María Daniela Anda León
Carolyn Huang
Miriam Berretta
Promise Nduku
Andile Madonsela
Shannon Shisler

Systematic Review on the Effect of Skills-Based Active Labor Market Interventions on Migration Outcomes

October 2023

Systematic Review 48

Social protection

3ie International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) develops evidence on how to effectively transform the lives of the poor in low- and middle-income countries. Established in 2008, we offer comprehensive support and a diversity of approaches to achieve development goals by producing, synthesizing and promoting the uptake of impact evaluation evidence. We work closely with governments, foundations, NGOs, development institutions and research organizations to address their decision-making needs. With offices in Washington DC, New Delhi and London and a global network of leading researchers, we offer deep expertise across our extensive menu of evaluation services.

3ie systematic reviews

3ie systematic reviews appraise and synthesize the available evidence on the effectiveness of social and economic development interventions in low- and middle-income countries. These reviews follow scientifically recognized review methods and are peer-reviewed and quality assured according to internationally accepted standards. 3ie is providing leadership in demonstrating rigorous and innovative review methodologies, such as using theory-based approaches suited to inform policy and programming in the dynamic contexts and challenges of low- and middle-income countries.

About this review

The *Systematic Review on the Effect of Skills-Based Active Labor Market Interventions on Migration Outcomes* was developed by 3ie with funding from the International Organization for Migration (IOM) - Guatemala through the United States Agency for International Development (USAID) funded project, Addressing the Root Causes of Irregular Migration in Guatemala. The content of this report is the sole responsibility of the authors and does not represent the opinions of IOM, USAID, the U.S. government, 3ie, its donors or its Board of Commissioners. Any errors or omissions are also the sole responsibility of the authors. Please direct any comments or queries to the corresponding author, María Daniela Anda León, at danda@3ieimpact.org.

Suggested citation: Anda-León, María Daniela, Carolyn Huang, Miriam Berretta, Promise Nduku, Andile Madonsela, and Shannon Shisler. 2023. *Systematic Review on the Effect of Skills-Based Active Labor Market Interventions on Migration Outcomes*. 3ie Systematic Review 48. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: https://doi.org/10.23846/SR00048

3ie systematic review executive editor: Birte Snilstveit
Managing editor: Tanvi Lal
Publications and web design: Akarsh Gupta
Production manager: Mallika Rao

© International Initiative for Impact Evaluation (3ie), 2023

Systematic Review on the Effect of Skills-Based Active Labor Market Interventions on Migration Outcomes

María Daniela Anda León
International Initiative for Impact Evaluation (3ie)

Carolyn Huang
3ie

Miriam Berretta
3ie

Promise Nduku
South Africa Centre for Evidence (SACE)

Andile Madonsela
SACE

Shannon Shisler
3ie

3ie Systematic Review 48

December 2023

Acknowledgements

We are grateful to 3ie and SACE colleagues who provided input at various stages of the review and supported on the administrative, communication and finance management. We are particularly grateful to Constanza Gonzalez Parrao, Maren Duvendack and Anil Thota for their thorough review of the protocol and final report. We are grateful to Juan Pascual who contributed to the extraction and assessment processes, and to Edward Kayongo who contributed to the screening and retrieval processes. We would also like to thank the members of our advisory group: David McKenzie, Alan Hirsch, and Domenico Tabasso for their thoughtful suggestions. The authors of this systematic review would like to thank colleagues from IOM Guatemala, for their time, flexibility and input. Finally, thanks to all the team that worked in the execution of the search and screening protocols for the Evidence Gap Map Addressing Root Causes and Drivers of Irregular Migration.

Executive summary

Background

One of the root causes of irregular migration is economic insecurity, which has been addressed by governments, international aid, and development institutions through human capital development interventions such as skills-based active labor market policies (ALMPs). The theory behind this strategy assumes that providing new skills to labor market participants and entrants can help them find employment opportunities, which in turn has a dichotomous impact on migration. On one hand, new and more available opportunities in the country of origin increase the opportunity cost of migrating; on the other, beneficiaries improved economic security gives them access to safe ways through which to migrate to destinations where their acquired skills can also increase their employability. We synthesize the evidence on the effectiveness of this type of intervention on migration-related outcomes.

Research objectives

Our primary objective for this review was to identify, assess, and synthesize evidence on the effect of skills-based ALMPs on migration outcomes. We aim to make the evidence in the field more accessible and facilitate its use in informing policy and practice decisions regarding irregular migration programming. To achieve this objective, we addressed the following questions:

1. How effective are skills-based active labor market interventions in improving intermediate outcomes that influence eventual migration (intention to migrate, and knowledge, perceptions, attitudes, and expectations) or migration behavior outcomes (attempted migration, any migration, international migration, migration flow, migration stock) in low and middle-income countries (L&MICs)?
2. Are there any unintended consequences of such interventions?
3. Do effects vary by context, intervention type, design or population characteristics (e.g., sex)?
4. What are contextual barriers to and facilitators of intervention effectiveness?
5. What is the cost-effectiveness of these interventions?
6. How can future research enrich this evidence base?

Methods

We conducted a theory-based mixed-methods systematic review using studies identified through the systematic literature search of key academic databases and grey literature sources conducted for the Evidence Gap Map (EGM), Addressing Root Causes and Drivers of Irregular Migration by Berretta and colleagues (2023). The EGM included published and unpublished literature since 1990, and both experimental and quasi-experimental study designs were eligible. Full text screening was completed independently in duplicate. The EGM team conducted backward and forward citation tracking, and the authors published a blog with an open call for relevant studies. Studies from the EGM meeting the following criteria were included in this synthesis: impact evaluations of skills-based ALMPs conducted in L&MICs of origin, exploring the effects on migration, attempting migration and/or intermediate outcomes that influence eventual migration such as intention to migrate and knowledge, perceptions, attitudes, and expectations about migration.

Two coders assessed the risk of bias of included studies and extracted data independently in duplicate. We computed Standard Mean Differences (SMD) of impact estimates and pooled studies together using random effects meta-analysis for comparable outcomes. When feasible, we explored potential sources of heterogeneity in the outcomes.

We conducted a targeted search for qualitative studies, descriptive quantitative studies, project documents, and process evaluations of the programs evaluated in included impact evaluations. This programmatic evidence base was synthesized using a thematic approach on factors related to context, intervention design and implementation, and population characteristics.

Included Studies

We included ten impact evaluations of skills-based ALMPs in L&MICs of origin, all of which used random allocation of treatment to measure program impacts, and six of which used a mixed methods approach or provided enough details on the context, design and implementation of the program to be included in both quantitative and qualitative syntheses. We found seven additional records that met the inclusion criteria for qualitative studies. The included quantitative evaluations covered programs from Sub-Saharan Africa (n = 6), Latin America and the Caribbean (n = 2), and South Asia and East Asia (n = 2). Only two of the programs being evaluated in this evidence base were specifically targeting potential migrants. These two were also the only studies reporting on international migration specifically. One in particular used measures of irregular migration. The rest of included studies reported on outcomes such as migration unspecified as to whether it is regular or irregular.

Results

We found no evidence of an effect of skills-based ALMPs on migration outcomes, including intention to migrate, attempted migration, and actual migration behavior. All the estimates were based on relatively small numbers of studies, limiting our power to detect effects, and nine of the ten impact evaluations were assessed as having a high risk of bias. Given these limitations, the results should be interpreted with caution and taken as a first attempt to synthesize the evidence of an emerging field for rigorous impact evaluations. Upcoming programing and research can take these findings and implications as a starting point, in particular reflecting on the reasons that we provide as possible explanations for the lack of findings.

Results from the qualitative synthesis were also based on a small number of studies, limiting our ability to draw strong conclusions. The findings suggest that uptake of skills-based ALMPs may increase through a sufficient range of awareness-raising campaigns and support to understand the objectives of the program. It is helpful when training activities are tailored to the local context and consider participant preferences (e.g., in relation to program length). Mechanisms to ensure that collaborators clearly understand their authority and responsibilities are key to the successful completion of skills-based active labor market programs. The timely and predictable delivery of certificates to beneficiaries upon completion of training can enhance positive program impacts, while obstacles and delays to accessing tools or fees are potential bottlenecks. Participation may be enhanced by providing safe and accessible training locations and by covering transportation and subsistence costs. Support

from family and local authorities may be important in ensuring participation whilst networks may become key for participants to secure employment after the completion of training. Finally, when designing interventions, implementers may need to consider how they will include participants from different socioeconomic statuses and account for external constraints, such as health status or household and childcare duties.

We found evidence of cost-effectiveness reported in one of the ten impact evaluations, but the analysis was not related to migration outcomes. Rather, it was looking at the cost-effectiveness for earning outcomes. Two studies also reported data related to program costs, but the rest of the studies did not include any information on cost.

Implications for research

Conducting rigorous research on migration programming can be challenging for a variety of reasons. While some evidence exists on the effects of skills-based ALMPs on migration-related outcomes, the current evidence base is insufficient to understand whether, how, or why these programs are working, particularly because most of the programs in the evidence base for this review were not explicitly targeting potential migrants. Commissioning more research across a range of geographies that is specific to migration programing will help fill gaps in the current evidence base.

Future studies using mixed-methods approaches can help comprehensively address questions about how and why interventions are successful (or not). The addition of cost analyses can provide further evidence about the relative cost-effectiveness of programs so that the most can be made of finite resources. Programs might also consider measuring longer-term program impacts and utilizing technology (e.g., mobile apps) to assist with data collection and mitigate attrition and non-response issues.

Finally, many studies do not comprehensively report on their methodologies and/or results. Impact evaluations can make more useful contributions to our knowledge base if they register and report a pre-analysis plan, establish an appropriate counterfactual, clearly describe all constructs and methods in detail, report all statistical information needed to calculate effect sizes, and avoid *p*-hacking by reporting on all outcomes, regardless of statistical significance.

Implications for policy and practice

Policymakers and practitioners have an important role to play in addressing these evidence gaps by identifying opportunities for evaluation in current and future investments. We suggest that when designing or commissioning skills-based ALMP programs, decision-makers should consider collecting primary data to establish the acceptability and appropriateness of interventions and confirm that assumptions in any program theories of change reflect the concerns and needs of program participants and are contextually appropriate. Migration decision-making is often driven by a combination of factors in addition to economic insecurity. A program that fails to account for all primary drivers may lead to low uptake and diminish the rigor of impact evaluations. In addition, successful programs may not change an individual's desire to migrate but may actually increase their capacity to migrate. If there are few legal channels to do so, individuals may choose irregular pathways. Program implementers should anticipate and plan to ensure that programs do not unintentionally exacerbate these dynamics. It will also be useful to identify opportunities to include impact evaluations in future programming.

