does not adequately account for this. This can be a problem in both randomised controlled trials and quasi-experimental studies where treatment allocation is clustered. Both the unit of assignment to treatment and comparison groups and the unit of analysis will be coded for all studies. To correct for variation associated with cluster-level assignment, adjustments will be made using the correction

[9] https://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php

formula suggested in Hedges (2007). If sufficient information to apply this formula is not provided by the primary authors, then we will follow the approach described in Higgins and Green (2011).

DEALING WITH MISSING AND INCOMPLETE DATA

We will attempt to extract data on all variables of interest listed in the data extraction tool (see Appendix 5). If we have studies that are missing data considered essential for the review, we will make thorough attempts to contact the original investigators. Missing data will be requested, along with information on whether or not it can be assumed to be missing at random. If the necessary data to compute effect sizes cannot be retrieved from authors, we will consider imputing the missing data with replacement values. In this event, we will make explicit the methods used (Higgins & Green 2011). Proportions of missing participants will be coded and included in the risk of bias assessment. Reasons given for missing data will be provided in the discussion section of the full review report, and the potential impacts of missing data on the findings of the review will also be discussed.

SYNTHESIS OF OVERALL EFFECTS

The intention is to use a statistical approach to synthesising results across primary studies. It is assumed that data extracted from the studies will be pooled using meta-analysis. However, pooling of extracted data will depend on such factors as the heterogeneity of the studies and study populations. Key features of the participants, interventions, and outcomes will be described in summary tables, along with effect size estimates and methodological quality characteristics. The analysis will be conducted using the specialised built-in meta-analysis function within EPPI-Reviewer 4.

METHODS OF SYNTHESIS (1): META-ANALYSIS

Meta-analysis is the most rigorous method to synthesise quantitative evidence (Lipsey & Wilson 2001; Borenstein et al 2009). As a statistical approach it aggregates the numerical findings such as effect sizes of primary research to report a pooled overall numerical value. This numerical value – the pooled effect size – expresses the overall finding derived from the combined primary research results. The pooled effect size reflects the direction and magnitude of the observed primary effect sizes, which are allocated different weight in the analysis depending on sample sizes and variance.

We will report calculated effect sizes in tabular format as well as using forest plots. Where sufficient contextual homogeneity prevails, effect sizes will be averaged across studies by using an inverse variance weighting of the individual effect size. This weighting will result in the individual effect sizes of studies with larger study samples being given more weight in the combined, pooled effect size. The meta-analysis will be carried out using random effects statistical models.

The studies included in the review are likely to include impact evaluation designs based on a random (or quasi-random) method of group assignment, and those that use non-randomised procedures. Some caution is needed when synthesising effect sizes from studies using different designs and estimation techniques. Where appropriate, we will summarise across designs and make explicit our rationale for doing so. In addition, we will calculate separate estimates of treatment effects for

randomised and non-randomised studies. If sufficient studies are identified, we will also provide separate summary effect sizes for each distinguishable class of randomised and non-randomised design. If relevant, the synthesis will also separately analyse studies with different kinds of counterfactuals (e.g. 'treatment' versus 'no treatment' studies will be combined separately from 'treatment' versus 'alternative treatment').

ASSESSMENT OF HETEROGENEITY

In order to visibly examine variability in the effect-size estimates, we will use forest plots to display the estimated effect sizes from each study along with their 95% confidence intervals. Subsequently, and acknowledging the limitations of a quantification of heterogeneity and the different strengths of statistical approaches, the following test for heterogeneity will be conducted: calculation of the Q statistic as a statistical test of heterogeneity (Hedges & Olkin 1985); calculation of the i^2 and Tau^2 statistic to provide estimates of the magnitude of the variability across study findings caused by heterogeneity (Higgins 2002; Higgins 2003).

SENSITIVITY ANALYSIS

To test the robustness of the results of the meta-analysis, a number of sensitivity analyses will be conducted. Broadly, this will involve collecting data on, and assessing sensitivity of, findings to (i) the methods of the primary studies and (ii) the methods of the review.

The included studies are likely to vary methodologically. If sufficient studies are identified, we will conduct sensitivity analyses to examine the influence of these variations on the summary measures, in order to offer possible explanations for the differences between studies when interpreting the results. We will aim to examine whether the results are sensitive to study design, the risk of bias associated with the study, the degree of missing/incomplete data, the way outcomes were measured and the timing at which they were measured, the treatment effect estimator and comparator, and the use of adjusted analysis.

A series of sensitivity analyses will also be conducted to examine whether any decisions made during our own analyses substantially alter the review findings: for example, the specific statistical procedures and methods we selected to compute each effect size, decisions relating to transformation between effect size metrics, the way outlier effect sizes and sample sizes were handled, and whether or not we replaced missing data with substituted values.

As explained above, the main objective of the sensitivity analysis will be to serve as visual tool that allows for informal comparisons whether the results of our meta-analyses are sensitive to methodological decisions of our review team. However due to the controversy of pooling studies of random and non-random design, as well as of different risks of bias, we will follow up the sensitivity analyses of these two variables with a one-way random effects ANOVA model calculated in EPPI-reviewer. That is, the mean effect size and standard error for each group of studies is calculated to test whether these means are statistically significant from one another (Lipsey & Wilson 2001).

MODERATOR ANALYSIS

If there are sufficient data, we will conduct moderator analyses to try to explain variation in effect sizes. Moderator analyses will be reported in tabular format below each meta-analysis. Analyses will be calculated using the same one-way random effects ANOVA model as applied for sensitivity analyses.

Below, we pre-specify a list of *a priori* moderator variables assumed to moderate the true effect of the interventions but caution that these are subject to change following stakeholder engagement based on the presentation of the evidence map.

We will investigate whether findings differ according to key contextual factors including the PROGRESS-plus categories[10]. The PROGRESS-plus framework is used in order to bring an equity lens to the review and synthesis process. We will engage with our stakeholder group as to identify the most important moderator analyses to conduct once data extraction is complete.

- Geographical region, including the country classification employed in the United Nation's Roadmap for Promoting Women's Economic Empowerment (UN 2013):
    - High fertility agrarian societies
    - Declining fertility urbanising societies
    - Declining fertility formalising economies
    - Ageing societies
- Place of residence
- Ethnicity
- Occupation
- Religion
- Education
- Social capital
- Socio-economic position
- Age
- Disability
- Sexual orientation
- Other vulnerable groups

We will also analyse several population sub-groups of interest including the population of categories used in the United Nation's Roadmap for Promoting Women's Economic Empowerment (UN 2013):

- Young women vs. older women

[10] Note that gender has been excluded from the PROGRESS-plus framework as this review will not investigate the impact of interventions on men.

- Non-poor women vs. poor women vs. very poor women[11]
- Rural women vs. urban women
- Level of social organisation (e.g. collectives vs. individuals)

However, the main emphasis will be on identifying design features of interventions that are critical to intervention success. Here, we present a preliminary list of design variables to be tested. This list will be further refined following stakeholder engagement. Following the data extraction stage of the full review, we will further iterate the list of relevant design features based on the available evidence. This iteration will similarly be validated and refined through stakeholder engagement.

- Gender-sensitive programme design and service provision
    - Taking into consideration travel constraints (e.g. distance to homes)
    - Taking into consideration time constraints (e.g. when the programme is offered)
    - Taking into consideration caregiving constraints
    - Taking into consideration the gender of programme implementers
- Changes to social/professional norms (e.g. changing norms around women's appearance in public as an intervention mechanism)
- Dosage of the intervention (e.g. length of business training, size of loan)
- Quota/reservation approaches to ensure women's participation
- Demand-driven approaches (i.e. demands by markets)
- Financial input from employer
- Interventions working through subjective economic empowerment:
    - Increased economic self-reliance
    - Increased self-confidence
    - Increased risk taking
- Interventions addressing cognitive and social determinants of economic behaviour:
    - Protecting women from family and community pressures
    - Decision-making support
    - Increasing women's choice and autonomy
- Interventions addressing pervasive gender biases embedded in organisations and working environments
- Interventions applying explicit behavioural designs:
    - Commitment devices
    - Framing and identity cues
    - Designing for risk preference of population (e.g. women being more risk averse)
    - Micro-incentives (e.g. performance-based rewards, social tokens)
    - Reminders

[11] This will follow the definition of non-poor, poor, and very poor applied in the UN Roadmap for Women's Economic Empowerment (2013).

The above variables related to intervention design features will be aggregated into higher-level categories to allow for a meaningful QCA or moderator analysis.

Lastly, we will also conduct moderator analyses using the economic sectors which are targeted for an increased participation of female labour in the included studies.

As indicated earlier, it is our intention to use QCA as a substitute to the moderator analyses. The decision as to whether a substantive moderator analysis (ANOVA) or a QCA presents the most relevant synthesis approach to identify intervention design features will depend on the identified heterogeneity in the sample of included studies. If neither type of quantitative synthesis is possible, we will use narrative synthesis for analysing sub-groups. Our proposed approach to QCA follows.

METHODS OF SYNTHESIS (2): QUALITATIVE COMPARATIVE ANALYSIS

QCA investigates the configuration of different conditions and their association with intervention effectiveness (Thomas et al 2014). A condition refers to different design features in the context of our review, e.g. the use of behavioural techniques, gender sensitive design, and the intensity of the intervention. Using QCA enables us to analyse whether configurations of different design features are present (or not) when the intervention has been effective (or not) in increasing women's labour market participation. In comparing different types of configurations, QCA applies Boolean logic to establish necessary or sufficient conditions for intervention effectiveness. We will refer to these conditions as the active ingredients of women's labour market participation interventions in our review.

The advantage of this synthesis approach is the presentation and investigation of overlapping pathways to causality, which is to be expected in our review as it is unlikely that a single intervention or design feature is the single cause of positive labour market participation effects. While meta-analysis is restricted to evaluate the effects of different variables individually, QCA allows us to evaluate the effects of different configurations of conditions. For example, conducting a moderator analysis as part of our meta-analysis, we could examine whether intervention intensity has a significant impact on intervention effectiveness *or* whether using behavioural approaches has a significant impact on intervention effectiveness, and so forth. In contrast, QCA allows us to compare different configurations of these design features (e.g. using a behavioural approach *and* a gender-sensitive design) and how these configurations—not the conditions themselves—are associated with intervention effectiveness. We assume that providing different configurations of intervention design features will provide more policy-relevant synthesis results that can inform future policy and programme designs.

Our approach to QCA will follow the six step method for QCA proposed by Thomas and peers (2014) based on Rihoux and Ragin (2009). The authors propose the following six steps, but suggest that step five – "consideration of the 'logical remainders' cases" – might not be necessary when using QCA in the context of systematic reviews:

(1) Building the data table
(2) Constructing a 'truth table'

(3) Resolving contradictory configurations
(4) Boolean minimisation
(5) Consideration of the 'logical remainders' cases
(6) Interpretation

We will use the studies included in the systematic review as our sole source to collect relevant conditions to construct the initial data table. In this process of identifying relevant conditions, we will first consult the studies reporting the most and least effective interventions and extract the design features of these interventions. We will further consider the results of the second review advertised under this research call, which investigates barriers and facilitators to women's labour market participation. If the timelines of both reviews overlap, design features identified in the complementary review will be considered for the QCA of our review.

The process of extracting relevant intervention design features will apply the same consistency checks as the general data extraction process outlined above. The data extraction tool in Appendix 5 includes a separate section to record relevant design features which subsequently will be used as a first round of conditions for constructing the data table. The data table will map the relevant conditions and outcomes for each individual study. We assume to conduct a fuzzy set QCA in which interventions might be partially attributed to conditions. Following analysis of the data table, we will develop the truth table which displays the conditions, configurations, and the number of studies with membership in each configuration set (Thomas et al 2014). The truth table itself will then undergo careful analysis to investigate the logical coherence of the identified configuration. Once contradictory configurations are resolved, we will use EPPI-reviewer and its built-in Boolean minimisation algorithms to identify the most consistent and simple configurations of conditions. At this stage in the analysis we will be able to present a range of configurations, i.e. a combination of design features of labour market participation interventions, which we will interpret for relevance and further consideration in collaboration with our advisory group.