There is a dearth of impact evaluations on programs that explicitly attempt to address migration and target populations that consider irregular migration as a means for a better life. However, this requires intentionality and commitment – for example, by identifying activities or phases of a program that could facilitate random assignment and testing of specific program components for impact, or using eligibility criteria for quasi-experimental approaches. For example, the pilot or initial phases of rolling out a large program may present an opportunity to use naturally occurring implementation-scaling delays for running an impact evaluation.

When appropriate and where evidence generation is a priority, policymakers should also ensure that evaluated programs are of adequate scale so that impact evaluations are sufficiently powered for an evaluation. Relatedly, ensure there is adequate funding for impact evaluations within the larger investments, and consider how potential investment impacts can be sustained after the life of the program. Successfully building skills and strengthening local workforce capability may address “demand-side” issues in the short term, but will fail to fully address economic root causes if there are no long-term solutions for “supply-side” issues like lack of local labor market and economic opportunities.

Contents

Acknowledgements ..... **i**
Executive summary ..... **ii**
List of figures and tables ..... **vii**
1. Background **1**
1.1 The problem, condition, or issue 1
1.2 The intervention 2
1.3 Theory of Change 2
1.4 Rationale for the review 5
2. Research questions **5**
3. Methods **5**
3.1 Criteria for including and excluding studies in the review (PICOS) 6
3.2 Search strategy 8
3.3 Selection and coding of studies 8
3.4 Analytical approach for quantitative and qualitative data 9
4. Results **11**
4.1 Characteristics of the evidence base 11
4.2 Quantitative impacts 25
4.3 Qualitative findings (context, barriers, and enablers) 31
5. Discussion **40**
5.1 Summary and application of findings 40
5.2 Implications for future research, policy, and practice 46
5.3 Limitations of this review 49
6. Acknowledgements **49**
6.1 Contributions of authors 49
6.2 Declarations of interest 49
6.3 Plans for updating this review 49
Appendix A: Search strategy **50**
Appendix B: Data extraction tools **52**
Appendix C: Critical appraisal tool **56**
Appendix D: Risk of bias criteria definitions **67**
References **69**

List of figures and tables

Figure 1: Theory of Change 4
Figure 2: Composition of included studies by study design 12
Figure 3: Risk of bias assessment of included impact evaluations 24
Figure 4: Risk of bias assessment by criterion assessed 25
Figure 5: Forest plot showing the observed outcomes and the estimate of the random-effects model of skills-based ALMPs on intention to migrate 26
Figure 6: Forest plot showing the observed outcomes and the estimate of the random-effects model of skills based ALMPs on attempted migration 27
Figure 7: Forest plot showing the observed outcomes and the estimate of the random-effects model of skills-based ALMPs on migration 29
Figure 8: Forest plots showing the observed outcomes and the estimates of the random-effects models of skills-based ALMPs on migration by gender 31

Table 1: Select current and salient policy responses to root causes of irregular migration 1
Table 2: PICOS inclusion criteria 6
Table 3: Characteristics of programs in included impact evaluation 15
Table 4: Characteristics of included impact evaluations 19
Table 5: Characteristics of included programmatic studies linked to impact evaluations 23
Table 6: Overview table of descriptive and analytical themes 32
Table 7: Synthesis of the quantitative and qualitative analysis answering the research questions of this review 45

1. Background

1.1 The problem, condition, or issue

In a globalized world, international migration serves important development purposes. Besides being an internationally recognized human right as the "natural expression of people's desire to choose how and where to lead their lives, which is a fundamental component of human development" (UNSD 2022, 12), some evidence shows that it may also improve development in countries of origin through remittances (Ghosh 2006; Faist 2008; Hossain 2022). However, when individuals are forced to migrate out of necessity or survival and there are limited means outside of formal channels, international migration can increase the vulnerability of already-disadvantaged populations.

Irregular migration affects millions of people around the world (Yayboke and Gallego 2019), putting them at greater risk of financial and/or labor exploitation, physical harm, violence, or death (Vutha, Pide, and Dalis 2011; Yayboke and Gallego 2019; United Nations Office on Drugs and Crime 2021; ILO 2022). This has induced governments and international organizations to invest significant resources in addressing the "root causes" of irregular migration that create unfavorable conditions in countries of origin (e.g., economic disparity, conditions exacerbated by climate change, political instability, insecurity and transnational crime) and humanitarian crises such as conflicts, wars, or persecution (Loschmann, Kuschminder, and Siegel 2014; National Security Council 2021; Rose et al. 2021; Vutha, Pide, and Dalis 2011; Yayboke and Gallego 2019; UNHCR 2022a).

There are often multiple factors behind an individual's decision to migrate so that migration behavior cannot be attributed to a single or primary reason. These factors are jointly considered and may include broader drivers (Gent 2002). Several large-scale policies have been designed to address what have been identified as "root causes" (Table 1).

Table 1: Select current and salient policy responses to root causes of irregular migration

Program	Resources invested	Root causes addressed	Beneficiaries
The Netherlands 2016-2021 Addressing Root Causes of Conflict, Instability and Irregular Migration (ARC) program (ECORYS 2020)	EUR €90 million	Security, rule of law, peace processes, political governance, and socioeconomic reconstruction	Burundi, Democratic Republic of Congo, Ethiopia, Jordan, Lebanon, Mali, Somalia, South Sudan, Sudan, and Syria
	EUR €37 million	Governance, rule of law, access to markets and employment, peace, and security	Pakistan and Afghanistan
The EU Emergency Trust Fund for Africa (EUTF for Africa) (Knoll and Sheriff 2017; European Commission. Directorate General for International Partnerships. 2022)	EUR €4.2 billion	Diverse causes of instability, irregular migration and forced displacement to support all aspects of stability, security and resilience.	Sub-Saharan Africa
The US Root Causes Strategy (Office of Management and Budget 2022)	USD $987 million	Economic insecurity, inequality, governance, human rights and free press, and gender-based violence and trafficking	Central America

There is insufficient empirical research examining whether such “root cause” interventions effectively decrease irregular migration, despite the large programs that adopt these approaches (Berretta, Anda Leon, et al. 2023). Rather, the existing evidence base is primarily descriptive, reporting on the reasons individuals choose to migrate, characteristics of who decides to migrate, and the broader development impacts of migration (Goldin et al. 2018; IMF 2015; Obokata, Veronis, and McLeman 2014; Pitoski, Lampoltshammer, and Parycek 2021). Systematic evidence on the effectiveness of programs addressing root causes of irregular migration is, therefore, still scant. The objective of this systematic review is to synthesize the evidence on the effectiveness of one such type of intervention addressing the root causes of irregular migration, by using the quantitative impact evaluations identified in the evidence gap map *Addressing root causes and drives of irregular migration* by Berretta and colleagues (2023), and identifying related qualitative evidence to further explore elements of the context, program design and implementation that can shed light on why, how, when and for whom this type of intervention works (or does not).

The line of root-cause programing that we focus on is active labor market policies (ALMP), defined as demand-side interventions in the countries of origin that aim to create and improve employment opportunities for potential migrants. These interventions include skills-based training or apprenticeships programs, job search assistance programs, employment pipelines/pathways, public works schemes and self-employment promotion efforts. In this review we focus on skills-based training or apprenticeships programs as described in the following section in more detail. Extant literature on ALMPs has focused on the effects on earnings and employment in local or national labor markets (Card, Kluve, and Weber 2015; McKenzie 2017), and is mostly concentrated in high-income countries (Dar and Tzannatos 1999; Betcherman, Olivas, and Dar 2004). Indeed, the evidence gap map by Berretta and colleagues (2023) found no systematic effectiveness reviews looking at the effects of ALMPs on migration outcomes. Thus, to our knowledge, this is the first systematic review of the literature on the effectiveness of such policies on migration outcomes for low- and middle-income countries (L&MICs).

1.2 The intervention

We included interventions that address the root cause of migration related to economic insecurity through active labor market policies (ALMPs) that aim to provide labor market entrants and participants with a set of new or improved skills for their job, referred to skills-based ALMPs. However, these interventions may or may not have been explicitly designed to affect migration behavior specifically. For analytical purposes, we focused on a specific type of ALMP: skills-based interventions which seek to provide training on technical, professional, financial, or business skills. These may include classroom-delivered (vocational) training or experiential, on-the-job training (apprenticeships).

1.3 Theory of Change

The theory of change linking skills-based ALMPs and irregular migration assumes that the net benefits of migrating through irregular channels are reduced when unfavorable systemic conditions at home, such as economic insecurity or lack of employment opportunities, are addressed. Such improvement in economic security increases the prospects of staying in the origin country, which can be considered the opportunity cost of migrating.

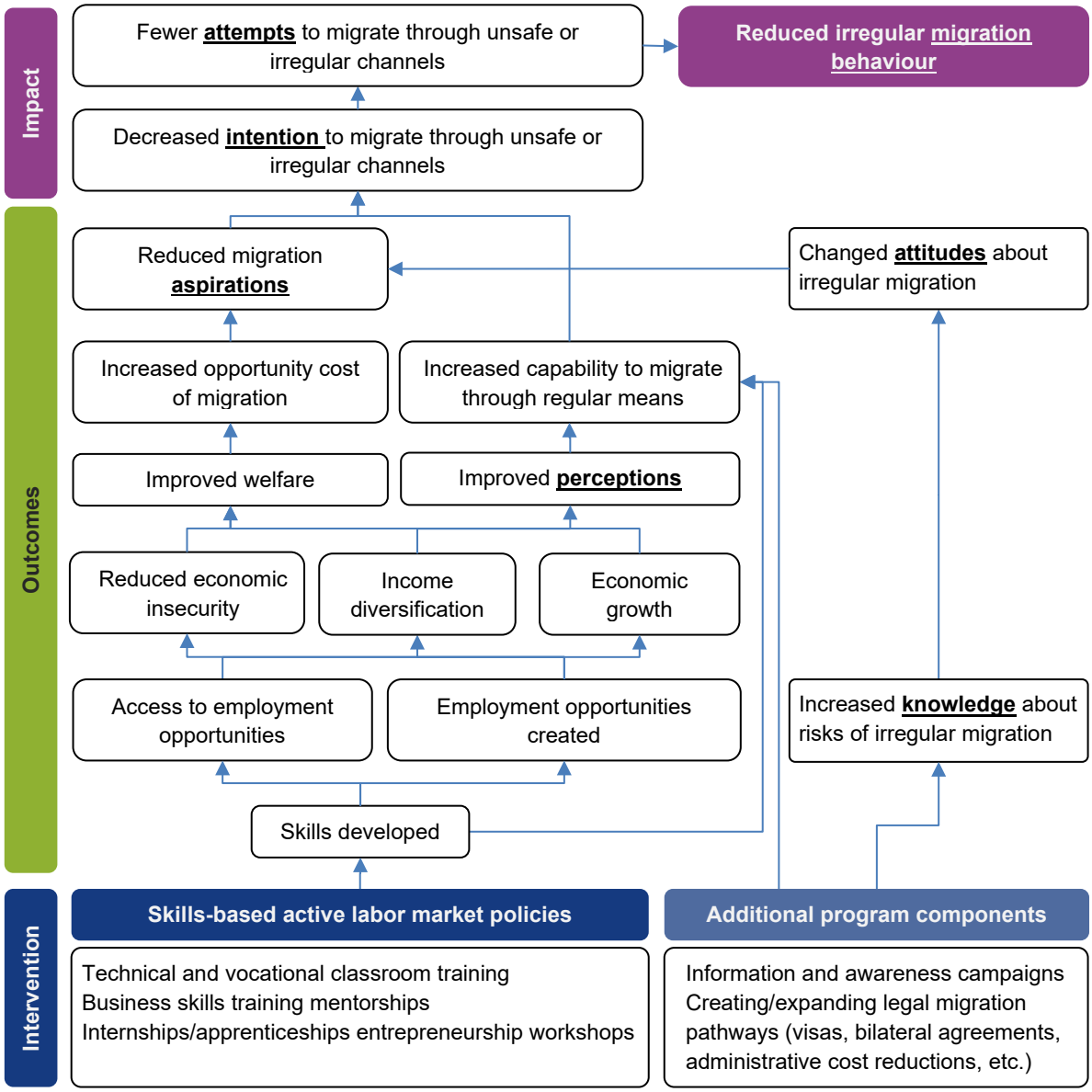
Our theory of change is adapted from Carling's (2002) theory of migration aspirations and abilities – further adapted by Carling and Talleraas (2016) and Carling and Schewel (2018) – who present a framework of an individual's decision-making process. According to this conceptual model, migration decisions made on a "voluntary" basis are driven by poor conditions, limited prospects, and/or perceptions in origin countries of stagnation or hopelessness and leading to a desire for change, which may affect migration aspirations.

We hypothesize that skills-based ALMPs provide new or improved skills to labor market entrants and participants that create employment opportunities, either through facilitating access to formal labor markets or self-made, entrepreneurship opportunities. This leads to improvements in local conditions and prospects such as greater economic security and income diversification. These factors may increase the opportunity cost of migrating, resulting in decreased migration aspirations and ultimately a reduction in migration behavior. Irregular migration may also decrease as a result of an increased capability to migrate through regular means (Massey et al. 1993; Kleemans 2015). Such changes may reduce decisions and intentions to travel through unsafe or irregular channels that were based on a lack of better opportunities (Figure 1).

We present several assumptions in our theory of change, acknowledging this list is unlikely to be exhaustive. Assumptions may pertain to the targeted population, context, program design, and implementation and they may affect one or more causal linkages in the impact pathway. If any assumptions are unmet, we would expect to observe no detectable differences in migration outcomes. For example, if the skills-based program is insufficiently designed from the onset, we would not expect assumptions around skills developed or access to employment opportunities to be met, or changes to occur further down the causal chain.

Our main assumption is that migration aspirations are mainly driven by economic factors (Soto et al. 2021).If this assumption is unmet, improvements in welfare could lead to unintended increases in irregular migration. For example, if migration aspirations are not primarily driven by economic security, but rather by a strong desire for change as a result of other factors such as climate change or violence. individuals will possess improved means to achieve migration aspirations through skills development and improved welfare or financial ability to afford migration costs.

Figure 1: Theory of Change



Assumptions

Design

- Effective targeting of potential migrants
- Sensitization and advertisement campaigns are effective
- The curriculum includes skills demanded by labor markets (internally and internationally)
- The length of the intervention is appropriate to provide the set of skills and in line with participants' preferences

Population

- Migration aspirations are mostly driven by economic considerations (among all other drivers)
- Participants do not have access to other sources of skills-development services
- Participants will prefer regular means over irregular migration when accessible
- Participants will take new and recently accessible employment opportunities

Implementati

- Positive program enrolment and uptake
- Compliance with the intervention protocols
- Training location is accessible, safe, and convenient.
- The timing of the training does not interfere with other activities of participants' livelihoods

Context

- Regular means to migrate are available
- Acquired skills are valuable in national and international labor markets
- There are no barriers to entrepreneurship aside from skills development

1.4 Rationale for the review

This systematic review is expected to inform decision-making regarding skills-based ALMPs designed to address the lack of economic security as a root cause of irregular migration. Key decision-makers in the field have indicated interest in this area and can utilize the results of this review to inform interventions creating economic opportunities and developing skills in the workforce that aim to improve migration outcomes.

2. Research questions

Our primary objective for this review was to identify, assess and synthesize evidence on the effect of skills-based ALMPs on migration outcomes. We aim to facilitate the use of evidence in informing policy and practice decisions within the field of irregular migration programming. To achieve this objective, we answered the following questions:

1. How effective are skills-based active labor market interventions in improving **intermediate outcomes that influence eventual migration** (intention to migrate, and knowledge, perceptions, attitudes and expectations) in low and middle-income countries (L&MICs)?
2. How effective are skills-based active labor market interventions in improving **migration behavior** (attempted migration, any migration, international migration, migration flow, migration stock) in L&MICs?
3. Are there any unintended consequences of such interventions?
4. Do effects vary by context, intervention type, design or population characteristics (e.g., age, sex, socio-economic status, etc.)?
5. What are contextual barriers to and facilitators of intervention effectiveness?
6. What is the cost-effectiveness of these interventions?
7. How can future research enrich the evidence on the effects of active labor market interventions designed to improve migration in L&MICs?

3. Methods

To respond to these research questions, we conducted a theory-based mixed-methods systematic review using best practices outlined by Snilstveit (2012) and by Cochrane and the Campbell Collaboration (Shemilt et al. 2013; Kugley et al. 2017; J. P. T. Higgins et al. 2019).

The quantitative evidence included in this review was based on the systematic literature search of key academic databases and grey literature sources conducted for the Evidence Gap Map (EGM), 'Addressing Root Causes and Drivers of Irregular Migration,' see Berretta and colleagues (2023) for search details. We assessed the quality of included studies which evaluated the effects of skills-based ALMPs and summarized visually and in a narrative format its results related to migration outcomes. Whenever the number of studies and levels of heterogeneity in intervention, outcomes and context suggested that it was reasonable to pool effect sizes together, we also performed a meta-analysis to estimate an average effect size.

We implemented an additional search for linked publications of the programs evaluated in included studies. This was done to inform a qualitative synthesis of the evidence and address research questions related to unintended consequences, the intervention context and barriers and facilitators of change (research questions 3 through 5). More details about the methods used in this systematic review can be found in the review protocol (Anda Leon et al. 2023).

While the evidence base for this review was scarce, our precisely defined intervention scope made it possible to pool effect sizes together and estimate an average treatment effect using statistical methods. We further complemented the quantitative analysis on the effectiveness of ALMPs with qualitative evidence identified to address research questions on contextual considerations and design features. However, there are caveats of working with small sample sizes. We highlight them in the results and discussion sections.

3.1 Criteria for including and excluding studies in the review (PICOS)

We included studies from academic and grey-literature sources that utilized experimental or quasi-experimental designs to measure the impact of skills-based ALMPs on migration outcomes in L&MICs. Impact evaluations needed to be published in 1990 or later, and even though the search strategy was conducted in English, studies were not excluded based on language[1].

We also included programmatic evidence from qualitative studies, descriptive quantitative studies, process evaluations and project documents but only when they were linked to the interventions that were evaluated in the included experimental and quasi-experimental studies.

Table 2: PICOS inclusion criteria

Criteria	Included	Excluded
Criteria for inclusion in this systematic review		
Participants	People of any age and gender residing in low- and middle-income countries (L&MICs)	People in high-income countries
Intervention(s)	Skills-based active labor market policies (ALMPs) including classroom and on-the-job training interventions. This may include technical and vocational education training (TVET), business skills training, mentorships, internships/ apprenticeships, and entrepreneurship workshops. In order to be included, multi-component interventions need to have a substantive training component, i.e., training is provided to all participants in at least one treatment arm and the study provides a clear description of the content in the training curriculum.	Other ALMPs such as job search assistance programs, direct employment opportunities or public works schemes not focused on developing new skills or upskilling, self-employment promotion efforts and all else. Multi-component interventions with a training element that is not substantive because it is either not provided to all participants or it is not clear what the curriculum of the training was.
Comparison	• Comparison group constructed using randomization or statistical methods • Business as usual, including pipeline and waitlist controls • An alternative intervention	No comparator

[1] All restrictions related to publication date, language and publication status are from the EGM by Berretta and colleagues (2023), as that search was the basis for this review.

Criteria	Included	Excluded
Outcome(s)	Intermediate outcomes that influence eventual migration: • Intention to migrate • Knowledge, perceptions, attitudes, and expectations Final migration outcomes: • Attempted migration • Any migration behavior measure unspecified as to international and/or irregular • International migration flow (number of international migrants moving from their country of origin to another over a specific time period.) • International migration stock (total number of international migrants present in a given country at a particular point in time)	All else
Study designs	Experimental and quasi-experimental impact evaluations. See Berretta et al. (2023) for more details about eligible impact evaluation designs.	Efficacy trials, before-after with no control group, feasibility/ acceptability studies, reviews.
Qualitative Programmatic Evidence Inclusion Criteria		
Intervention(s)	• Skills-based active labor market policies or programs evaluated by the studies included in the systematic review	Skills-based active labor market policies or programs not evaluated by studies in the systematic review
Study designs	Descriptive studies, process evaluations, and other qualitative studies linked to programs in included impact evaluations: • *A qualitative study* using mixed- methods or qualitative methods to collect and analyze primary data on all of the following: the research question, procedures for collecting data, procedures for analyzing data, and information on sampling and recruitment. • *A descriptive quantitative study* using quantitative methods to collect primary data and conduct descriptive quantitative analysis on all of the following: the research question, procedures for collecting data, procedures for analyzing data, and information on sampling and recruitment, including at least two sample characteristics. • *A process evaluation* answering research questions about whether an intervention is being implemented as intended and perceptions of what is working well, and why. It may include collection of qualitative and quantitative data from different stakeholders to examine subjective questions, such as perceptions of intervention success or objective issues, such as how an intervention was operationalized.	Qualitative evidence not linked to the specific programs evaluated in the impact evaluations.