The result of the QCA are intended as an explanatory supplement to the meta-analysis findings. The QCA is designed to address objective (3) of our systematic review in case we will not be able to conduct meaningful moderator analyses due to prevailing heterogeneity in the sample of included studies.

ASSESSMENT OF PUBLICATION BIAS

All reasonable attempts will be made to reduce publication bias by searching for and including unpublished studies in the review. If sufficient studies are identified, we will test for possible publication bias using funnel plots. Given the subjectivity in assessing funnel plot symmetry, we will in addition assess publication bias and its impact on the findings of the review using the 'trim and fill' method (Duval & Tweedie 2000).

3.6 OVERALL ASSESSMENT OF QUALITY OF THE EVIDENCE

We will apply the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) tool[12] to assess the overall quality of the evidence included in the review and the strengths of the review's recommendations. GRADE evaluates the quality of the primary evidence included in the review based on a range of factors including: primary study limitations, inconsistency of the identified effect sizes, indirectness of the evidence (e.g. not evaluating final outcomes), and risk of bias ratings (Guyatt et al 2008). Quality of evidence rankings range from high to moderate, low, and very low. To assign a strength of recommendation ranking to review findings, GRADE combines the quality of evidence scores with additional variables such as uncertainty about the balance between desirable and undesirable effects, or uncertainty or variability in values and preferences. Strength of recommendation rankings are divided into 'strong' and 'weak'. We will construct a summary table for each included intervention illustrating the review's findings regarding intervention effectiveness and GRADE's rating of the quality of evidence and the strength of recommendation for each finding (Guyatt et al 2011).

3.7 TREATMENT OF QUALITATIVE DATA

We will not include qualitative research studies in our review. Where qualitative research is reported as part of included impact evaluations, we will extract information about intervention implementation and context. This information will be used to guide our grouping and analysis of included studies.

[12] http://www.gradeworkinggroup.org/

4 CONTEXTUALISATION

We will use QCA in order to contextualise the findings generated by the meta-analysis of intervention effects. This acknowledges explicitly that there are multiple intervention pathways that can lead to the same desired outcomes (i.e. increased labour market participation by women). Intervention pathways to success can be influenced by a range of factors including intervention design, contextual factors, and characteristics of the intervention. It is therefore not sufficient to rely only on the net effects of interventions identified in the meta-analysis as the sole review conclusion. Rather, the QCA allows us to seek 'causal receipts' (Thomas et al 2014) mapping different configurations of design features, contextual factors, and population characteristics and investigating the association of different configurations with intervention success. In these configurations, we are particular interested in the relative effectiveness of different design features (and their combinations) and how these features address different contextual factors. That is, our contextual analysis is primarily concerned with the ability of interventions to address contextual factors through different design features in order to increase the probability of intervention success. A contextual analysis that uses the contextual factors themselves as a unit of analysis and variable for intervention success is beyond the scope of this review as it would require a more configurative review approach[13].

Contextual variables used in the QCA will be categorised according to PROGRESS-PLUS guidelines (Welch et al 2013).

[13] We expect the second systematic review funded under this call (which investigates the barriers and facilitators to women's labour market participation) to provide this type of analysis and are interested in exploring collaboration and sharing data and findings across both reviews.

5 DISSEMINATION

We assume the audience and users of this review to be policy-makers in LMICs and international development departments worldwide, the non-governmental sector, and researchers focussed on development and gender empowerment. Identified end-users of this review so far include the South African government, particularly the Departments of Women (DoW), of Labour (DoL), and of Planning, Monitoring and Evaluation (DPME) and, in the UK, DFID. In Uganda, the Office of the Prime Minister, the Ministry of Finance, Planning and Economic Development, and the Ministry of Labour and Gender have been identified as end-users. We will involve users from the beginning of the review process with an emphasis on engagement during the protocol development, the draft reporting and interpretation of findings, and the design and implementation of the dissemination strategy. This process is already in preparation.

We have secured this strong user involvement across a range of government departments already with a representative of DoW, DPME, and DFID joining our advisory group. We are in the process of setting up individual meetings with these government departments to involve them more closely in the review. The interactive evidence interface as a tool to present the evidence map presents a key input in this engagement process. We assume that it will allow users from an early review stage onwards to directly engage with the evidence and the direction that the review should be taking to meet their needs. Our advisory group is complemented by representatives of the intergovernmental and non-governmental sectors.

The final review report will be published following funders' guidelines. We assume that the EPPI-Centre as well as DFID's R4D databases present relevant publication sources, which could be complemented by 3ie's database of systematic reviews and the Africa Evidence Network. Academic journal articles will be published in high-profile journals such as World Development. We intend to host a review dissemination workshop in September 2017. Lastly, we will also publish a two-page policy brief communicating the key findings of the systematic review. This brief will also be complemented by the production of a review infographic.

In addition, we have created a 'Diary of a Systematic Review' webpage[14] in which we document the review process to engage with a more lay audience. This page will feature regular review progress updates and reflective blogs from the review team. This online media engagement will be completed with the creation of a review hashtag to be used by the Africa Centre for Evidence Twitter account.

[14] http://www.africaevidencenetwork.org/supported-by-ace/diary-of-systematic-reviews/

6 TIMETABLE

Stage of review	Start date	End date
Preparing the protocol	01 November 2016	10 February 2017
Production of evidence map	01 November 2016	10 February 2017
Peer review of protocol (allow 2 months)	13 February 2017	24 March 2017
Searching for studies	01 November 2016 24 March 2017	30 January 2017 31 March 2017
Assessing study relevance	01 January 2017 24 March 2017	10 February 14 April 2017
Extracting data from studies	24 March 2017	02 June 2017
Assessing study quality	01 May 2017	10 June 2017
Synthesising studies	22 May 2017	11 August 2017
Preparing draft report	22 May 2017	15 September 2017
Disseminating draft report (allow 3 months)	15 September 2017	31 December 2017
Revising report	18 September 2017	3 November 2017
Submission for publication with the EPPI-Centre	06 November 2017	24 November 2017

7 REFERENCES

Berlinski, S., Galiani, S. & McEwan, P. J. (2011). Preschool and maternal labor market outcomes: evidence from a regression discontinuity design. *Economic Development and Cultural Change*, 59(2): 313-44.

Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: John Wiley & Sons.

Bohnet I. (2016). *What works: gender equality by design.* Cambridge: Belknap Press.

Brody, C., De Hoop, T., Vojtkova, M., Warnock, R., Dunbar, M., Murthy, P., Dworkin, S. (2015). Economic Self-Help Group Programs for Improving Women's Empowerment: A Systematic Review. *Campbell Systematic Reviews* 2015:19.

CGD. (2016). *Revisiting what works: women, economic empowerment, and smart design*. Washington: Center for Global Development. Available from: http://www.cgdev.org/sites/default/files/CGD-Roadmap-Update-2016.pdf

Chinen, M., de Hoop,T., Balarin, M., Alcázar, L. (2016). *Vocational and business training to increase women's participation in higher skilled occupations in low- and middle-income countries: protocol for a systematic review.* Campbell Collaboration. Available from: https://www.campbellcollaboration.org/library/women-training-participation-in-higher-valued-occupations.html

Cho, Y., Honorati, M. (2014). Entrepreneurship programs in developing countries: A meta regression analysis. *Labour Economics*, 28: 110-130.

Deeks, J. J., Altman, D. A., Bradburn, M. J. (2001). Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In M. Egger, G. Davey Smith, D. G. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (pp. 285-312). London: British Medical Journal Publishing Group.

Duval, S., Tweedie, R. (2000). Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2): 455-463.

Field, E., Jayachandran, S., Pande, R. (2010). Do Traditional institutions constrain female entrepreneurship? A field experiment on business training in India. *American Economic Review,* 100(2): 125-29.

Gibbs, A., Willan, S., Misselhorn, A., Mangoma, J. (2012). Combined structural interventions for gender equality and livelihood security: a critical review of the evidence from southern and eastern Africa and the implications for young people. *Journal of the International AIDS Society*, 15(3).

Gopalaswamy, A.K., Babu, M.S., Dash, U. (2016). *Systematic review of quantitative evidence on the impact of microfinance on the poor in South Asia*. London: EPPI-Centre, Social Science Research Unit, UCL Institute of Education, University College London.

Gustavo, A., Gadsden, P., Galiani, S., Gertler, P., Herrera, A., Kariger, P., Seira, E. (2014). *The impact of daycare on maternal labour supply and child development in Mexico.* 3ie Impact Evaluation Report 6. New Delhi: International Initiative for Impact Evaluation (3ie).

Guyatt, G.H., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, P., Schunemann, H.J. (2008). GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ: British Medical Journal*, 336(7650): 924.

Guyatt, G. Oxman A., Akl, E. A., Kunz, R., Vist, G., Brozek, J., Norris, S., Falck-Ytter, Y., Glasziou, P., deBeer H., Jaeschke R., Rind D., Meerpohl J., Dahm, P., Schünemann, H. J. (2011). GRADE guidelines: 1. Introduction to GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology,* 64*:* 383-394.

Hedges, L. (2007). Effect sizes in cluster-randomised designs. *Journal of Educational and Behavioural Statistics*, 32: 341-370.

Hedges, L. V., Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL.: Academic Press.

Hedges, L., Pigott, T. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods,* 9(4): 426-445.

Higgins, J. P. T., Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21: 1539-1558.

Higgins J. P. T., Thompson, S. G., Deeks, J. J., Altman, D.G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, 327: 557-560.

Higgins, J. P. T. & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley-Blackwell.

Ibanez, M., Vollmer, S., Khan, S., et al. (2016). *Title registration: Interventions to promote female economic empowerment in low-and middle-income countries: A systematic review of reforms to labor market, land, inheritance and participation of women in decision making*. Campbell Collaboration. Available from: https://www.campbellcollaboration.org/library/women-economic-empowerment-low-and-middle-income-countries.html

IGC. (2016). *Emerging opportunities for women in Khyber Pakhtunkhwa. Industry (Manufacturing and Construction), Mining, Tourism, and Agriculture*. International Growth Centre. Oxford, UK: IGC.

ILO. (2016). *Women at work. Trends* 2016. Geneva: ILO.

Keef, S.P., Roberts, L.A. (2004). The meta-analysis of partial effect sizes. *British Journal of Mathematical and Statistical Psychology*, 57(1): 97-129.

Lipsey, M. W., Wilson, D. (2001). *Practical meta-analysis*. London: Sage.

McKinsey Global Institute. (2015). *The power of parity: How advancing women's equality can add $12 trillion to global growth*. Available from: http://www.mckinsey.com/insights/growth/how_advancing_womens_equality_can_add_12_trillion_to_global_growth

Kabeer, N. (2012). *Women's economic empowerment and inclusive growth: labour markets and enterprise development*. SIG Working Paper 2012/1. Available from: https://www.idrc.ca/sites/default/files/sp/Documents%20EN/NK-WEE-Concept-Paper.pdf

Kahneman, D. (2011). *Thinking, fast and slow*. London: Macmillan.

Kluve, J., Puerto, S., Robalino, D.A., Romero, J.M., Rother, F., Stöterau, J., Weidenkaff, F., Witte, M. (2016). *Do Youth Employment Programs Improve Labor Market Outcomes? A Systematic Review*. IZA Discussion Paper No. 10263. Available from: http://ftp.iza.org/dp10263.pdf

Overseas Development Institute (ODI). (2016). *Women's economic empowerment. Navigating enablers and constraints*. London: ODI.

Peters, H.E., Astone, N.M., Malik, A.A., Maret, F., Heller, C. (2016). *Women's economic empowerment: A review of evidence on enablers and barriers.* Urban Institute. Available from: http://www.urban.org/research/publication/womens-economic-empowerment-review-evidence-enablers-and-barriers

Piza, C., Cravo, T., Taylor, L., Gonzalez, L., Musse, I., Furtado, I., Sierra, A.C., Abdelnour, S., (2016). *Business support for small and medium enterprises in low- and middle-income countries: a systematic review*. 3ie Systematic Review 25. London: International Initiative for Impact Evaluation (3ie).

Revenga, A., Shetty, S. (2012). Empowering Women Is Smart Economics. *Finance & Development*, 49(1): 40-43.

Rihoux, B., Ragin, C. (2009) *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) And Related Techniques*. London: Sage.

Sen, A. (2001). *Development as freedom*. Oxford, UK: Oxford Paperbacks.

Snilstveit, B., 2012. Systematic reviews: from 'bare bones' reviews to policy relevance. *Journal of development effectiveness*, 4(3): 388-408.

Sterne, J.A., Hernán, M.A., Reeves, B.C., Savović, J., Berkman, N.D., Viswanathan, M., Henry, D., Altman, D.G., Ansari, M.T., Boutron, I.,Carpenter, J.R. (2016). ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, 355: 4919.

Stewart, R., van Rooyen, C., Korth, M., Chereni, A., Rebelo Da Silva, N., de Wet, T. (2012). *Do micro-credit, micro-savings and micro-leasing serve as effective financial inclusion interventions enabling poor people, and especially women, to engage in meaningful economic opportunities in low- and middle-income countries. A systematic review of the evidence*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Thaler, R.H., Sunstein, C.R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Yale University Press: New Haven, USA.

Thomas, J., Brunton, J., Graziosi, S. (2010). *EPPI-Reviewer 4.0: Software for research synthesis.* London: Social Science Research Unit, Institute of Education, University of London.

Thomas, J., O'Mara-Eves, A., Brunton, G. (2014). Using qualitative comparative analysis (QCA) in systematic reviews of complex interventions: a worked example. *Systematic Reviews*, 20(3): 67.

Tripney, J., Hombrados, J., Newman, M., Hovish, K., Brown, C., Steinka-Fry, K., Wilkey, E. (2013). Technical and Vocational Education and Training (TVET) Interventions to Improve the Employability and Employment of Young People in Low- and Middle-Income Countries: A Systematic Review. *Campbell Systematic Reviews* 2013:9.

Tripney, J., Roulstone, A., Vigurs, C., Hogrebe, N., Schmidt, E., Stewart, R. (2015) Interventions to improve the labour market situation of adults with physical and/or sensory disabilities in low- and middle-income countries. *Campbell Systematic Reviews* 2015:20.

United Nations. (2013). *A roadmap for promoting women's economic empowerment*. United Nations. Available from: http://www.womeneconroadmap.org/sites/default/files/WEE_Roadmap_Report_Final.pdf

United Nations. (2015). *Progress of the world's women 2015-2016: transforming economies, realizing rights*. United Nations Women. Available from: http://progress.unwomen.org/en/2015/pdf/UNW_progressreport.pdf

Vaessen, J., Rivas, A., Duvendack, M., Palmer-Jones, R., Leeuw, F.L., van Gils, G., Lukach, R., Holvoet, N., Bastiaensen, J., Hombrados, J.G., Waddington, H. (2014). The effect of microcredit on women's control over household spending in developing countries: a systematic review. *Campbell Systematic Reviews*, 10(8).

Welch, V., Petticrew, M., Tugwell, P., Moher, D., O'Neill, J., Waters, E. and White, H. (2013). PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. *Revista Panamericana de Salud Pública*, 34(1): 60-67.

World Bank. (2012). *World Development Report 2012: Gender equality and development.* Washington DC: World Bank.

Yoong, J., Rabinovich, L., Diepeveen, S. (2012). *The impact of economic resource transfers to women versus men: a systematic review*. Technical report. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

8 APPENDIX

APPENDIX 1: AUTHORSHIP OF THIS REPORT

DETAILS OF ADVISORY GROUP MEMBERSHIP

Harsha Dayal	Research Director: South African National Department of Planning Monitoring and Evaluation
Thomas De Hoop	Senior Researcher: American Institute for Research
Sumana Hussain	Economic Advisor: UK Department for International Development
Marcel Korth	Senior lecturer: Witwatersrand University Senior consultant to the United Nations Office on Drugs and Crime on gender-based violence
Bernedette Muthien	Commissioner and Chair of Research & Policy Development Statutory Committee Constitutional Commission for Cultural, Religious & Language Rights South Africa
Josephilda Nhlapo-Hlope	Outcome Facilitator Social Cohesion: South African National Department of Planning Monitoring and Evaluation
Jenny Schreiner (interim)	Director-General: South African National Department of Women

DETAILS OF REVIEW GROUP MEMBERSHIP

Laurenz Langer	(Africa Centre for Evidence, University of Johannesburg)
Janice Tripney	(EPPI-Centre, University College London)
Yvonne Erasmus	(Africa Centre for Evidence, University of Johannesburg)
Natalie Tannous	(Africa Centre for Evidence, University of Johannesburg)
Charity Chisoro	(Africa Centre for Evidence, University of Johannesburg)
Mary Opondo	(Africa Centre for Evidence, University of Johannesburg)
Luvuyo Zigana	(Africa Centre for Evidence, University of Johannesburg)
Ekwaro Obuku	(African Centre for Systematic Reviews and Knowledge Translation, Makerere University)
Carina van Rooyen	(Africa Centre for Evidence, University of Johannesburg)
Ruth Stewart	(Africa Centre for Evidence, University of Johannesburg)

ACKNOWLEDGEMENTS

We would like to thank our review advisory group—Harsha Dayal, Thomas De Hoop, Sumana Hussain, Marcel Korth, Bernedette Muthien, and Josephilda Nhlapo-Hlope—for their insightful comments and inputs on the protocol. Sincere appreciation is also expressed to Alison Bullen for leading and conducting the scientific search for academic literature. We are also grateful for helpful feedback and support received from members of the EPPI-Centre: Mukdarut Bangpan, Sandy Oliver, Claire Stansfield, Dylan Kneale, Jeff Brunton, Sergio Graziosi, and James Thomas.

CONFLICTS OF INTEREST

There were no conflicts of interest in the writing of this report.

APPENDIX 2: SUMMARY OF INCLUSION AND EXCLUSION CRITERIA

CRITERIA USED FOR THE EVIDENCE MAP:

Inclusion Criteria	Exclusion Criteria
Population	
Women: The study sample must include women aged 15 years and older. The study sample must either be majority female or the study results must be disaggregated by gender. *Geographical location:* Low- or middle-income country as classified by the World Bank at time of data collection for the study.	*Women:* The study sample is not majority female or the study results are not disaggregated by gender. *Geographical location:* High-income country as classified by the World Bank at time of data collection for the study.
Employment Setting	
The study must evaluate the effects of relevant interventions applied in economic sectors with high or growing productivity *and/or* which are male-dominated. This refers to: • Commercial agriculture; • Energy (mining and quarrying, electricity, gas and water supply); • Trade; • Transportation; • Accommodation and food; • Business administration services; • Finance; • Electronics and ICT; • Maritime services; • Wood pulp and forestry; • Construction; • Manufacturing; • Higher education/Science and Technology; Studies assessing vertical labour market segregation are included regardless of the economic sector.	All other economic sectors are not relevant for inclusion.
Intervention	
We will include *any* intervention that aims to overcome the barriers to women's labour	We will exclude macro-level interventions such as investment in basic to tertiary education,

market participation in LMICs. Section 1.5 outlined the main categories of interventions that we expect to encounter in this review as well as provided examples of interventions for each category: (1) Interventions overcoming discrimination by markets & work institutions (2) Interventions overcoming constraints in access to credit, finance, and assets (3) Interventions overcoming constraints in employability (4) Interventions overcoming a lack of social capital & norms (5) Interventions overcoming a lack of economic capital (6) Interventions overcoming behavioural factors	health care, citizenship, social welfare, and economic growth which are known to benefit labour market participation rates of the general population.
Outcomes	
To be included in the evidence map, studies must evaluate the impact of interventions on one of the below three final outcomes: (1) Participation in formal or informal employment (in higher growth and/or male-dominated sectors) (2) Entrepreneurial success (following outcome 1: participation in higher growth and/or male-dominated sectors) (3) Economic empowerment (following outcome 1: participation in higher growth and/or male-dominated sectors)	All other outcomes are not eligible for inclusion, including intermediate outcomes reported in studies that do not assess one of the final outcomes.
Methods	
We will include studies using either of the following quantitative experimental or quasi-experimental study designs: (a) Designs using a random or quasi-random method of group assignment in which one of the following is true: • Units (individuals or clusters of individuals) are randomly assigned to treatment and control groups by the investigator using a fully random procedure, such as computerised random number generation;	We will exclude all other types of study designs. For example, this refers to evaluations designs such as those without a control group or without multiple data points.

• A quasi-random procedure presumed to produce comparable groups has been used by the investigators. For example, allocation by date of birth or next person to walk in the door (i.e. the method of allocation falls short of full randomisation); • Regression discontinuity designs in which participants are assigned by the investigator to intervention or control groups solely on the basis of a cut-off score on a pre-programme measure. (b) Designs employing non-random methods of assignment, in which one of the following is true: • The investigator controls group exposure and assigns participants using a non-random procedure (e.g. alphabetically by surname); • The investigator constructs the comparison group after the start of the intervention (e.g. by exploiting existing survey data); • A natural experiment in which units exposed to the treatment and control conditions are determined by nature (e.g. change in policy or divergence in practice between regions) or by other factors outside the control of the investigators); • Assignment to conditions (treatment versus comparison) is by means of self-selection by participants or by administrator selection (e.g. by welfare officials).	

CRITERIA USED FOR THE FULL SYSTEMATIC REVIEW:

Inclusion Criteria	Exclusion Criteria
Population	
Women: The study sample must include women aged 15 years and older. The study sample must either be majority female or the study results must be disaggregated by gender. *Geographical location:* Low- or middle-income country as classified by the World Bank at time of data collection for the study.	*Women:* The study sample is not majority female or the study results are not disaggregated by gender. *Geographical location:* High-income country as classified by the World Bank at time of data collection for the study.
Employment Setting	
The study must evaluate the effects of relevant interventions applied in economic sectors with high or growing productivity *and/or* which are male-dominated. This refers to: • Commercial agriculture; • Energy (mining and quarrying, electricity, gas and water supply); • Trade; • Transportation; • Accommodation and food; • Business administration services; • Finance; • Electronics and ICT; • Maritime services; • Wood pulp and forestry; • Construction; • Manufacturing; • Higher education/Science and Technology; Studies assessing vertical labour market segregation are included regardless of the economic sector.	All other economic sectors are not relevant for inclusion.
Intervention	
We will include *any* intervention that aims to overcome the barriers to women's labour market participation in LMICs. Section 1.5 outlined the main categories of interventions that we expect to encounter in this review as well as provided examples of interventions for each category:	We will exclude macro-level interventions such as investment in basic to tertiary education, health care, citizenship, social welfare, and economic growth which are known to benefit labour market participation rates of the general population.

(1) Interventions overcoming discrimination by markets & work institutions (2) Interventions overcoming constraints in access to credit, finance, and assets (3) Interventions overcoming constraints in employability (4) Interventions overcoming a lack of social capital & norms (5) Interventions overcoming a lack of economic capital (6) Interventions overcoming behavioural factors	
Outcomes	
To be included in the full systematic review, studies must evaluate the impact of interventions on the final outcome: (1) Participation in formal or informal employment (in higher growth and/or male-dominated sectors)	All other outcomes are not eligible for inclusion, including intermediate outcomes reported in studies that do not assess the final outcome.
Methods	
We will include studies using either of the following quantitative experimental or quasi-experimental study designs: (a) Designs using a random or quasi-random method of group assignment in which one of the following is true: • Units (individuals or clusters of individuals) are randomly assigned to treatment and control groups by the investigator using a fully random procedure, such as computerised random number generation; • A quasi-random procedure presumed to produce comparable groups has been used by the investigators. For example, allocation by date of birth or next person to walk in the door (i.e. the method of allocation falls short of full randomisation); • Regression discontinuity designs in which participants are assigned by the investigator to intervention or control groups solely on the basis of a cut-off score on a pre-programme measure.	We will exclude all other types of study designs. For example, this refers to evaluations designs such as those without a control group or without multiple data points.