As indicated previously, we included studies evaluating skills-based ALMP interventions implemented in L&MICs of origin, even if they did not target areas with high prevalence of out-migration. That is, migration might not have been the focus of the programs and the impact evaluations. In such cases, we anticipate that studies may have measured migration generally but may not have distinguished the type of migration (e.g., international versus

internal, or through regular versus irregular channels). For this reason, we adopted an over-inclusive approach and excluded studies where it was clear that authors reported the results on internal migration only. The outcomes could be measured using a variety of indicators such as rates, proportions, occurrence, etc. Whenever available, we preferred outcomes associated with irregularity (e.g., irregular migration rates, intention to migrate irregularly), but based on the limited evidence on irregular migration (Berretta, Anda Leon, et al. 2023), we also extracted alternate outcomes such as intention to and final migration through regular channels, or migration unspecified as to whether it is regular or irregular.

3.2 Search strategy

We did not perform any new searches of impact evaluations for this systematic review given that the search for the EGM was conducted less than a year ago (between December 2022 and April 2023). The EGM search strategy was developed through a comprehensive and systematic process (Kugley et al. 2017). Details of the EGM search strategy are provided in Appendix A and in the EGM protocol (Berretta, Huang, et al. 2023).

We conducted a targeted search for qualitative studies, descriptive quantitative studies, project documents, and process evaluations of the interventions from the 10 included skills-based ALMPs impact evaluations. We followed systematic search guidelines developed by Snilstveit and colleagues (2012). The qualitative study identification process involved two stages of work, namely searching and screening. Using the names of programs from the included studies, we conducted internet searches on Google and Google Scholar as well as in the funder and implementer[2] websites of the identified programs.

3.3 Selection and coding of studies

3.3.1 Screening

The EGM screening processes included independent duplicate screening of titles and abstracts and of the full text of potentially includable studies. Coders met to reconcile decisions and when a disagreement was unresolved a member of the core team weighed in.

For this systematic review, we screened all impact evaluations from the EGM that had been coded as assessing ALMPs, applying the specific inclusion criteria to this review outlined in section 3.1. We also screened these impact evaluations for inclusion into the qualitative synthesis based on the use of mixed methods or reporting detailed descriptions of the context, design, and implementation of the programs. Finally, we single screened records obtained from the targeted search for qualitative programmatic evidence.

[2] List of funders and implementers includes the United States Agency for International Development (USAID); the Government of Norway; the European Union (EU); Mastercard Foundation; Private Enterprise Development in Low-Income Countries (PEDL); International Growth Centre (IGC); Global Fund; Japan Social Development Fund (JSDF); Poverty Alleviation Fund; International Initiative for Impact Evaluation (3ie); Center for Economic Policy Research (CEPR); York University Abu Dhabi; NORC at the University of Chicago; Winrock International; Government of Ghana; BRAC; Technical Education and Vocational Education and Training Authority (TEVETA)-Kenya; World Bank (WB); Ministry of Commerce-South Sudan; Partners of the Americas; Council for Technical and Vocational Education Training (COTVET)-Ghana; Innovations for Poverty Action (IPA)-Kenya.

3.3.2 Data extraction and coding procedures

We coded included impact evaluations for bibliographic, geographic information and substantive data, as well as standardized methods information. In addition, two reviewers independently extracted data on outcomes, population (including gender/age disaggregation, when available), and effect sizes corresponding to the outcomes indicated in section 3.1. Any discrepancies were reconciled through consensus, or with a third team member when necessary. We present data extraction templates in Appendix B.

For the programmatic evidence, we maintained independence of the data extraction and analysis processes from the coding of impact evaluations (Noyes et al. 2019). That is, a different researcher coded additional information and applied separate data extraction tools. These coding tools were applied to linked publications such as process evaluations, descriptive studies and other qualitative studies, and to qualitative evidence from included mixed methods impact evaluations.

3.3.3 Critical appraisal

Two reviewers independently appraised all of the included quantitative impact evaluations using 3ie's critical appraisal tool (Appendix C) to assess their internal validity (i.e., the extent to which observed changes in outcomes can be attributed to the intervention instead of other contextual factors or participants characteristics). 3ie's tool expands the bias domains of the Cochrane's ROBINS-I tool and Rob 2.0 (J. P. Higgins et al. 2016; Sterne et al. 2019) and covers potential risks of selection bias, confounding, implementation infidelity, performance bias, outcome measurement bias and reporting bias (see Appendix D for a conceptualization of these criteria). We assessed the risk of bias of included impact evaluations to identify potential threats to internal validity which could introduce bias into the results or conclusions.

We produced an overall rating for each study as either "High risk of bias", "Some concerns" or "Low risk of bias", drawing on the decision rules in RoB 2.0 (Higgins et al. 2016):

- "High risk of bias": if any bias mitigation measures were assessed as "No" or "Probably No."
- "Some concerns": if one or more bias mitigation measures were assessed as "Unclear", and none were "No" or "Probably No."
- "Low risk of bias": if all the bias mitigation measures were assessed as "Yes" or "Probably Yes."

Qualitative studies from the targeted search were not assessed for risk of bias. This decision was taken based on the fact that the search for qualitative evidence was not comprehensive to all literature available about ALMPs and migration outcomes, and there was a risk of excluding critical risk of bias studies and end with low numbers of studies to extract qualitative results from. The main purpose of the qualitative synthesis in this review was to complement the quantitative analysis and not necessarily answer the question on effectiveness but rather identify design, implementation, contextual and population factors that might have enabled or hindered the effectiveness of the programs.

3.4 Analytical approach for quantitative and qualitative data

3.4.1 Quantitative data synthesis

To provide summary effect estimates, we pooled together studies that we assessed to be sufficiently similar with respect to the type of outcomes being measured. We worked with

independent effect sizes; prioritizing outcomes based on comparability among studies when authors report more than one impact estimate for each of our analyses.

We combined studies using meta-analysis when two or more effect sizes reported on a similar outcome construct and the comparison group state was judged to be similar, similar to the approach taken by Lwamba and colleagues (2021). When we found too few studies or they were considered too heterogeneous in terms of outcomes, we present a discussion of individual effect sizes.

As heterogeneity exists in theory due to the variety of intervention characteristics and contexts included, we estimated inverse-variance weighted, random effects meta-analytic models (Higgins et al. 2020) using the metafor package (version 2.4.0; Viechtbauer 2010) in R software (version 4.3.0; R Core Team 2023).

Since all included studies reported results from multivariable linear regression models, we computed standardized mean differences (SMDs) following the approach suggested by Keef and Roberts (2004) using the regression coefficient and either the pooled standard deviation of the outcome, standard errors, t-statistics or significance levels, in that order upon availability of the data. We then adjusted SMDs using Hedges' method to deal with potential biases in cases where sample sizes are small, as described in our protocol (Anda Leon et al. 2023).

We assessed heterogeneity graphically using forest plots and by calculating the Q-statistic, I^2, and Tau^2 to provide an estimate of the amount of variability in the distribution of the true effect sizes (Borenstein 2009). We also explored heterogeneity through the use of moderator analyses through meta-regressions whenever the data allowed, that is when there were at least three effect sizes for continuous moderators and a minimum of two effects per cell for categorical moderators. We assessed aspects of the program and evaluation design such as duration of the intervention, modality of training, geographical region, migration programing focus, multicomponent nature, length of follow up period, mean migration for the control group and mean age of the pooled sample at baseline. We also aimed to examine whether the quantitative results were moderated by the overall risk of bias ratings, to explore whether there were systematic differences in intervention effects between primary studies with different risk of bias scores.

We ran sensitivity analysis by removing studies from the meta-analysis one-by-one and assessing changes in results. Studentized residuals and Cook's distances were used to examine whether studies were outliers and/or overly influential in the context of the model (Viechtbauer & Cheuuer and Cheung, 2010)[3].

We could not test for the presence of publication bias given that we did not find at least 10 studies for any of the intervention-outcome pairs we analyzed (Dalton, Bolen, and Mascha 2016). Finally, we did not need to correct for unit of analysis errors since all included studies using clustered designs reported standard errors appropriately.

[3] Studies with a studentized residual larger than the 100 × (1 - 0.05/(2 × k)) the percentile of a standard normal distribution are considered potential outliers (i.e., using a Bonferroni correction with two-sided α = 0.05 for k studies included in the meta-analysis). Studies with a Cook's distance larger than the median plus six times the interquartile range of the Cook's distances are influential.

3.4.2 Qualitative data analysis

We complemented statistical meta-analyses with a qualitative evidence synthesis (Noyes et al. 2019). After having completed the detailed coding of all of the included studies, we classified the codes according to factors related to context, intervention design and implementation, and population characteristics.

We adopted a thematic synthesis approach to identify analytical themes on intervention mechanisms and contexts that mitigate or reinforce intervention effects. Following Thomas and Harden (2008), we used inductive coding techniques to first identify common descriptive themes in the reported findings of the primary studies. We used EPPI-Reviewer's coding software to illustrate the link between the inductive codes in the primary studies and the identified descriptive themes. In a second step, following the identification of descriptive themes, we configured these into higher level analytical themes. Again, this configuration from descriptive to analytical themes was conducted in EPPI-Reviewer and we produced an overview table of both types of themes and their linkages, presented in the results section. .

The process of generating inductive codes, descriptive themes, and final analytical themes was configured around four analytical lenses derived from research questions 3 to 5 of this review. These refer to the interplay of context, intervention design, intervention implementation, and population characteristics with program effects, outlined in more detail in our protocol (Anda Leon et al. 2023).

3.4.3 *Integration of quantitative and qualitative findings*

We used quantitative impact evaluations to synthesize the effects of skills-based ALMPs on migration outcomes to address research questions 1 and 2. We then used qualitative evidence linked to programs evaluated in included impact evaluations to assess the influence of context, design and implementation, and population characteristics in the effectiveness and unintended consequences of these interventions to address research questions 3 through 5. In sum, the meta-analysis conducted with the quantitative data was complemented by a thematic synthesis of qualitative data extracted from programmatic sister publications of included impact evaluations.

4. Results

4.1 Characteristics of the evidence base

In this section, we provide an overview of the characteristics and distribution of the evidence base. We start by providing the results of the search and screening of the literature followed by a summary of the characteristics of interventions, included impact evaluations, and programmatic evidence.