(b) Designs employing non-random methods of assignment, in which one of the following is true: • The investigator controls group exposure and assigns participants using a non-random procedure (e.g. alphabetically by surname); • The investigator constructs the comparison group after the start of the intervention (e.g. by exploiting existing survey data); • A natural experiment in which units exposed to the treatment and control conditions are determined by nature (e.g. change in policy or divergence in practice between regions) or by other factors outside the control of the investigators); • Assignment to conditions (treatment versus comparison) is by means of self-selection by participants or by administrator selection (e.g. by welfare officials).	

APPENDIX 3: WORLD BANK CLASSIFICATION OF WORLD ECONOMIES

Region	Low-income economies	Lower-middle income economies	Upper-middle income economies
Europe and Central Asia		Armenia, Kosovo, Kyrgyz Republic, Moldova, Tajikistan, Ukraine, Uzbekistan	Albania, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Georgia, Kazakhstan, Macedonia FYR, Montenegro, Romania, Russian Federation, Serbia, Turkey, Turkmenistan
South Asia	Afghanistan, Nepal	Bangladesh, Bhutan, India, Pakistan, Sri Lanka	Maldives
Middle East and North Africa		Djibouti, Egypt, Morocco, Syrian Arab Republic, Tunisia, West Bank and Gaza, Yemen	Algeria, Iran, Iraq, Jordan, Lebanon, Libya
East Asia and Pacific	Democratic Republic of Korea	Cambodia, Indonesia, Kiribati, Lao PDR, Micronesia, Mongolia, Myanmar, Papua New Guinea, Philippines, Samoa, Solomon Islands, Timor-Leste, Tonga, Vanuatu, Vietnam	American Samoa, China, Fiji, Malaysia, Marshall Islands, Palau, Thailand, Tuvalu

Region	Low-income economies	Lower-middle income economies	Upper-middle income economies
Sub Saharan Africa	Benin, Burkina Faso, Burundi, Central African Republic, Chad, Comoros, Democratic Republic of Congo, Eritrea, Ethiopia, The Gambia, Guinea, Guinea-Bissau, Liberia, Madagascar, Malawi, Mali, Mozambique, Niger, Rwanda, Senegal, Sierra Leone, Somalia, South Sudan, Tanzania, Togo, Uganda, Zimbabwe	Cameroon, Cape Verde, Republic of Congo, Côte d'Ivoire (Ivory Coast), Ghana, Kenya, Lesotho, Mauritania, Nigeria, São Tomé and Principe, Sudan, Swaziland, Zambia	Angola, Botswana, Equatorial Guinea, Gabon, Mauritius, Namibia, South Africa
Latin America and Caribbean	Haiti	Bolivia, El Salvador, Guatemala, Honduras, Nicaragua	Argentina, Belize, Brazil, Chile[15], Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, Grenada, Guyana, Jamaica, Mexico, Panama, Paraguay, Peru, St. Lucia, St. Vincent and the Grenadines, Suriname, Uruguay[16], Venezuela

As of 1st July 2016, https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

[15] Pre 1 July 2013

[16] Pre 1 July 2013

APPENDIX 4: SEARCH STRATEGY

ACADEMIC DATABASES TO BE SEARCHED USING THE MASTER SEARCH STRING:

- Web of Science
- Econlit (EBSCO)
- ERIC (EBSCO)
- Business Source Complete (EBSCO)
- PRISMA database
- Sociological Abstracts (CSA)

KEY CONCEPTS

1. Developing Countries
2. Women
3. Type of study
4. Intervention

Search 1, 2, 3 and 4 in Title, Abstract, Keyword, Subject Heading.

Example of search string applied and piloted in Web of Science:

1 AND 2 AND 3 AND 4 = 8,372 citations (search ran 11 January 2017)

1. SEARCH FOR 'DEVELOPING COUNTRIES' TERMS

Africa OR Asia OR Caribbean OR "West Indies" OR "South America" OR "Latin America" OR "Central America" or Afghanistan OR Albania OR Algeria OR Angola OR Antigua OR Barbuda OR Argentina OR Armenia OR Aruba OR Azerbaijan OR Bahrain OR Bangladesh OR Barbados OR Benin OR Byelarus OR Byelorussian OR Belarus OR Belorussian OR Belorussia OR Belize OR Bhutan OR Bolivia OR Bosnia OR Herzegovina OR Hercegovina OR Botswana OR Brasil OR Brazil OR Bulgaria OR "Burkina Faso" OR "Burkina Fasso" OR "Upper Volta" OR Burundi OR Urundi OR Cambodia OR "Khmer Republic" OR Kampuchea OR Cameroon OR Cameroons OR Cameron OR Camerons OR "Cape Verde" OR "Central African Republic" OR CAR OR Chad OR Chile OR China OR Colombia OR Comoros OR "Comoro Islands" OR Comores OR Mayotte OR Congo OR Zaire OR "Costa Rica" OR "Cote d'Ivoire" OR "Ivory Coast" OR Croatia OR Cuba OR Cyprus OR Czechoslovakia OR "Czech Republic" OR Slovakia OR "Slovak Republic" OR Djibouti OR "French Somaliland" OR Dominica OR "Dominican Republic" OR "East Timor" OR "East Timur" OR "Timor Leste" OR Ecuador OR Egypt OR "United Arab Republic" OR "El Salvador" OR Eritrea OR Estonia OR Ethiopia OR Fiji OR Gabon OR "Gabonese Republic" OR Gambia OR Gaza OR Georgia OR Ghana OR "Gold Coast" OR Greece OR Grenada OR Guatemala OR Guinea OR Guam OR Guiana OR Guyana OR Haiti OR Honduras OR Hungary OR India OR Maldives OR Indonesia OR Iran OR Iraq OR Jamaica OR Jordan OR Kazakhstan OR Kazakh OR Kenya OR Kiribati OR Korea OR Kosovo OR Kyrgyzstan OR Kirghizia OR "Kyrgyz Republic" OR Kirghiz OR Kirgizstan OR "Lao PDR" OR Laos OR Latvia OR Lebanon OR Lesotho

OR Basutoland OR Liberia OR Libya OR Lithuania OR Macedonia OR Madagascar OR "Malagasy Republic" OR Malaysia OR Malaya OR Malay OR Sabah OR Sarawak OR Malawi OR Nyasaland OR Mali OR Malta OR "Marshall Islands" OR Mauritania OR Mauritius OR "Agalega Islands" OR Mexico OR Micronesia OR "Middle East" OR Moldova OR Moldovia OR Mongolia OR Montenegro OR Morocco OR Mozambique OR Mocambique OR Myanmar OR Myanma OR Burma OR Namibia OR Nepal OR "Netherlands Antilles" OR "New Caledonia" OR Nicaragua OR Niger OR Nigeria OR "Northern Mariana Islands" OR Oman OR Muscat OR Pakistan OR Palau OR Palestine OR Panama OR Paraguay OR Peru OR Philippines OR Philipines OR Phillipines OR Phillippines OR Poland OR Portugal OR "Puerto Rico" OR Romania OR Rumania OR Roumania OR Russia OR Russian OR Rwanda OR Ruanda OR "Saint Kitts" OR "St Kitts" OR Nevis OR "Saint Lucia" OR "St Lucia" OR "Saint Vincent" OR "St Vincent" OR Grenadines OR Samoa OR "Samoan Islands" OR "Navigator Island" OR "Navigator Islands" OR "Sao Tome" OR "Saudi Arabia" OR Senegal OR Serbia OR Montenegro OR Seychelles OR "Sierra Leone" OR Slovenia OR "Sri Lanka" OR Ceylon OR "Solomon Islands" OR Somalia OR Sudan OR Suriname OR Surinam OR Swaziland OR Syria OR Tajikistan OR Tadzhikistan OR Tadjikistan OR Tadzhik OR Tanzania OR Thailand OR Togo OR "Togolese Republic" OR Tonga OR Trinidad OR Tobago OR Tunisia OR Turkey OR Turkmenistan OR Turkmen OR Uganda OR Ukraine OR Uruguay OR "USSR" OR "Soviet Union" OR "Union of Soviet Socialist Republics" OR Uzbekistan OR Uzbek OR Vanuatu OR "New Hebrides" OR Venezuela OR Vietnam OR "Viet Nam" OR "West Bank" OR Yemen OR Yugoslavia OR Zambia OR Zimbabwe OR "developing country" OR "developing countries" OR "developing nation" OR "developing nations" OR "developing world" OR "less-developed countr*" OR "less developed countr*" OR "less-developed world" OR "less-developed world" OR "lesser-developed countr*" OR "lesser developed countr*" OR "lesser-developed nation" OR "lesser developed nation*" OR "lesser developed world" OR "lesser-developed world" OR "under-developed countr*" OR "under developed countr*" OR "under-developed nation*" OR "under developed nation*" OR "under-developed world" OR "underdeveloped world" OR "under developed world" OR "underdeveloped countr*" OR "under-developed countr*" OR "Under developed countr*" OR "under developed nation*" OR "under-developed nation*" OR "underdeveloped nation*" OR "lower middle income countr*" OR "lower middle-income countr*" OR "lower middle income nation*" OR "lower middle-income nation*" OR "upper middle-income countr*" OR "upper middle income countr*" OR "upper middle-income nation*" OR "upper middle income nation*" OR "low-income countr*" OR "low income countr*" OR "low-income nation*" OR "low income nation*" OR "lower income countr*" OR "lower-income countr*" OR "lower income nation*" OR "lower-income nation*" OR "Low- and Middle- Income countr*" OR "Low and Middle Income Countr*" OR "underserved country" OR "underserved countries" OR "underserved nation" OR "underserved nations" OR "underserved world" OR "under served country" OR "under served countries" OR "under served nation" OR "under served nations" OR "under served world" OR "deprived country" OR "deprived countries" OR "deprived nation" OR "deprived nations" OR "deprived world" OR "poor country" OR "poor countries" OR "poor nation" OR "poor nations" OR "poor world" OR "poorer country" OR "poorer countries" OR "poorer nation" OR "poorer nations" OR "poorer

world" OR "developing economy" OR "developing economies" OR "less developed economy" OR "less developed economies" OR "lesser developed economy" OR "lesser developed economies" OR "under developed economy" OR "under developed economies" OR "underdeveloped economy" OR "underdeveloped economies" OR "middle income economy" OR "middle income economies" OR "low income economy" OR "low income economies" OR "lower income economy" OR "lower income economies" OR lmic OR lmics OR "third world" OR "lami country" OR "lami countries" OR "transitional country" OR "transitional countries" LMIC OR LMICs OR LIC OR LICs OR LMICs OR LMIC OR UMICs OR UMIC) OR ("khmer" AND "republic") OR ("cape" AND "verde") OR ("central" AND "african" AND "republic")

2. SEARCH FOR 'WOMEN' TERMS

Woman OR Women OR Women's OR Mothers OR young mother OR Female OR Females OR Wife OR Wives OR 'Young Women' OR 'older girls' OR Femini* OR Gender OR Maternal OR Maternity OR daughter OR Daughters

OR

((Student OR students OR adolescents OR adolescent OR youth OR adolescence OR "young adults" OR "young adult" OR teenagers OR teenage OR teenager OR teenaged OR NEETs OR NEET OR "factory worker*" OR "wage worker*" OR employee* OR "schoolleaver*" OR "school leaver*" OR "high school graduate" OR "not in education") AND (girls OR girl OR woman OR women OR female OR females))

3. SEARCH FOR 'TYPES OF STUDIES' TERMS

Systematic review*" OR "Observational stud*" OR "Longitudinal stud*" OR "impact stud*" OR "Impact evaluation*" OR "comparison stud*" OR "Longitudinal Analysis*" OR "impact analysis" OR "observational analysis" OR "non-comparison stud*" OR Imp-Act OR "random* control* trial*" OR "random* trial*" OR "comparison group*" OR "control group*" OR RCT OR experiment* OR "program* evaluation*" OR "experimental control*" OR "pilot scheme*" OR "Pilot stud*" OR "pilot program*" OR "time series" OR "comparative analysis" OR Quasi-experiment* OR post-test* OR posttest* OR "post test*" OR pre-test* OR pretest* OR "pre test*" OR "project apprais*" OR "cluster random* trial*" OR "propensity score matching" OR PSM Or "regression discontinuity design" OR RDD OR "difference in difference*" OR DID OR meta-analy* OR "meta analy*" OR "control* random* trial*" OR "case control*" OR matching OR "interrupted time series" OR "random* allocation*" OR "instrumental variable*" OR "research synthesis" OR "scoping review" OR "rapid evidence assessment*" OR "systematic literature review*" OR "Rapid Review*" OR QED OR "group design" OR "intervention group" OR "intervention groups" OR "controlled stud*" OR "comparative stud*" OR "between group" OR "between groups" OR "group difference*" OR "Quasi-experiment*" OR "experimental group*" OR "control community" OR "intervention community" OR "intervention communities" OR "control communities" OR "intervention

condition*" OR "control condition*" OR "controlled condition*" OR "control participant*" OR "experimental condition*")

OR

(Design N3 quantitative) OR (Design N3 “comparison group”) OR (Design N3 counterfactual) OR (Design N3 “counter factual”) OR (Design N3 counter-factual) OR (Design N3 experiment*) OR (Study N3 quantitative) OR (Study N3 “comparison group”) OR (Study N3 counterfactual) OR (Study N3 “counter factual”) OR (Study N3 “counter-factual”) OR (Study N3 experiment*) OR (Analysis N3 quantitative) OR (Analysis N3 “comparison group”) OR (Analysis N3 counterfactual) OR (Analysis N3 “counter factual”) OR (Analysis N3 counter-factual) OR (Analysis N3 experiment*) OR (random* N3 allocat*) OR (effectiveness N3 intervention*) OR (before N2 after study) OR (matched N3 group)

4. SEARCH FOR ‘INTERVENTION’ TERMS

“Job Market” OR “labour Power” OR “Labor Power” OR Employee* OR “Labor Market” OR “Labour Market” OR “Labour Force” OR “labor Force” OR “Labour Economy” OR “Human Capital” OR “Economic Capital” OR “Labour Demand” OR “Labour Supply” OR “Wage Labour” OR “Wage labor” OR “employment reservation” OR “labour reservation” OR “Job reserv*” OR “Wage Gap*” OR “Wage differential*” OR “Low wage*” OR “Work* Condition*” OR “labour incentive*” OR “monetary incentive*” OR “financial incentive*” OR “minimum wage” OR “minimum wages” OR “employment tax*” OR “labour tax*” OR “labor tax*” OR “job insecurity” OR “employment security” OR “labour market regulation*” OR “labor market regulation*” OR “labour law*” OR “labor law*” OR “labour regulation*” OR “labor regulation*” OR “labour polic*” OR “labor polic*” OR “Employment polic*” OR “employment law” OR “employment laws” OR “labour reform” OR “labor reform” OR “labour reforms” OR “labor reforms” OR “job security” OR “labor standard*” OR “labour standard*” OR “labor code*” OR “labour code*” OR “labor legislation” OR “labour legislation” OR “Labour Supply” OR “Labour market participation” OR “Labour Force Participation” OR “labor force participation” OR “Labour Market” OR “Labor Market” OR “labor market participation” OR “Employment Trend*” OR “Employment Pattern*” OR “Employment Potential” OR “Economic Conditions” OR “Economic Development” OR Infrastructure OR “Economic Opportunit*” OR “economic right*” OR “Public spending” OR “Public Finance” OR “finance law” OR “finance polic*” OR “Public Works” OR “Job creation” OR “employment guarantee*” OR “Trade polic*” OR “trade law*” OR “export polic*” OR “export law*” OR “Public Sector” OR “tariff* reduction*” OR “tariff* change*” OR “Trad* Openness” OR “Trade reform*” OR “Trade liberalisation” OR “Trade liberalization” OR “Preferential Trade Agreement*” OR “Trade PTA*” OR “Free Trade Agreement*” OR “Trade FTA*” OR “special economic zone*” OR “export processing zone*” OR “free trade zone*” OR “free zone*” OR “foreign trade zone*” OR “industrial park*” OR “industrial estate*” OR “urban enterprise zone*” OR FTZ* OR EPZ* OR “Monetary polic*” OR “Public Spending” OR “Fiscal Stimulation” OR SME OR SMME OR “small enterprise*” OR start-up OR “medium enterprise*” OR “medium siz* enterprise*” OR “High Skill*” OR “formal Enterprise*” OR

“small and medium-sized enterprise*” OR “Labour Economy” OR “Human Capital” OR “Economic Capital” OR “Labour Demand” OR “Wage Labour” OR “Wage labor” OR “labour Power” OR “Labor Power” OR “labour quota*” OR “employment quota*” OR “employee quota*” OR “gender quota*” OR “affirmative Action” OR “Economic Empowerment” OR “Procurement norm*” OR “procurement Standard*” OR “Tax regulation*” OR “Tax certification” OR “tax simplification” OR “Income tax credit*” OR “demand-side” OR “business development” OR “Financial service*” OR “Personal finance” OR “Personal financial” OR “Domestic investment*” OR “Personal wealth” OR Microenterpris* OR “Micro-enterpris*” OR “Productive resource*” OR “micro credit” OR Micro-lease OR Monetary OR Bank* OR Insurance OR “Economic participation” OR “Tax penalt*” OR “Second Earner*” OR “Tax incentive*” OR “Tax Regulation*” OR “Tax Certification” OR franchise OR lease OR “market access” OR “value chain*” OR “technical and vocational education and training” OR “Technical Education” OR “Vocational Training” OR “Vocational Education” OR “on the job training” OR “on-the-job training” OR “Technical and Vocational Voucher*” OR “Job Placement” OR “Retraining” OR “Re-training” OR Employability OR “Employment Support” OR “Training Support” OR “Capacity Building” OR “Appropriate Skills” OR “Role Models” OR “Business Network*” OR “labour network*” OR “enterprise network*” OR “industry network*” OR “Career option*” OR “Career Progression” OR “Business Leadership” OR “Job Counselling” OR “Career Guidance” OR “Career counselling” OR “job fair” OR “mentor*” OR “business advisory” OR “Business Skill*” OR “entrepreneur* training” OR internship* OR intern OR interns OR “job placement” OR “Soft Skills” OR Apprentic* OR “labour market information” OR “Job search*” OR “Job Seek” OR “Active Labour Market Polic*” OR ALMP OR “job matching” OR “employment support” OR “educational voucher*” OR “entrepreneur* training” OR “self-help group*” OR “empowerment group*” OR “peer support” OR “Participatory learning” OR Traineeships OR traineeship OR TVET OR “technical training” OR “tech* prep*” OR “technician education” OR “technical stud*” OR “technical cent*” OR “technical school*” OR “technical course*” OR “technical program*” OR “technical college*” OR “technical degree*” OR “technical diploma*” OR “technical qualification*” OR “vocational stud*” OR “vocational retraining” OR “vocational work experience” OR “vocational cent*” OR “vocational school*” OR “vocational course*” OR “vocational program*” OR “vocational college*” OR “vocational degree*” OR “vocational diploma*” OR “vocational qualification*” OR “vocational framework*” OR “industrial education” OR traineeship* OR “trade course*” OR “job training” OR “job-related training” OR “job-site training” OR “in-service training” OR “retraining” OR “training program*” OR “skill* training” OR “skill* development” OR “staff development” OR “work place learning” OR “work based learning” OR “work related learning” OR “work* education” or “work place education” OR “work based education” OR “work related education” OR “work* training” OR “work place training” OR “work based training” OR “work related training” OR “work* program*” OR “work place program*” OR “work based program*” OR “work related program*” OR “work experience program*” OR “workforce development” OR “labour market” or “labor market” OR “employment based education” OR “employment based training” OR “employ* training” OR “employ* education” OR “employ* development program*” OR “employ* program*” OR “employ* course*” OR “unemploy* training” OR

“training for unemployed” OR “training for the unemployed” OR “occupation* education” OR “occupation* training” OR “occupation* program*” OR “occupation* course*” OR “business education” OR “office occupations education” OR “contract training” OR “school to career program*” OR “school to work program*” OR “career* education” OR “youth program*” OR “company training” OR “company-based learning” OR “investment in training” OR “Occupational Mobility” OR “Business Training” OR “Business Loan*” OR “Business Grant*” OR “Business ownership” OR “Business Credit” OR Asset* OR “land tenure” OR “Property” OR “title deeds” OR Livestock OR Capital OR micro-Saving* OR “Micro saving*” OR “flexible work” OR “flexible Hours” OR “Flexible employment” OR “Flexible Labour” OR “Flexible labor” OR “local productive system*” OR “collective action*” OR “economic cluster*” OR “industry cluster*” OR “enterprise cluster*” OR “matching grant*” OR formalization OR “business environment” OR “property registration” OR “regulatory framework*” OR “mandatory employee benefit*” OR “unemployment benefit” OR “unemployment benefits” OR “payroll tax” OR “payroll taxes” OR “hour restrictions” OR “hiring rigidity” OR “hiring rigidities” OR “firing rigidity” OR “firing rigidities” OR “termination benefit” OR “termination benefits” OR “job insecurity” OR “employment security” OR “labour rigidity” OR “labor rigidity” OR “labor rigidities” OR “labour rigidities” OR “collective bargaining” OR “labour market regulation*” OR “labor market regulation*” OR “labour regulation*” OR “maternity protection” OR “employment search*” OR “work search*” OR “employment seek*” OR “work seek*” OR “Financ* literacy” OR “financ* inclusion” OR “financial awareness” OR “financial capabilit” OR “financial competence” OR “financial education” OR “financial knowledge” OR “financial literacy” OR “gender-responsive budget*” OR “female-owned enterprise*” OR “women-owned enterprise*” OR “female-owned business” OR “women-owned business” OR “female-owned industry*” OR “women-owned industr*” OR “female-owned factor*” OR “women-owned factor*” OR “business practice*” OR “labour practice*” OR “employment practice*” OR “procurement practice*” OR “business standard*” OR “labour standard*” OR “employment standard*” OR “procurement standard*” OR “business norm*” OR “labour norm*” OR “employment norm*” OR “procurement norm*” OR “Combined structural intervention*” OR “bundled service*” OR “livelihood program*” OR “poverty graduation program*” OR microfinance OR micro-finance OR “micro finance” OR microcredit OR micro-credit OR “micro loan” OR “micro-loan” OR microloan OR microloans OR “micro-loans” OR funding OR “Financial Support” OR Transport OR Sanitation OR Water OR telecommunications OR telecommunication OR “Business Technolog*” OR Electricity OR Electrification OR “clean stoves” OR “clean energy” OR “household fuel” OR power OR ICT OR “information and communication technolog*” OR “mobile technolog*” OR phone* OR telephone* OR credit OR Saving* OR Finance OR lending OR “fair trade” OR “ethical trade” OR “sustainability standard*” OR “Corporate social responsibility” OR CSR OR “corporate responsibility” OR “code of conduct*” OR “Gender Mainstream*” OR “gender awareness” OR “outreach program*” Or “advocacy Program*” OR “awareness campaign*” OR “awareness raising” OR Empowerment OR “Support group*” OR “advocacy group*” OR “Job Market” OR scholarship* OR subvention* OR stipend* OR donation OR bursary OR bursaries OR “tuition relief” OR “merit aid” OR “merit based aid” OR “merit-based aid” OR “merit award” OR

advisory OR tutor* OR Loan* OR Grant* OR Credit

OR

(budget* OR “resource allocation” OR “public fund*”) AND (gender)

OR

(land OR property) AND (right* OR conversion OR freehold OR titl* OR codification OR recognition OR customary OR certification)

OR

(career OR skill* OR work OR Performance) AND (Chang* or increas* or rise* or rising or rose or rais* or augment* or growth* or grow* or grew or improv* or gain* or motivat* or promot* or encourag* or enhanc* or boost* or achiev* or success* or succeed* or accomplish* or thrive* or thriving or achiev* or attain OR enhance OR Upgrade)

OR

(Leave OR Care) AND (Matern* OR Patern* OR parent* OR Child OR Elder* OR disab*)