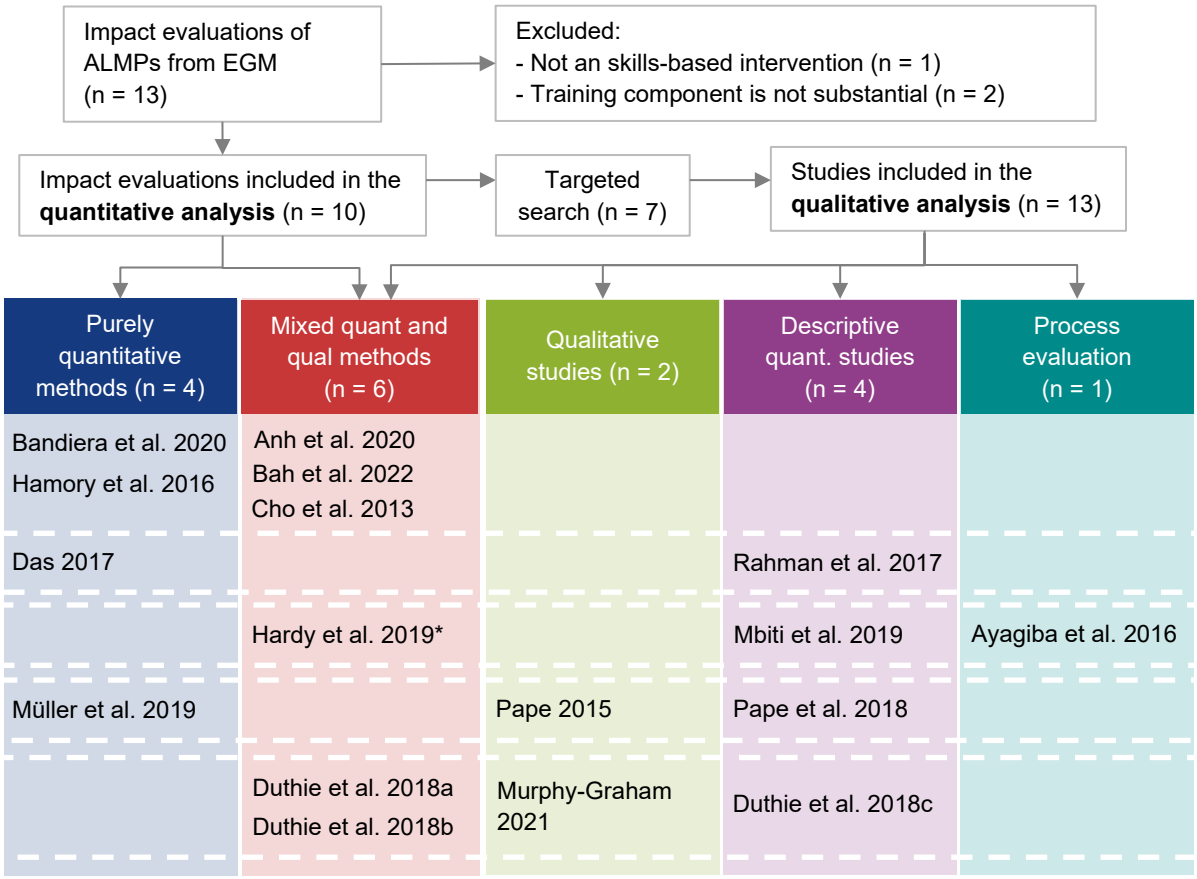
4.1.1 Search results

We examined studies included in the EGM by Beretta and colleagues (f2023). The EGM includes 13 impact evaluations looking at the effects of ALMPs. We screened those studies for inclusion with regard to the specific type of ALMP interventions within the scope of this review: skills-based on-the-job or classroom training interventions. Three of the 13 studies were excluded because the programs being evaluated did not have a substantial training component (i.e., the training was not offered to all beneficiaries, or the training curriculum

was not clearly reported; Figure 2)[4].

After screening, we included 10 impact evaluations in our quantitative analysis, six of which reported qualitative findings or enough detail on the context, design and program implementation to be included in the qualitative analysis as well. We found seven additional programmatic records that met the inclusion criteria for qualitative studies through internet searches of Google and Google scholar as well as the websites of the funders and implementers of the programs. Exclusion reasons in the screening of programmatic records pertained to the document type (e.g., project PowerPoint Presentations, newspaper articles, blogs, project announcements on funder or implementer websites and other informal reports). We included 13 studies for the thematic analysis of qualitative evidence about context, design, implementation and population of the interventions evaluated in the evidence base used in our quantitative analysis (six impact evaluations providing details on these themes and seven linked programmatic records)

Figure 2: Composition of included studies by study design



Note: *Hardy and colleagues (2019) did not use a mixed-methods approach but was included in the qualitative synthesis of this review given the detailed information provided on the context, design and implementation of the program. Rows marked by the white doted lines are grouping linked studies.

[4] Antwi and Phillips (2012) evaluate a wage reform for public health workers in Ghana, while Nepal (2016) and Regmi and Colleagues (2018) evaluate the Poverty Alleviation Fund in Nepal which had a training component, but its content was not clearly described by the papers and not all beneficiaries received it.

4.1.2 Characteristics of evaluated programs

The skills-based active labor market programs in the evidence base used in this review were implemented in Sub-Saharan Africa (n = 6), Latin America and the Caribbean (n = 2), and South Asia and East Asia (n = 2). Programs targeted similar populations but varied in the modality of implementation (on-the-job or apprenticeships vs. access to, or actual classroom-based training), content of the training, and intervention duration (Table 3).

Population characteristics

The programs were similar in terms of targeting young adults, with the average age at baseline ranging from 16 for the Skill Training for Advancing Resources (STAR) program evaluated by Das (2017), to 31 for the USAID/Cambodia Countering Trafficking in Persons (CTIP) Program evaluated by Anh and colleagues (2020; Table 3). The STAR program evaluated by Das (2017) was the only study examining differential effects by age groups, estimating the impact on migration outcomes separately for participants older than 17 years old and for participants 17 or younger.

In general, program participants had little experience with the labor market and high potential to benefit from training. For example, Bandiera and colleagues (2020), described beneficiaries of the vocational training program in Uganda as “individuals with limited labor market experience and much scope to learn about their job prospects" (p.3), while Cho and colleagues (2013) argued that “youth often lack the formal education or skills required to access salaried employment” (p.1) when reflecting upon beneficiaries of the technical and vocational education and training program in Malawi. The South Sudan Youth Business Start-Up Grant Program had literacy requirements as part of the program design. As a result, the participants were mostly educated, over half of the participants completed some secondary or post-secondary education, and only 15 per cent of the participants never attended any formal education (Pape et al. 2018).

Two programs specifically targeted potential migrants – a job-seeking platform with soft-skills training targeting youth at risk for labor-trafficking in Cambodia, evaluated by Ahn and colleagues (2020), and an information campaign with vocational skills training for male youth most likely to migrate in The Gambia, evaluated by Bah and colleagues (2022). The rest of the programs targeted vulnerable youth in Uganda (Bandiera et al. 2020), orphans or school dropouts (Cho et al. 2013), disadvantaged and unemployed youth (Das 2017), out-of-school youth (Hamory et al. 2015), or at-risk youth in general, without focusing specifically on participants’ migration decisions.

Intervention modality

Half of the studies evaluated classroom-based training programs such as vocational training as a standalone treatment. One study facilitated access to vocational training through vouchers in Kenya, but did not provide training directly (Hamory et al. 2015). The rest focused on apprenticeship approaches, either as standalone or combined with classroom-based training. The study assessing the STAR program in Bangladesh (Das 2017) evaluated both modalities and presented evidence on the additional effect of one component over the other. The A Ganar program, implemented in Honduras and Guatemala, also included both modalities, but these were not evaluated separately as all participants received both components (Duthie et al. 2018a and 2018b).

Interventions delivered through classroom training varied in their curriculum. Some focused on a mixture of technical, professional, and personal skills, such financial literacy, market assessments, and basic English language communication (Das 2017); life skills, language, math and information technology (Duthie et al. 2018a and 2018b); and business-and-life skills (Müller et al. 2019). One focused on skills relating to obtaining and performing in formal labor market opportunities, such as job searching, interview skills, money management, and action plans (Ahn et al. 2020). The rest of the classroom-based interventions focused on vocational training for technical occupations. For example, courses in construction, welding, carpentry and joinery, electrical installation and wiring, appliance and mechanical repair, plumbing, gas fitting, personal care services, or tailoring (Bah et al. 2022; Bandiera et al. 2020). Finally, one intervention provided access via tuition vouchers for private and government-run courses in multiple sectors instead of provision of the course itself (Hamory et al. 2015). This includes in sectors and services such as construction, textiles, mechanics, hair dressing and beauty, computers, secretarial, business.

Duration of the intervention

The average intervention duration was about 13 months. The programs with the shortest duration were the apprenticeship program in Malawi evaluated by Cho and colleagues (2013) and the classroom-based vocational training in The Gambia evaluated by Bah and colleagues (2022), implemented in three and four months respectively. The program with the longest duration was the National Apprenticeships Program (NAP) in Ghana, where apprenticeships typically lasted three years (Hardy et al. 2019). The average length of the training programs varied similarly between the two training modalities, with classroom interventions implemented in between four and 20 months and apprenticeships lasting between three and 36 months (Table 3).

Table 3: Characteristics of programs in included impact evaluation

Study	Country	Program name	Targeted population	Intervention description	Industry	Intervention duration	Funding agency	Implementing agency
Ahn et al. 2020*	Cambodia	USAID/Cambodia Countering Trafficking in Persons (CTIP) Program	Youth aged 18-39 from prevalent trafficking areas	Web-based job-search platform called “Bong Pheak”, job search training, technical and vocational education, job skills, financial literacy, goal setting, and grants to start a small business	Hospitality, construction	N/A	USAID	NORC and Winrock International
Bah et al. 2022*	The Gambia	N/A	Males aged 18-33 most likely to migrate from the village in the next 12 months	Information campaigns with testimonials about the risks of the journey, information and assistance for migration to a safer destination, and vocational skills training	Construction, manufacturing	6 months	European Union (EU)	Not reported
Bandiera et al. 2020*	Uganda	N/A	Disadvanta-ged youth aged 18-25	Vocational training, worker referrals to an established employer in a “good” sector	Construction, manufacturing, personal care services	6 months	Mastercard Foundation, PEDL, the IGC	BRAC and five reputable vocational training institutes (VTIs)
Cho et al. 2013	Malawi	Technical Education and Vocational Education and Training Authority (TEVETA) program	Vulnerable youth aged 15-24, mainly defined as orphans or school dropouts	Apprenticeship placements with master craftspeople, and vocational training	Arts, retail	3 months	Global fund	Government of Malawi through TEVETA