OR

(Work* OR employment OR Business OR Unemployment OR career OR employee OR job OR profession OR occupation) AND (Violence OR abuse OR exploitation OR harassment OR equity OR equality OR childcare OR “infant care” OR “child daycare” OR daycare OR “day care” OR nursery OR nurseries OR “nursery school” OR “nursery schools” OR pre-school OR pre-schools OR kindergarten OR “family caregiving,” OR “informal caregiving,” OR “unpaid caregiving” OR “Social Constraint*” OR Norm* OR fairness OR Inequality OR Discrimin* OR Dispar* OR “Self Esteem” OR self-esteem OR Self Confidence OR self-confidence OR motivation OR Equity OR Mainstreaming OR Exploit* OR equality OR information OR skill* OR training OR coaching OR Empowerment)

OR

(Pay* OR Remuneration OR Salar* OR Benefits OR Incentive* OR Financial OR Money OR Monetary OR Reward* OR Wage* OR Bonus OR Pension OR earning*) AND (change* OR Increase* OR Rise* OR Augmentation* OR Grow*)

OR

(subsid* OR subsidy OR subsidies OR subsidized OR subsidised) AND (Wage* OR Labour OR Employment)

OR

(Work* OR employment OR Business OR Unemployment OR career OR employee OR job) (Household OR residential OR domestic) AND (energy OR fuel) AND (choice OR switch OR switching)

OR

(industrialization OR industrialisation OR "industrial policy" OR "industrial policies" OR "investment policy" OR "investment policies" OR "domestic investment*" OR "foreign investment*" OR "foreign direct investment*" OR "fiscal policy" OR "fiscal policies" OR "monetary policy" OR "monetary policies")

GREY LITERATURE SEARCH STRATEGY

IMPACT EVALUATION AND SYSTEMATIC REVIEW REPOSITORIES

- UN Women Economic Empowerment: A road map – Database of Empirical Evaluations (http://www.womeneconroadmap.org/the_database)
- UN Gender Equality Evaluation Portal (http://genderevaluation.unwomen.org/en)
- 3ie Database of Systematic Reviews (http://www.3ieimpact.org/evidence/systematic-reviews/)
- enGENDER IMPACT A Gateway to Gender-Related Impact Evaluations (http://www.worldbank.org/en/topic/gender/publication/engender-impact-a-gateway-to-gender-related-impact-evaluations)
- 3ie Registry for International Development Impact Evaluations (RIDIE) (http://ridie.3ieimpact.org/)
- 3ie Register of Impact Evaluation Published Studies (RIEPS) (http://www.3ieimpact.org/evidence/impact-evaluations/)
- Campbell Collaboration Library: systematic reviews (http://www.campbellcollaboration.org)
- EPPI-Centre Systematic Reviews Library (https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=56)

SPECIALISED DATABASES

- Abdul Latif Jameel Poverty Action Lab (J-PAL) Evaluation and Publication Database (http://www.povertyactionlab.org/)
- ELDIS (http://www.eldis.org/)
- Innovations for Poverty Action (IPA) Database (http://www.poverty-action.org/project-evaluations/search)
- International Centre for Research on Women (http://www.icrw.org/)
- Labordoc (ILO) (http://labordoc.ilo.org/)
- RePEc (Research Papers in Economics)/IDEAS Economics and Finance Research: (http://ideas.repec.org/)

- Research for Development (http://r4d.dfid.gov.uk/)
- USAID Development Experience Clearinghouse (https://dec.usaid.gov/)
- World Bank Poverty Impact Evaluations Database (http://www1.worldbank.org/prem/poverty/ie/evaluationdb.htm)
- World Bank Independent Evaluation Group (IEG) (http://ieg.worldbankgroup.org)
- Youth Employment Inventory (YEI) (http://www.youth-employment-inventory.org/)

ORGANISATIONAL WEBSITES

- African Development Bank Evaluation Reports (http://www.afdb.org/en/documents/evaluation-reports)
- Asian Development Bank (ADB) Evaluation Resources (http://www.adb.org/site/evaluation/resources)
- Bill & Melinda Gates Foundation (http://www.gatesfoundation.org/)
- BRAC (http://www.brac.net/)
- Centre for Global Development (http://www.cgdev.org/)
- Ford Foundation (https://www.fordfound.org/)
- Havard Women and Public Policy Program (http://wappp.hks.harvard.edu/about-wappp)
- Hewlett Foundation (http://www.hewlett.org/)
- Inter-American Development Bank Office of Evaluation and Oversight (http://www.iadb.org/en/office-of-evaluation-and-oversight)
- International Growth Centre (http://www.theigc.org/)
- Institute of Labour Economics (IZA) (http://www.iza.org)
- Institute of Development Studies (IDS) (http://www.ids.ac.uk/)
- Millennium Challenge Corporation (MCC) (http://www.mcc.gov/pages/results/evaluations)
- National Bureau of Economic Research (NBER) (http://www.nber.org)
- Overseas Development Institute (ODI) (http://www.odi.org.uk/)
- Oxfam (https://www.oxfam.org/)
- Poverty and Economic Policy Research Network (PEP) (Project List http://www.pep-net.org/publications/pep-projects/)
- Self Employed Women's Association (SEWA) (http://www.sewa.org/)
- UNDP International Policy Centre for Inclusive Growth (IPC-IG) (http://www.ipc-undp.org/)
- UNESCODoc (http://www.unesco.org/new/en/unesco/resources/online-materials/publications/unesdoc-database/)
- World Bank Labor Markets (http://www.worldbank.org/en/topic/labormarkets)

BACKWARD CITATION TRACKING

We will screen the reference list of all included studies as well as existing systematic reviews for relevant studies. The following reviews and evidence maps will be searched in this way:

- ILO What Works In Youth Employment Evidence Map (http://www.wwinye.org/wwinye/evidence-gap-map)
- 3ie Youth & Transferable Skills Evidence Gap Map (http://gapmaps.3ieimpact.org/evidence-maps/youth-transferable-skills-evidence-gap-map)
- Brody, Carinne, et al. 2015. "Economic Self-Help Group Programs for Improving Women's Empowerment: A Systematic Review." Campbell Systematic Reviews 11.19.
- Buvinic, M., Furst-Nichols, R. and Pryor, E.C., 2013. "A Roadmap for promoting women's economic empowerment." United Nations Foundation and ExxonMobil Foundation.
- Buvinic, M. and O'Donnell, M., 2016. "Revisiting What Works: Women, Economic Empowerment and Smart Design."
- Kluve, J., et al. 2016. "Do Youth Employment Programs Improve Labor Market Outcomes? A Systematic Review."
- McKenzie, D. and Woodruff, C., 2013. "What are we learning from business training and entrepreneurship evaluations around the developing world?". The World Bank Research Observer, p.lkt007.
- ODI. 2014. "Review of evaluation approaches and methods used by interventions on women and girls' economic empowerment."
- ODI. 2016. "Women' s economic empowerment—Navigating enablers and constraints."
- Peters et al 2016. "Women's Economic Empowerment: A Review of Evidence on Enablers and Barriers."
- Piza, C, et al. 2016. "Business support for small and medium enterprises in low- and middle-income countries: a systematic review". 3ie Systematic Review 25. London: International Initiative for Impact Evaluation (3ie).
- Tripney, J., et al. 2013. "Technical and Vocational Education and Training (TVET) Interventions to Improve the Employability and Employment of Young People in Low-and Middle-Income Countries: A Systematic Review". Campbell Systematic Reviews 2013: 9. Campbell Collaboration.

FORWARD CITATION TRACKING

We will run a citation search in Google Scholar to identify literature that has cited the studies included in our review. All citations for each included study will be screened.

HAND-SEARCHING OF SPECIALISED JOURNALS

- Economic Development and Cultural Change
- Feminist Economics
- Journal of Development Economics
- Journal of International Development
- World Development

APPENDIX 5: DATA EXTRACTION TOOL

This tool will be translated into a coding set on EPPI-reviewer allowing for additional sub-levels of codes.

Questions	Answers
Section A: Administration	
Name of reviewer	Details (specify)
Linked reports	None / not known
	Linked (specify)
	Unclear (specify)
Language of report *If more than one report, choose the main report.*	English
	Other (specify)
Section B: Study characteristics	
Form of publication *If more than one report, choose the main report.*	Journal article
	Grey Literature (specify)
	Dissertation/thesis (specify)
	Other (specify)
Year of publication *If more than one report, choose the main report.*	1990-1994 (specify)
	1995-1999 (specify)
	2000-2004 (specify)
	2005-2009 (specify)
	2010-2014 (specify)
	2015-2017 (specify)
Broad aims of the study / research question	Explicitly stated (specify)
	Implicit (specify)
	Unclear/not stated (specify)
Study funding	Government or government-related agency (specify)
	Donor country agency
	Non-governmental organisation (NGO)/non-profit (specify)

Questions	Answers
	Multilateral agency (e.g. development bank or WHO) (specify)
	Academic/research institution (specify)
	Employer (specify)
	Other (specify)
	Unclear/not stated (specify)
When was the study conducted?	Initial year (specify)
	Final year (specify)
	Unclear/not stated (specify)
Section C: Subject characteristics	
Region *Details about the country/countries where programme is implemented are collected in Section D.*	Low-income country
	Lower-middle income country
	Upper-middle income country
	Africa
	Asia
	Europe
	Latin America & Caribbean
	Oceania
Region classification	High fertility agrarian societies
	Declining fertility urbanising societies
	Declining fertility formalising economies
	Ageing societies
Total number of study participants *This question refers to the number of subjects who were originally assigned to the treatment and control groups. Later in the tool there will be a question about*	Total 250 or less (specify)
	Total 251-500 (specify)
	Total 500 or more (specify)
	Unclear/not stated (specify)

Questions	Answers
the number of subjects who were actually observed/measured. *Typically you can use baseline sample sizes for the assigned N, but the authors may have removed the subjects with incomplete data. In this event, indicate that the figure you provide is approximate.* *Report total N.*	
Age *Select all that apply.*	Working aged adults aged 15-64 years (specify)
	Older adults aged 65+ (specify)
	Youth (15-35)
	Unclear/not stated (specify)
Is the study focussed specifically on young people?	Yes (specify)
	No
	Unclear (specify)
Is the study focussed specifically on transitions from school (incl out-of-school) to work??	Yes (specify)
	No
	Unclear (specify)
Is the study focussed specifically on returning to work after childbirth?	Yes (specify)
	No
	Unclear (specify)
Is the study focussed specifically on migrating from rural to urban areas?	Yes (specify)
	No
	Unclear (specify)
Is the study focussed specifically on transitions from unemployment back into employment?	Yes (specify)
	No

Questions	Answers
	Unclear (specify)
Is the study focussed specifically on transitions from farm labour to non-farm labour	Yes (specify)
	No
	Unclear (specify)
Is the study focussed specifically on transitions from informality to formality?	Yes (specify)
	No
	Unclear (specify)
Is the study focussed specifically on transitions from conflict/trauma back into society?	Yes (specify)
	No
	Unclear (specify)
Sex	Females only
	Mixed (specify)
	Unclear (specify)
Descriptive characteristics of the sample	Years of schooling (specify)
	Income indicators (specify)
	Asset indicators (specify)
	Social capital indicators (specify)
	Employment indicators (specify)
Based on the above, identify the sample as per UN definition:	Very poor women
	Poor women
	Non-poor women
Other useful information about study participants.	Details (specify)
	None
Section D: Intervention characteristics	

Questions	Answers
Formal name (if any)	Not applicable (no formal name)
	Details (specify)
	Unclear (specify)
Scale (or availability) of the intervention *(specific to this particular study)*	International
	National
	Provincial/Regional (specify)
	Local (specify)
	Other (specify)
	Unclear/not stated (specify)
Which level of social organisation is targeted?	Individuals
	Households
	Community groups/collectives
	Entire communities/villages
	Others (specify)
Maturity of the intervention	Pilot (specify)
	Iteration following pilot/prior experience
	Scale up
Country or countries where the intervention is implemented *(specific to this particular study)*	Details (specify)
	Unclear (specify)
Area of country where the intervention is implemented (e.g. particular regions or cities) *(specific to this particular study)*	Details (specify)
	Unclear (specify)
Primary location *(specific to this particular study)*	Primarily urban
	Primarily rural
	Both
	Unclear/not stated (specify)
Type of intervention	Interventions to balance work & family responsibilities
	Increase women's financial returns
	Changing business culture/practice