Study	Country	Program name	Targeted population	Intervention description	Industry	Intervention duration	Funding agency	Implementing agency
Das 2017*	Bangladesh	Skill Training for Advancing Resources (STAR)	Youth aged 14-18 years or disabled youth aged 15-21 years from poor households, who are out of school for at least a year	Apprenticeship placements with master craftsperson, vocational training, classroom-based training on specific trades, financial literacy, basic communication in English	Manufacturing, Arts, Professional services, personal care services	6 months	Japan Social Development Fund (JSDF) through the World Bank	BRAC
Duthie et al. 2018a	Honduras	A Ganar	At-risk youth aged 17-24	Sports-based field and classroom training on soft and professional skills, vocational training, internship/ apprenticeships, service training, mentoring, entrepreneur workshops, and employment referrals	Unspecified	7-9 months	USAID	Partners of the Americas
Duthie et al. 2018b	Guatemala	A Ganar	At-risk youth aged 16-24	Sports-based field and classroom training on soft and professional skills, vocational training, internship/ apprenticeships, service training, mentoring, entrepreneur workshops, and employment referrals	Unspecified	7-9 months	USAID	Partners of the Americas

Study	Country	Program name	Targeted population	Intervention description	Industry	Intervention duration	Funding agency	Implementing agency
Hardy et al. 2019	Ghana	National Apprenticeship Programme (NAP)	Youth aged 15-30	Apprenticeship placements, fee-free apprenticeship trainings	Personal care services, construction	18-48 months	Government of Ghana	Council for Technical and Vocational Education and Training (COTVET)
Hamory et al. 2015*	Kenya	N/A	Out-of-school youth aged 17-28	Vocational training vouchers to cover the tuition costs, material, uniforms and trade test fees for private and public vocational education programs, information on benefits of vocational training	Unspecified	20 months	Not reported	Innovations for Poverty Action-Kenya
Müller et al. 2019*	South Sudan	Youth Startup Business Grant Program	Youth aged between 18 and 34 with focus on young women	Unconditional cash grant, one-week business and life-skills training	Unspecified	12 months	Not reported	World Bank+ Ministry of Commerce

Note: Studies marked with an asterisk (*) use factorial design (i.e., there were treatment arms testing different combinations of activities).

4.1.3 Characteristics of included impact evaluations

While the evidence gap map implemented a systematic and comprehensive search for academic and grey literature, the impact evaluations included in this review came exclusively from grey literature sources, including technical reports (n = 4), working papers (n = 5) and a doctoral dissertation (n = 1; Table 4). We did not identify any impact evaluation looking at the effects of skills-based active labor market interventions on migration outcomes published in peer-reviewed academic journals.[5]

All included impact evaluations used an experimental design. Two of these, the ones that targeted irregular migration and complemented training with information campaigns in The Gambia (Bah et al. 2022) or job assistance in Cambodia (Ahn et al. 2020), used a cluster randomized design, randomizing at the settlement and commune level, respectively. The rest of the studies allocated treatment at the individual level, randomizing among eligible applicants (Bandiera et al. 2020; Müller et al. 2019; Duthie et al. 2018a and 2018b; Hardy et al. 2019; and Hamory et al. 2015) or eligible participants identified from a nationally representative sample (Cho et al. 2013) or from door-to-door visits (Das 2017). Five studies used mixed methods and provided qualitative evidence based on in depth interviews with participants (Ahn et al. 2020, Bah et al. 2022, Duthie et al. 2018a and 2018b) or the implementing agency (Cho et al. 2013).

Of the 10 impact evaluations, two reported effects on knowledge, attitudes and perceptions, three reported effects on intentions to migrate, four evaluated the effect of programs on attempted migrations, and six presented effects on migration behavior. Two studies looked specifically at international migration (Bah et al. 2022; Hamory et al. 2015), while the remaining studies did not report what percentage of the observed migration behavior was internal versus international. Given the nature of the intervention targeting potential migrants in the study by Bah and colleagues (2022), and the aim of the project to reduce irregular migration, the authors reported outcomes such as intention to migrate irregularly or having attempted to migrate through irregular channels. No other study looked at irregular migration related outcomes.

Cho and colleagues (2013), Das (2107), Müller and colleagues (2019), and Hardy and colleagues (2019), presented impact estimates by gender. Bah and colleagues (2022), on the other hand, targeted only male participants. This allowed us to estimate impacts separately for men and women whenever authors reported estimates on comparable outcomes.

Three studies provided cost evidence, one of which reported results of a cost analysis comparing earning gains of beneficiaries to the program costs (Das 2017). The author who reported a cost analysis suggested that on-the-job training interventions might be highly cost effective with regards to the effect on earnings. Bandiera and colleagues (2020) provided the average cost per trainee for a classroom based vocational training program but did not compare this to the value of long-term employment outcomes. Finally, Cho and colleagues (2013) discussed the cost of the average stipend dispersed to participants. We did not find any study reporting a cost-effectiveness analysis on any migration-related outcomes (Table 4).

[5] The EGM search strategy was implemented between December 2022 and April 2023. Since then, and until the last revision of this report, one of the included impact evaluations had been published in an academic journal. The reference to the publication is: Bah, T.L. et al. (2023) 'Can information and alternatives to irregular migration reduce "backway" migration from The Gambia?', Journal of Development Economics, 165, p. 103153. Available at: https://doi.org/10.1016/j.jdeveco.2023.103153.

Attrition rates ranged from as low as 7.65% in Das (2017), up to 30% in Cho and colleagues (2013), and 45% in Müller and colleagues (2019), but the attrition rates were not different between treatment and control groups for any of the studies. Five studies checked attrition balance across baseline characteristics, and one of them found evidence of a correlation between attrition and observables that were then controlled for in the analysis (Müller et al., 2019).

Finally, two studies reported obtaining ethical clearance (Bah et al. 2022; Bandiera et al. 2020). This is not to suggest that other included studies did not obtain such clearance, only that it was not explicitly stated in the impact evaluation report.

Table 4: Characteristics of included impact evaluations

Study	Program country	Type of publication	Evaluation design	Evaluation period	Research question	Migration-related outcome description	Main results	Cost evidence
Ahn et al. 2020*	CTIP-Cambodia	Technical report - USAID	Cluster RCT	Data collected at endline (0 months)	What is the take-up and effect of an internet-based job-matching platform combined with soft-skills training and vocational training on the awareness about human trafficking and willingness to migrate?	Main outcomes: Perception on whether human trafficking is a huge problem, whether migration can be a huge risk for human trafficking, attitude towards risk, and willingness to migrate	Participants were more likely to believe that human trafficking was a big problem in Cambodia, but they did not change their willingness to migrate.	Not reported
Bah et al. 2022*	Information campaigns and vocational training-The Gambia	Working paper – WB	Cluster RCT	6 months	Does providing better information and testimonials about the risks of the journey, facilitating migration to a safer destination, or offering vocational skill training to enhance domestic employment opportunities reduces irregular migration intentions and behavior?	Main outcomes: Knowledge of irregular migration channels, over-optimism about backway migration, likely or surely migration within five years, consider backway, attempted backway, migrated to Europe	Vocational training reduced intentions to migrate the backway and the number of steps taken toward moving. However, the backway migration rate from The Gambia collapsed, even in the control group, resulting in no space for a treatment effect on irregular migration.	Not reported

Study	Program country	Type of publication	Evaluation design	Evaluation period	Research question	Migration-related outcome description	Main results	Cost evidence
Bandiera et al. 2020	Vocational training and job assistance-Uganda	Working paper - PEDL	Individual RCT	12 months	How individual search strategies vary with exogenous variation of: (i) the vocational skills (sector-specific vocational training); and (ii) information about labor market prospects (light-touch job assistance), and how these translate into long-run outcomes for workers.	Secondary outcomes: Has attempted to migrate to find a job (one of five components of a search intensity index)	Trained participants increase their belief over job-offer arrival and the distribution of expected earnings and search more intensively along multiple margins (time devoted to job search and channels used including attempting to migrate to find a job), but with time they change their job search strategy because of a lack of call backs. In the long-run, skilled workers have better employment outcomes.	Training costs provided at $470 per trainee. Policy implications suggest that job-assistance may have substantially higher returns than skills-based ALMP if designed and target optimally
Cho et al. 2013*	TEVETA program-Malawi	Working paper – IZA	Individual RCT	1 month	How does on-the-job development of technical skills affect self-reported skills and knowledge, and subjective measures of well-being? How do high drop-out rates affect estimates? And how do the results differ by gender?	Secondary outcomes: Migrated permanently or temporarily, for work, school or other	Training significantly increased the self-reported skills and knowledge that the training was meant to impart, and improved subjective measures of well-being. These results were not sensitive to drop-outs. The program reduced the likelihood of migrating away in search of employment, which is consistent with trainees making some longer-run investments.	The stipend provided for the participants (of 4300 MWK, or USD$28 on average) was not sufficient to cover transportation and lodging costs.
Das 2017	STAR-Bangladesh	Peer-reviewed thesis/ dissertation – UC Berkeley	Individual RCT	6 months	What is the effect of on-the-job training on labor market outcomes (employment and earnings)? and how does this effect vary if	Secondary outcome: Migration (internal or international)	On-the-job training increases both self- and wage employment and has positive effects on earnings. Additional effect of classroom training was statistically insignificant	A cost-benefit analysis of on-the-job training component of the program

Study	Program country	Type of publication	Evaluation design	Evaluation period	Research question	Migration-related outcome description	Main results	Cost evidence
					classroom training is compounded with on-the-job training?		for all outcomes except for hours of work in self-employment and earnings from this employment. Training did not affect migration.	shows that it is highly cost effective. Assuming that the life of benefit is 44 years, and that benefits do not change overtime, benefit-cost ratio for on-the-job training is estimated to be 6.34
Duthie et al. 2018a*	A Ganar-Honduras	Technical report – USAID	Individual RCT	18 months	Does the program increase the likelihood that youth will obtain and maintain jobs, return to school, start their own business or reduce risky behavior?	Secondary outcomes: Consider migrating out of Honduras and if so whether the respondent has attempted to migrate and, if so, whether to the United States	No effect was found on employment. Positive impacts on job quality including higher wages, higher prevalence of benefits and higher job satisfaction. Participants were more likely to attempt migration.	Not reported
Duthie et al. 2018b*	A Ganar-Guatemala	Technical report - USAID	Individual RCT	18 months	Does the program increase the likelihood that youth will obtain and maintain jobs, return to school, start their own business or reduce risky behavior?	Secondary outcomes: Consider migrating out of Guatemala and if so whether the respondent has attempted to migrate and, if so, whether to the United States	No effect was found on employment, but it increased the likelihood of having a work contract and starting a business. It also increased the desire to return to school. The program did not have an impact on migration rates.	Not reported
Hardy et al. 2019*	NAP-Ghana	Working paper – WB	Individual RCT	12 months	What are the short-run effects of apprenticeship	Secondary outcome: Migrated (0/1)	Apprenticeships shift youth out of wage work and into self-	Not reported

Study	Program country	Type of publication	Evaluation design	Evaluation period	Research question	Migration-related outcome description	Main results	Cost evidence
					training on labor market outcomes?		employment. However, the loss of wage income is not offset by increases in self-employment profits in the short-run. Further, the program offer increased the probability of migration.	
Hamory et al. 2015	Vocational training vouchers-Kenya	Technical report – 3ie	Individual RCT	2 months	What is the effect of vocational education on formal sector employment and labor market earnings, participation in the informal and agricultural sectors, entrepreneurship decisions, and a range of other life outcomes?	Residence outside of Kenya at time of survey or since program launch	There was limited evidence that the program increased earnings. A positive effect on wages was found among wage earners. No effect was found on the probability of living outside of Kenya.	Not reported
Müller et al. 2019	Youth Startup Business Grant Program-South Sudan	Working paper – WB	Individual RCT	15 months	What are the socioeconomic, behavioral, and psychological consequences of a business start-up grant program cancellation?	Migration index: Standardized weighted average of having moved since baseline, living outside, in a refugee camp, in an IDP camp, and having the wish to move	Originally planned treatment improved consumption, savings, and psychological well-being. However, participants who vainly expected to receive the cash grant showed reduced levels of consumption and women among this subgroup also experienced strong reductions in their trust level there was some evidence that they were less likely to migrate. No other effect was found on propensity to migrate.	Not reported

Note: Studies marked with an asterisk (*) included mix-methods approaches for data collection or provided detailed information in their description of the context and program design.