Questions	Answers
	Macroeconomic changes
	Provision of infrastructure
	Microfinance
	Cash transfers
	Economic assets
	Changes to land titles, business ownership, inheritance
	Bundled services/combined structural interventions
	Interventions to provide education/skills for women
	Interventions to provide access to economic opportunities
	Interventions to provide work experience for women
	Interventions to provide support to female businesses and entrepreneurs
	Social organisation
	Changes in norms and attitudes
	Gender sensitive design
	Behavioural nudges
Barriers addressed by the intervention	Discrimination by markets and work institutions
	Constraints in access to credit, finance and assets
	Constraints in employability & entrepreneurship
	Lack of social capital and norms
	Behavioural (social & cognitive) barriers
Does the paper refer to one or more specific theories for how the intervention should work?	Yes (specify)
	No
	Unclear/not stated (specify)
Is the 'treatment' a single 'activity' or a combination of activities? *If multi-component (also known as bundled or combined interventions), provide details of the different components. Multi-component interventions will typically target a number of different barriers/constraints faced by recipients.*	Single
	Multi-component (specify)
	Unclear/not stated (specify)

Questions	Answers
Dates of operation *Include both start and ending dates of the intervention where these are available, and where relevant note if it is ongoing.*	Not applicable (specify)
	Details (specify)
	Unclear/not stated (specify)
Who funded the intervention?	Government or government-related agency (specify)
	Donor country agency (e.g. DIFD) (specify)
	Non-governmental organisation (NGO)/non-profit (specify)
	Multilateral agency (e.g. development bank or WHO) (specify)
	Individual donor (e.g. charitable foundations or private sector)
	Academic/research institution (specify)
	Employer of beneficiaries (specify)
	Beneficiaries
	Other (specify)
	Unclear/not stated (specify)
Who designed the intervention?	Government or government-related agency (specify)
	Donor country agency (e.g. DIFD) (specify)
	Non-governmental organisation (NGO)/non-profit (specify)
	Multilateral agency (e.g. development bank or WHO) (specify)
	Individual donor (e.g. charitable foundations or private sector)
	Academic/research institution (specify)
	Employer of beneficiaries (specify)
	Beneficiaries
	Other (specify)
	Unclear/not stated (specify)
Who implemented the intervention?	Government or government-related agency (specify)
	Donor country agency (specify)
	Non-governmental organisation (NGO)/non-profit (specify)

Questions	Answers
This question is not referring to the delivery of the intervention (see below for question about service provision).	Multilateral agency (e.g. development bank or WHO) (specify)
	Individual donor (e.g. charitable foundations or private sector)
	Academic/research institution (specify)
	Employer of beneficiaries (specify)
	Other (specify)
	Unclear (specify)
Political/economic/social/cultural context (at time of intervention)	Details (specify)
	Unclear/not stated (specify)
Target group of intervention *Select all that apply, and only if targeting of the intervention is specifically mentioned (i.e. the study may focus on a specific population such as the low paid, but this is not what this question relates to).*	Not targeted specifically at any group
	Women
	Women returning to work after childbirth
	People within a certain age range, e.g. under 30 years (specify)
	Young people making transition from school (out of school) to work
	People within a specific geographical location (specify)
	People with disabilities (specify)
	People within certain ethnic groups (specify)
	People with low education (primary or lower) (specify)
	People with secondary education (or equivalent) (specify)
	People with higher education (above secondary) (specify)
	People on low incomes/disadvantaged (specify)
	People in urban locations (specify)
	People in rural locations (specify)
	People migrating from rural to urban areas (specify)
	People unemployed at intervention start
	People who lost employment, looking to return to work
	People already employed/entrepreneur at intervention start (specify)
	First-time jobseekers only (specify)

Questions	Answers
	People of a particular religion
	People focussed on a particular occupation
	Groups stated as vulnerable
	Other (specify)
	Unclear/not stated (specify)
Descriptive characteristics of the sample	Age (specify)
	Years of schooling (specify)
	Income indicators (specify)
	Asset indicators (specify)
	Social capital indicators (specify)
Based on the above, identify the sample as per UN definition:	Very poor women
	Poor women
	Non-poor women
Does the intervention design include gender considerations? *The intervention may or may not have been targeted specifically at women.*	No
	Yes (specify)
	Taking into consideration travel constraints (e.g. distance to homes)
	Taking into consideration time constraints (e.g. when the programme is offered)
	Taking into consideration caregiving constraints
	Taking into consideration the gender of programme implementers
	Changes to social / professional norms
	Quota/reservation approaches to ensure women's participation
	Interventions working through subjective economic empowerment (e.g. self-reliance/esteem)
	Increased risk taking
	Interventions addressing cognitive and social determinants of economic behaviour (e.g. Decision-making support; protection from external pressures)

Questions	Answers
	Other features supporting a gender-sensitive approach/addressing gender biases embedded in organisations/working environment (specify)
	Unclear/not stated (specify)
Does the intervention apply other design features?	Interventions applying explicit behavioural designs (e.g. commitment devices, reminders)
	Financial input from employer
	Demand-driven approaches (ie demands by markets)
	Bundling of interventions
Does the intervention design focus specifically on improving women's employment in higher growth/male-dominated sectors, and if so which ones?	No
	Manufacturing
	Trade
	Construction
	Energy (mining & quarrying, electricity, gas & water supply)
	Wood pulp and forestry
	Transportation
	Accommodation & food
	Business administration services
	Electronics and ICT
	Finance
	Commercial agriculture
	Maritime services
	Higher education/Science & Technology
	Unclear
	Other (specify)
Does the intervention design include awareness training about the intervention to targeted participants?	Not applicable (no targeting of the intervention)
	Yes (awareness training available)
	No
	Unclear/not stated (specify)
What was the take-up of the intervention?	Details (specify)
	Unclear/not stated

Questions	Answers
Is there a specific selection process? *In other words, is eligibility conditional on (targeted) individuals meeting further requirements, such as passing an interview or test?*	No (any woman/any woman within targeted population is eligible)
	Yes (specify)
	Unclear/not stated (specify)
Setting or site (for delivery of service/s)	Not applicable
	Classroom (college or training centre) (specify)
	Online (distance learning) (specify)
	Workplace (specify)
	Community site (specify)
	Other (specify)
	Mixed/multiple sites (specify)
	Unclear/not stated (specify)
Service provider/s	Not applicable (specify)
	Public institution/contractor (specify)
	Private institution/contractor (specify)
	NGO/non-profit (specify)
	Other (specify)
	Unclear/not stated (specify)
Payment system (to service provider/s)	Not applicable (specify)
	Lump-sum budget (specify)
	Payment for services delivered (specify)
	Payment by outcomes (specify)
	Other (specify)
	Unclear/not stated (specify)
Duration of treatment: overall average duration a single cohort stays in the programme	Not applicable (specify)
	One day or less (specify)
	One day to 1 week (specify)
	One week (and 1 day) to 1 month (specify)
	One month (and 1 day) to 3 months (specify)
	Three months (and 1 day) to 6 months (specify)

Questions	Answers
	Six months (and 1 day) to 1 year (specify)
	One year (and 1 day) to 2 years (specify)
	Two years (and 1 day) to 3 years (specify)
	Three years (and 1 day) to 5 years (specify)
	More than 5 years (specify)
	Other (please specify)
	Unclear/not stated (specify)
Frequency of contact between beneficiaries and provider or treatment activity	Less than weekly
	Once a week
	3-4 times a week
	Continuous
	Other (specify)
	Unclear/not stated (specify)
Intervention dosage: hours per week	Not applicable (specify)
	1-10 hours (specify)
	11-19 hours (specify)
	Over 20 hours (specify)
	Unclear/not stated (specify)
Other intervention dosage:	Size of loan/cash transfer/asset (specify)
	Other (specify)
Were incentives provided to intervention participants?	Yes, monetary benefits (e.g. stipend, transport allowance)
	Yes, non-monetary benefits (e.g. transport, childcare)
	No
	Unclear/not stated
Implementation of the intervention: was this monitored to assess whether this was delivered as intended?	Yes (monitored by programme staff)
	Yes (monitored by researcher)
	No
	Unclear/not stated (specify)
Were there any implementation problems?	Yes (specify)
	No

Questions	Answers
For example, high drop outs, erratic attendance.	Unclear/not stated (specify)
Role of study funder *Consider, for example, whether the funder was also involved in the design and delivery of the intervention, or was a member of the research team.*	Not applicable (details of study funding not reported)
	Not independent (specify)
	Independent
	Unclear/not stated (specify)
Role of evaluators *Consider, for example, whether the evaluator was also involved in the design and delivery of the intervention.*	Not stated
	Not independent (specify)
	Independent
	Unclear (specify)
Other relevant information about the programme (if any)	Details
Intervention (if any) received by the control/comparison group	Not applicable (time-series design, so only one group)
	No treatment
	Treatment as usual (specify)
	Alternative intervention (specify)
	Other (specify)
	Unclear/not stated (specify)
Section E: Key methodological characteristics	
Study design according to publication	Specify:
Unit of group assignment	Not applicable
	Individuals
	Groupings (clusters) of individuals (specify)
	Programme area, region, etc. (specify)
	Unclear/not stated (specify)
Method of group assignment *This question focuses on the initial method of assignment to groups. For*	Random (RCT design, with allocation done randomly after matching, blocking, stratification, etc.)
	Random (RCT design without matching, etc.)
	Quasi-random (e.g. investigator allocates by date of birth)

Questions	Answers
non-random methods of assignment, please indicate whether: • *the investigator controlled group exposure,* • *participants self-selected into the programme,* • *assignment was by means of administrator selection, or* • *study was a natural experiment (where exposure to the treatment and control conditions are determined by nature).*	Quasi-random (Regression Discontinuity Design)
	Non-random, but matched and/or statistically controlled (specify)
	Unclear (specify)
If study used matching, what variables are used? *Matching can be done on the propensity score or covariates. Check if study matched on (i) pre-test measures of some or all variables used later as outcome measures and (ii) other characteristics, such as demographic variables. Make a note if matching is done on endline characteristics only.*	Not applicable
	Details (specify)
	Unclear/not stated (specify)
If study used statistical controls, what variables are used? *Check if study controlled for (i) pre-test variables (i.e. measures of a dependent variable taken prior to treatment) and (ii) other characteristics, such as demographic variables.*	Not applicable
	Details (specify)
	Unclear/not stated (specify)
Proportion of sample with missing/incomplete data	No missing/incomplete data
	1-20% (specify)
	21-30% (specify)
	More than 30% (specify)
	Unclear/not stated (specify)
Section F: Outcome measurement	
Intermediate outcomes	Employability (specify)
	Access to employment opportunities (specify)
	Labour market participation enhancing behaviour(specify)
	Social capital (specify)

Questions	Answers
	Policy change (specify)
	Other (specify)
	Unclear (specify)
Final outcomes (labour market participation)	None
	Change in employment status (from unemployed to employed in high growth/male-dominated sector), measured in terms of employment probability or participation rate
	Change in employment sector (from traditional sector for women's employment to high growth/male-dominated sector)
	Change in employment status (from underemployment to full employment in high growth/male-dominated sector)
	Nature of employment
	Progression/career prospects
	Firm size
	Other (specify)
	Unclear (specify)
Type of measure (i.e. data source) *Where relevant, state whether data pre-existing or not (e.g. the investigators may use pre-existing survey data, or collect original data).*	In-depth interviews
	Survey, questionnaire
	Administrative data
	Other (specify)
	Unclear/not stated (specify)
Number of post-test measurements	One
	Two
	Three
	Other (specify)
	Unclear (specify)
Timing of outcome measurements *Select all that apply.* *Indicate if this is the time that has lapsed since baseline measurement, start of intervention or end of the intervention.*	Not stated
	0-6 months (specify)
	7-12 months (specify)
	13-18 months (specify)
	Over 18 months (specify)