4.1.4 Characteristics of included studies with qualitative evidence from programs

For the qualitative synthesis of evidence in this review, we used five included impact evaluations using mixed methods (Ahn et al. 2020; Bah et al. 2022; Cho et al. 2013; Duthie et al. 2018a and 2018b), one quantitative impact evaluation that provided detailed information on the context and program design (Hardy et al. 2019), and the seven studies identified from the targeted search for linked papers, which included qualitative studies (n = 2), descriptive quantitative studies (n = 4), and a process evaluation (n = 1; Table 5).

Table 5: Characteristics of included programmatic studies linked to impact evaluations

Study	Program - Country	Title	Type	Methods used	Informant
Pape 2015	Youth Startup Business Grant Program-South Sudan	Republic of South Sudan Youth Startup Business Grant Program-Volume I: Program rationale, design and implementation	Qualitative study	Interviews/ Survey	Beneficiaries and control group
Murphy-Graham 2021	A Ganar-Honduras and Guatemala	Life Skills Education for Urban Youth in Honduras and Guatemala: A Capability Analysis of the Sports-Based Job Training Program A Ganar.	Qualitative study	Interviews	Beneficiaries (Youth)
Pape et al. 2018	Youth Startup Business Grant Program-South Sudan	Impact of program cancellation due to conflict in South Sudan: a chapter of the South Sudan poverty assessment 2017	Descriptive quantitative studies	Interviews	Beneficiaries
Mbiti et al. 2019	NAP-Ghana	Training for success: targeting and incentives in apprenticeship training in Ghana. International Initiative for Impact Evaluation	Descriptive quantitative studies	Interviews/ Survey	Beneficiaries
Duthie et al. 2018c	A Ganar-Honduras and Guatemala	A Ganar alliance impact evaluation synthesis report Guatemala and Honduras	Descriptive quantitative studies	Interviews/ Case study	Beneficiaries/ beneficiaries parents/ Staff members (facilitators and technical trainers)
Rahman et al. 2017	STAR-Bangladesh	The effects of skill training on livelihoods: Evidence from BRAC's Intervention on School Dropout Adolescents	Descriptive quantitative studies	Interview/ Case study	Participants/ Non-participants
Ayagiba et al. 2016	NAP-Ghana	Performance audit report of the auditor-general on the governance of the national apprenticeship programme by the council for technical and vocational education and training	Process evaluation	Interviews/ Focus group discussions	Key personnel/ Implementing agencies/ Master Craft Persons/ Apprentices

4.1.5 Quality of impact evaluations and risk of bias

As all ten included quantitative studies used an experimental design, we used the critical appraisal tool version that assesses criteria for randomized control trials (RCTs). In Appendix D we detail the criteria used to assess the risk of bias of included studies. One study was assessed as having a low risk of bias overall (Bah et al. 2022), while the remaining nine studies were assessed as having a high risk of bias. The ratings of each specific risk of bias domain and overall score for each included impact evaluation is presented in Table 6, while the summary of assessments by domain is presented in Figure 3. The most common concern among high risk of bias studies was related to potential performance bias, with some studies noting that control participants often knew their assigned treatment status. It is possible that control participants may have sought other training opportunities once aware that they would not be receiving the intervention.[6] The nature of the interventions also made the blinding of participants challenging, if not impossible. However, simple measures can help prevent or attenuate performance bias. For example, Bah and colleagues (2022) reported not telling participants about the experiment (i.e., about beneficiaries being compared to a control group) or that alternative interventions were offered in other settlements. Several studies also noted contamination issues, for example by control participants managing to receive the training (e.g., Cho et al. 2013; Hamory et al. 2015). Using clusters to allocate participants to treatment and control groups can help achieve implementation fidelity by preventing spill-overs (Ahn et al. 2020; Bah et al. 2022).

Figure 3: Risk of bias assessment of included impact evaluations

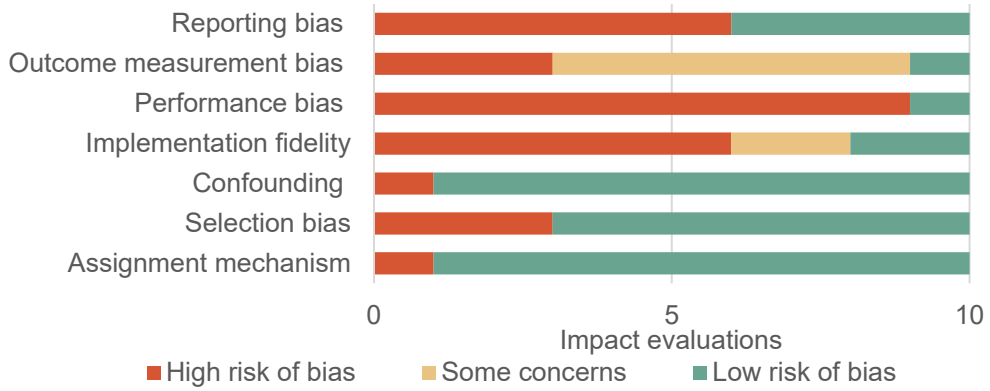
Study	Sources of bias							Overall RoB Rating
	Assignment mechanism	Selection	Confounding	Implementation fidelity	Performance	Outcome measurement	Reporting	
Ahn et al. 2020	🔴	🔴	🔴	🟡	🔴	🔴	🔴	High risk of bias
Bah et al. 2022	🟢	🟢	🟢	🟢	🟢	🟢	🟢	Low risk of bias
Bandiera et al. 2020	🟢	🟢	🟢	🔴	🔴	🟡	🟢	High risk of bias
Cho et al. 2013	🟢	🔴	🟢	🔴	🔴	🟡	🔴	High risk of bias
Das 2017	🟢	🟢	🟢	🟡	🔴	🟡	🔴	High risk of bias
Müller et al. 2019	🟢	🟢	🟢	🟢	🔴	🟡	🟢	High risk of bias
Duthie et al. 2018 a	🟢	🟢	🟢	🔴	🔴	🔴	🔴	High risk of bias
Duthie et al. 2018 b	🟢	🔴	🟢	🔴	🔴	🔴	🔴	High risk of bias
Hardy et al. 2019	🟢	🟢	🟢	🔴	🔴	🟡	🟢	High risk of bias
Hamory et al. 2016	🟢	🟢	🟢	🔴	🔴	🟡	🔴	High risk of bias

🟢 Low risk of bias - no issues identified
🟡 Some concerns - it is unclear if there is an issue
🔴 High risk of bias - potential risk of bias identified

Note: An unclear rating is typically given when the study does not provide enough information about the domain being assessed. For the overall rating to be "low", all criteria should be in green. The study is rated as "Some concerns" when at least one criterion is yellow and none red. The overall RoB is "high" when any of the domains are red.

[6] Geographical distance between treatment groups and blinding procedures can help ameliorate potential issues with performance bias.

Figure 4: Risk of bias assessment by criterion assessed



4.2 Quantitative impacts

Below we present the results of the synthesis of included quantitative impact evaluations on intermediate outcomes that influence eventual migration (knowledge, perceptions, attitudes and expectations of migration, and intention to migrate) and migration behavior (attempted migration, and migration). Results are presented as standardized mean differences. Average effects from meta-analyses are denoted by SMD while the standardized mean differences of single studies are denoted by g. As all the outcomes used for pooling studies and estimating average treatment effects were binary outcomes, positive SMDs reflect a higher likelihood of observing the outcomes for the treatment group than for the control group. The quantitative syntheses presented in this section each rely on a relatively small number of included studies, thus results should be interpreted with caution.

4.2.1 Intermediate outcome: Knowledge, perceptions, attitudes and expectations

Two of the studies in our sample reported effects on outcomes in this category.

Ahn and colleagues (2020) found that the job-seeking platform and soft-skills training intervention in Cambodia did not change the perception that migration can be associated with a big risk for human trafficking ($g = 0.02, [95\%\ CI: -0.07\ to\ 0.11]$)[7]. With $g = 0.02$, 50.8% of the "treatment" will be above the mean of the "control" group.

Bah and colleagues (2022) evaluated the effects on knowledge about "backway" or irregular migration and over-optimism on wages and on asylum likelihood, finding no significant effects ($g = 0.02, [95\%\ CI: -0.05\ to\ 0.09]$; and $g = -0.002, [95\%\ CI: -0.07\ to\ 0.07]$ respectively). However, these outcomes are of relevance mainly for the information campaign on the risks of irregular migration component delivered along the vocational training one.