Questions	Answers
If different outcomes are measured at different time points, provide details.	Other (specify)
	Unclear (specify)
Evidence of unintended effects	Doing harm (specify)
	Effect capture (specify)
	Other (specify)
	Unclear (specify)
Evidence of cost data being collected	Cost data reported
	Cost-effectiveness calculations being reported
Section H: Effect size calculation and data entry	
Section G: Narrative findings	
Authors report interventions as:	Effective (specify)
	Ineffective (specify)
	Other (specify)
	Unclear (specify)
Copy & paste authors conclusions	

APPENDIX 6: CRITICAL APPRAISAL TOOL

Methodological appraisal criteria	Response		
	Yes	No	Comment
IF RANDOMISED CONTROL TRIAL, START AFTER CONFOUNDING BIAS. FOR ALL OTHER STUDY DESIGNS, START HERE. **I. <u>Bias in selection of participants into the study</u>** *Are participants selected in a way that minimizes selection bias?*[17] <u>Appraisal indicators</u> Consider whether:			
i. there is an adequate description of how and why sample was chosen (i.e. identified/selected/recruited)			
ii. there is adequate sample size to allow for representative and/or statistically significant conclusions			
iii. participants in the control[18] *group were sampled from the same population as that of the treatment*			
iv. group allocation process minimised potential risk of bias (e.g. using computer algorithms)			
v. the selection of participants into the study (or into the analysis) is based on participant characteristics observed after the start of the intervention			
Low risk of bias / **Risk of bias** / **High risk of bias** / **Critical risk of bias**	*Worth to continue: Y/N?*		
II. <u>Bias due to confounding</u> *Is confounding potentially controllable in the context of this study?* <u>Appraisal indicators:</u> Consider whether:			

[17] *Selection bias can occur both in the way that individuals are **accepted** for participation in a study, and in the way that 'treatment' is **assigned** to individuals once they have been accepted into a study. This section deals with both these understandings of selection bias.*

[18] *The terms 'control' and 'comparison' group refer to any group with which the treatment of interest is compared that is presumed to represent conductions in the absence of that treatment, whether a true random control or not.*

Methodological appraisal criteria	Response		
	Yes	No	Comment
i. there is potential for confounding of the effect of the intervention in this study. If yes, provide example of confounding domain in comment box[19].			
ii. where matching was applied, it featured sufficient criteria[20]			
iii. where relevant, the authors conducted an appropriate analysis that controlled for all potential/remaining critical confounding domains after matching had been applied			
iv. *the authors avoided to adjust for variables identified after the intervention has been administered*			
v. the treatment and control group are comparable after matching/controls have been done. Select one of the following: ☐ No statistically significant differences ☐ Statistically significance differences ☐ Negligible descriptive differences ☐ Significant descriptive differences			
Low risk of bias | **Risk of bias** | **High risk of bias** | **Critical risk of bias**	*Worth to continue: Y/N?*		
IF RANDOMISED CONTROL TRIAL, SKIP I + II (ABOVE) AND START HERE! **Bias due to confounding (as a result of ineffective randomisation)** *Is allocation of treatment status truly random?* Appraisal indicators Consider whether:			
i. eligibility criteria for study entry are specified			
ii. there is a clear description of the randomisation process and methods are robust			
iii. the unit of randomisation and number of participants is clearly stated (pay special attention to treatment and control locations/ balance)			

[19] *Confounding domains are those for which, in the context of this study, adjustment is expected to lead to an important change in the estimated effect of the intervention*

[20] *Matching can be done on the calculated propensity score or covariates. If the latter, it should ideally be done on pre-test measures and other characteristics, such as demographic. Answer 'no' if the study only matched on pre-test measures of some or all variables used later as outcome measures OR matched only on endline characteristics.*

Methodological appraisal criteria				Response		
				Yes	No	Comment
iv. characteristics of both baseline and endline sample are provided[21] and at endline the treatment and control group are comparable. Select one of the following: ☐ No statistically significant differences ☐ Statistically significance differences ☐ Negligible descriptive differences *v.* ☐ Significant descriptive differences						
Low risk of bias	**Risk of bias**	**High risk of bias**	**Critical risk of bias**	*If critical risk of bias, treat as non-random study*		
III. <u>Bias due to departures from intended interventions</u> *Was the intervention implemented as laid out in the study protocol?* <u>Appraisal indicators</u> Consider whether:						
i. the critical co-interventions were balanced across intervention and control groups						
ii. treatment switches were low enough to not threaten the validity of the estimated effect of the intervention						
iii. implementation failure was minor and unlikely to threaten the validity of the estimated effect of the intervention						
iv. it is possible that the intervention was taken by the controls (contamination and possible crossing-over)[22]						
v. it is possible that knowledge of group allocation affects how the two study groups are treated during delivery and evaluation of the intervention[23]						
Low risk of bias	**Risk of bias**	**High risk of bias**	**Critical risk of bias**	*Worth to continue: Y/N?*		
IV. <u>Bias due to missing/incomplete data (attrition)</u> *Are the intervention and control groups free of critical differences in participants with missing/incomplete data?* <u>Appraisal indicators</u> Consider whether:						

[21] *Preferable condition: A RCT with appropriate randomization procedure can be included without showing baseline data, as both experimental groups can be assumed to be equal at baseline by design.*

[22] *Whilst challenging in terms of estimating impact, spill-overs might be an important finding in itself.*
[23] *Consider only in extreme cases in which preferential treatment is clearly evident; blinding is generally not expected in social interventions.*

Methodological appraisal criteria	Response		
	Yes	No	Comment
i. outcome data are reasonably complete (80% or above)[24]			
ii. If level of attrition (or other forms of missing/incomplete data) is more than 20%, are reasons for the missing data reported?			
iii. If level of attrition (or other forms of missing/incomplete data) is more than 20%, do the authors demonstrate similarly between remaining participants and those lost to attrition and are the proportion of participants with missing/incomplete data and reasons for missing/incomplete data similar across groups?			
iv. If level of attrition (or other forms of missing/incomplete data) is more than 20%, were appropriate statistical methods used to account for missing data? (e.g. sensitivity analysis)[25]			
v. If not possible to control for missing/incomplete data, are outcomes with missing/incomplete data excluded from analysis?			
Low risk of bias / **Risk of bias** / **High risk of bias** / **Critical risk of bias**	*Worth to continue: Y/N?*		
V. <u>Bias in measurement of outcomes</u> *Are measurements appropriate, e.g. clear origin, or validity known?* <u>Appraisal indicators</u> Consider whether:			
i. there was an adequate period for follow up[26]			
ii. the outcome measure (e.g. employment status, income) was clearly defined and objective[27]			
iii. outcomes were assessed using standardised instruments and indicators			
iv. outcome measurements reflect what the experiment set out to measure			
v. the methods of outcome assessment were comparable across groups			

[24] *The assumption here that the level of attrition (or other forms of missing/incomplete data) is sufficiently low to not require adjustment.*
[25] *Select 'no' if the study addresses missing/incomplete data through simple estimates of missing data and observations.*
[26] *In many social science interventions, follow-up is not required to coincide with the start of the treatment; further, longer period of follow up are often required to measure changes.*
[27] *Subjective measures (e.g. those based on self-report) are likely to have lower reliability and validity than objective measures*

Methodological appraisal criteria				Response		
				Yes	No	Comment
vi. were outcome assessors aware of the intervention received by study participants?[28]						
Low risk of bias	**Risk of bias**	**High risk of bias**	**Critical risk of bias**	*Worth to continue: Y/N?*		
VI. <u>Bias in selection of results reported</u> *Are the reported outcomes consistent with the proposed outcomes at the protocol stage?* <u>Appraisal indicators</u> Consider whether:						
i. it is unlikely that the reported effect estimate has been selected for publication due to it being a particularly notable finding among numerous exploratory analyses						
ii. it is unlikely that the reported effect estimate is prone to selective reporting from among multiple outcome measurements within the outcome domain						
iii. it is unlikely that the reported effect estimate is prone to selective reporting from among multiple analyses of the outcome measurements, including sub-groups analysis						
iv. if sub-group/ancillary/adjusted analyses are presented, are these pre-specified or exploratory?						
v. the analysis includes an intention to treat analysis? (If so, was this appropriate and were appropriate methods used to account for missing data?)[29]						
vi. do the authors report on all variables they aimed to study (as specified in their protocol or study aims/research questions)?						
Low risk of bias	**Risk of bias**	**High risk of bias**	**Critical risk of bias**			
OVERALL RISK OF BIAS:						

[28] *Consider only in extreme cases in which preferential treatment is clearly evident; blinding is generally not expected in social interventions.*

[29] *Usually in clinical RCTs, rare in social science: only rate if conducted.*

APPENDIX 7: EFFECT SIZE CALCULATIONS

Corrected SMD and corrected Standard Errors (SE) will be estimated as follows:

$$SMD_{corrected} = SMD_{uncorrected} * \left[1 - \frac{3}{4 * (n_t + n_c - 2) - 1}\right]$$

$$SE(SMD)_{corrected} = SE(SMD)_{uncorrected} * \left[1 - \frac{3}{4 * (n_t + n_c - 2) - 1}\right]$$

For regression-based estimates, we will follow Keef and Roberts (2004: p100-101; p129, equation A9) to correct for potential sample bias in the effect sizes.

Many of the impact evaluation designs that we expect to see in this review are likely to use complex statistical analyses, and there is a lack of standard methods for computing effect sizes from these designs. In most cases we expect our approach for computing effect sizes to be as follows.

CALCULATING STANDARDISED MEAN DIFFERENCES

For studies reporting matching-based estimates, the following formulae to compute *g* and its standard error will be used, where Y_t and Y_c are the post-intervention mean outcome in the treatment group and control group respectively.

$$SMD = \frac{Y_t - Y_c}{S_p}$$

To calculate S_p, the pooled standard deviation (the standard deviation of the outcome variable for both treated (S_t) and control (S_c) individuals), we will use the Hedges' approach described in Lipsey and Wilson (2001), where n_t and n_c are the sample size of the treatment and the control group.

$$S_p = \sqrt{\frac{(n_t - 1) * S_t^2 + (n_c - 1) * S_c^2}{n_t + n_c - 2}}$$

The standard error of *g* will be computed using the following formula.

$$SE\ (SMD) = \sqrt{\frac{n_t + n_c}{n_t * n_c} + \frac{SMD^2}{2 * (n_c + n_t)}}$$

Alternatively, in the event that the necessary information for calculating SE is not available, we will approximate it as follows, where *t* is the *t* statistic of the treatment effect.

$$SE\ (SMD) = \frac{SMD}{t}$$

For regression-based studies, we intend to use the following formula described in Keef and Roberts (2004), where β is the coefficient of interest (i.e. yielding the impact of the intervention) and σ is the standard deviation of the error term in a regression.

$$SMD = \frac{\beta}{\sigma}$$

Where σ is not reported by the study authors (highly likely), we will use the following formula to compute an equivalent using the sample standard deviation of the dependent variable and the sample size for both treatment and control groups.

$$S_p = \sqrt{\frac{(SD_y^2 * (n_t + n_c - 1) - \frac{(\beta^2 * (n_t * n_c))}{n_t + n_c}}{n_t + n_c}}$$

Standard errors will be approximated using the following formula where *t* is the *t* statistic for the regression coefficient.

$$SE\ (SMD) = \frac{SMD}{t}$$

CALCULATING RESPONSE RATIOS

For studies using matching, the following formulae to compute *g* and its standard error will be used.

$$RR = \frac{Y_t}{Y_c}$$

The standard error of *g* will be computed using the following formula, where *t* refers to either the *t* statistics/p-value of the regression coefficient or to the results of the *t* test for equality of means between groups after matching.

$$SE(RR) = \exp\frac{Ln(RR)}{t}$$

Alternatively, the standard error of *g* can also be computed as:

$$SE(RR) = S_p^2 * \left(\frac{1}{n_t * Y_t^2} + \frac{1}{n_c * Y_c^2}\right)$$

For regression-based studies, the following formula to compute *g* and its standard error will be used.

$$RR = \frac{Ys + \beta}{Yc}$$

The standard error of *g* will be computed using the same formula as presented for matched-based studies.