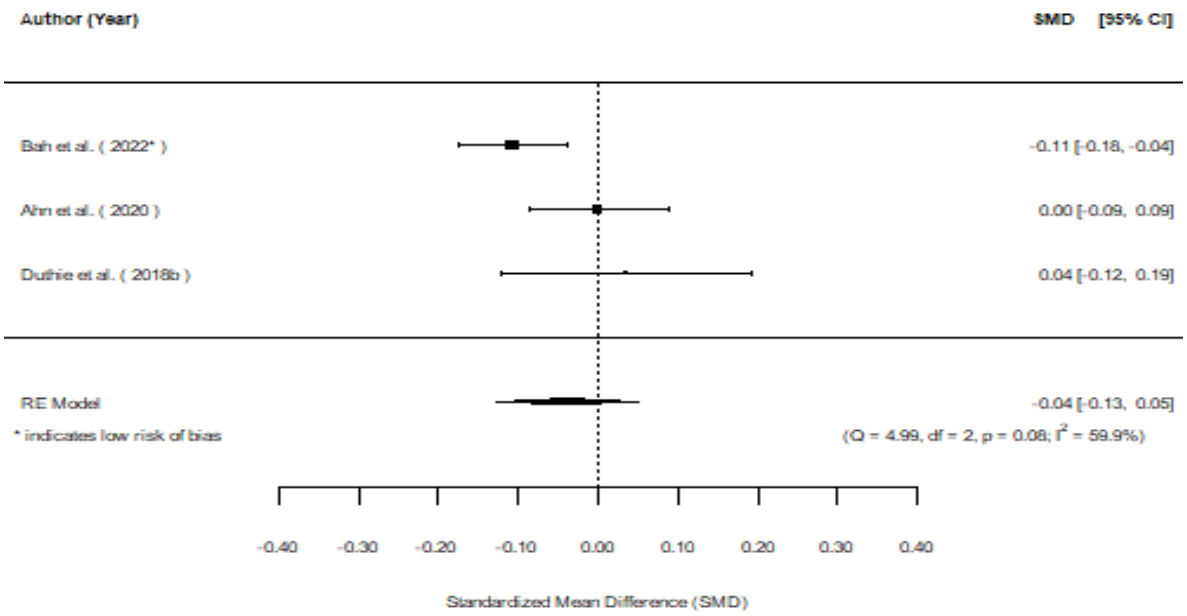
No study reported the effects of skills-based ALMPs interventions on perceptions and the desire for change, feelings of stagnation arising from life in the origin country, and challenges due to conditions that cannot be addressed, nor on expectations about wages in destination countries. These are theory-based intermediate outcome categories that were included the framework of the EGM by Berretta and colleagues (forthcoming), from which we built our review framework.

[7] The standardized mean difference for this effect was computed using the exact p-value of the coefficient provided directly by the authors ($p = 0.07$).

4.2.2 Intermediate outcome: Intention to migrate

Three studies evaluated the effect of skills-based ALMPs on intentions to migrate: the information campaign with vocational training for males likely to migrate in The Gambia (Bah et al., 2022), the CTIP job-seeking platform with soft-skills and workplace professionalism training in Cambodia (Ahn et al., 2020), and the A Ganar program providing job training in Guatemala (Duthie et al., 2018b). The observed outcomes ranged from -0.11 to 0.04. The intention to migrate was on average lower in intervention participants compared to control participants by $\text{SMD} = -0.04$ (95% CI: -0.13 to 0.05), but this difference was very small in magnitude and was not statistically significant ($p = .40$)[8]. A forest plot showing the observed outcomes and the pooled estimate based on the random-effects model is shown in Figure 4. Positive SMDs reflect a higher probability of intending to migrate for participants in the treatment group compared to the control group. One study found an effect on intentions to migrate significantly different from zero (Bah et al. 2022), where the vocational skills training delivered along an information campaign with testimonials about the risks of the journey to potential male migrants, was associated with a reduction in the intention of beneficiaries to migrate. Participants in the control group showed a probability of intending to migrate of 52% and the program reduced this for the group of beneficiaries by 7%, equivalent to a reduction of 0.11 standard deviations.

Figure 5: Forest plot showing the observed outcomes and the estimate of the random-effects model of skills-based ALMPs on intention to migrate



While tests of heterogeneity were not significant, some heterogeneity may still be present in the true outcomes ($Q(2) = 4.99$, $p = .08$, $\hat{\tau}^2 = 0.004$, $I^2 = 59.93\%$). There was no indication of outliers in the context of this model. However, because three studies reported effects for this outcome category, moderator analyses were not appropriate for categorical characteristics such as modality of the training or risk of bias assessment, and tests of publication bias are not valid. Only Bah and colleagues (2022) had a low risk of bias and was the only study reporting a reduction in intention to migrate. While tests for moderation are

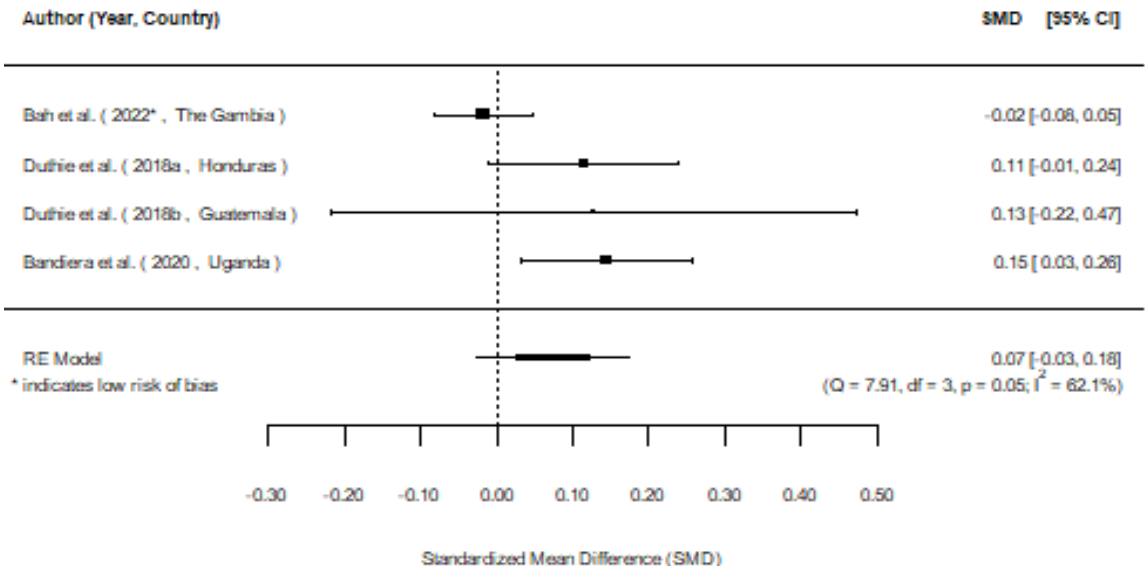
[8] The standardized mean difference for Anh and colleagues (2020) used in this analysis was computed using the exact p-value reported directly by the authors for this coefficient ($p = .99$)

exploratory given the small number of studies, we found that none of the continuous moderators tested were a significant source of heterogeneity. Indeed, the follow-up period in months ($\widehat{B} = 0.002, p = 0.55$ [95% CI: -0.005 to 0.009]), the intervention period in months ($\widehat{B} = 0.002, p = 0.75$ [95% CI: -0.01 to 0.02]), and the average age of participants at baseline ($\widehat{B} = 0.001, p = 0.95$ [95% CI: -0.02 to 0.02]) were not statistically significant predictors of the differences in impact estimates observed among the included studies for this analysis[9].

4.2.3 Final outcome: Attempted migration

Four studies evaluated the effects of skills-based ALPMs on attempted migration: the information campaign with vocational training for males likely to migrate in The Gambia (Bah et al. 2022), the A Ganar program providing job training in Honduras and Guatemala (Duthie et al. 2018a and 2018b), and the vocational training for disadvantaged youth on Uganda (Bandiera et al. 2020). The observed outcomes ranged from -0.02 to 0.15. Attempted migration was on average higher in intervention participants than control participants by $SMD = 0.07$ (95% CI: -0.03 to 0.18), but this difference was not statistically significant ($p = .14$). A forest plot showing the observed outcomes and the pooled estimate based on the random-effects model is shown in Figure 5. Positive SMDs reflect a higher probability of attempting migration for participants in the treatment group than for those in the control group. One study found an effect on attempting migration significantly different from zero (Bandiera et al. 2020), where the sector-specific vocational training in Uganda was associated with a significant increase of 0.15 standard deviations in the probability of attempting migration.

Figure 6: Forest plot showing the observed outcomes and the estimate of the random-effects model of skills based ALMPs on attempted migration



An examination of the studentized residuals revealed that one study (Bah et al. 2022) may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Bah and colleagues (2022) would increase the overall average effect ($SMD = 0.13, [95\%\ CI\colon\ 0.05\ to\ 0.21]$), and the effect would become significant (z =

[9] We were not able to test whether the proportion of participants who intended to migrate in the control group was a significant predictor of differences in the effect sizes because Duthie and colleagues (2018b) did not report this figure.

3.18, p = .002). However, this effect would be based on just three studies, all of which were assessed as having a high risk of bias, so we would caution against making too much of this result, noting that replication in high-quality studies would be needed.

The outcomes appear to be heterogeneous ($Q(3) = 7.91$, $p = .05$, $\hat{\tau}^2 = 0.01$, $I^2 = 62.09\%$). While we were able to test several moderators, given the number of included studies this should be considered an exploratory analysis and the results should be interpreted with caution[10].

Intervention duration in months and mean age of participants were the only statistically significant moderators:

- Each additional month of the intervention duration increased attempted migration by 0.05 standard deviation units ($\widehat{B} = 0.05, p = .01$ [95% CI: 0.02 to 0.10]).
- Each additional year in the average age of the sample decreased attempted migration by 0.04 standard deviation units ($\widehat{B} = -0.04, p = 0.01$ [95% CI: -0.07 to -0.01]).

None of the other moderators tested were a significant source of differences in the effects of the included studies. Indeed, follow-up period in months ($\widehat{B} = 0.01, p = 0.10$ [95% CI: -0.002 to 0.03]), and whether the program being evaluated took place in Latin America ($\widehat{B} = 0.06, p = 0.59$ [95% CI: -0.16 to 0.29]) were not statistically significant[11].

4.2.4 Final outcome: Migration

Two studies reported the effects of the program on international migration specifically, the study by Bah and colleagues (2022) looking at the information campaign and vocational training in The Gambia, and Hamory and colleagues (2015) looking at the program delivering vouchers for private vocational education programs and government training institutes in Kenya. We decided to combine these two with the effects reported on any migration (unspecified to whether it happened internally or internationally; n = 5)[12]. We used moderator analysis to explore differences between these two groups and estimated the model using only the five effect sizes as a sensitivity check.

---

[10] We could not run moderator analyses based on study design since all included studies used experimental approaches. With only one study assessed as having a low risk of bias (Bah et al. 2022) and the rest a high risk of bias, we were unable to run sensitivity analyses by the quality of the evidence. Only one included study specifically targeted migration, so we were also not able to test whether targeting migration exclusively was a potential source of variation in average outcomes. Finally, we were not able to explore if including components other than training was a source of heterogeneity because only one of the studies included in this model utilized an additional component (i.e., Bah and colleagues (2022) included information about the risks involved in migrating irregularly to Europe).

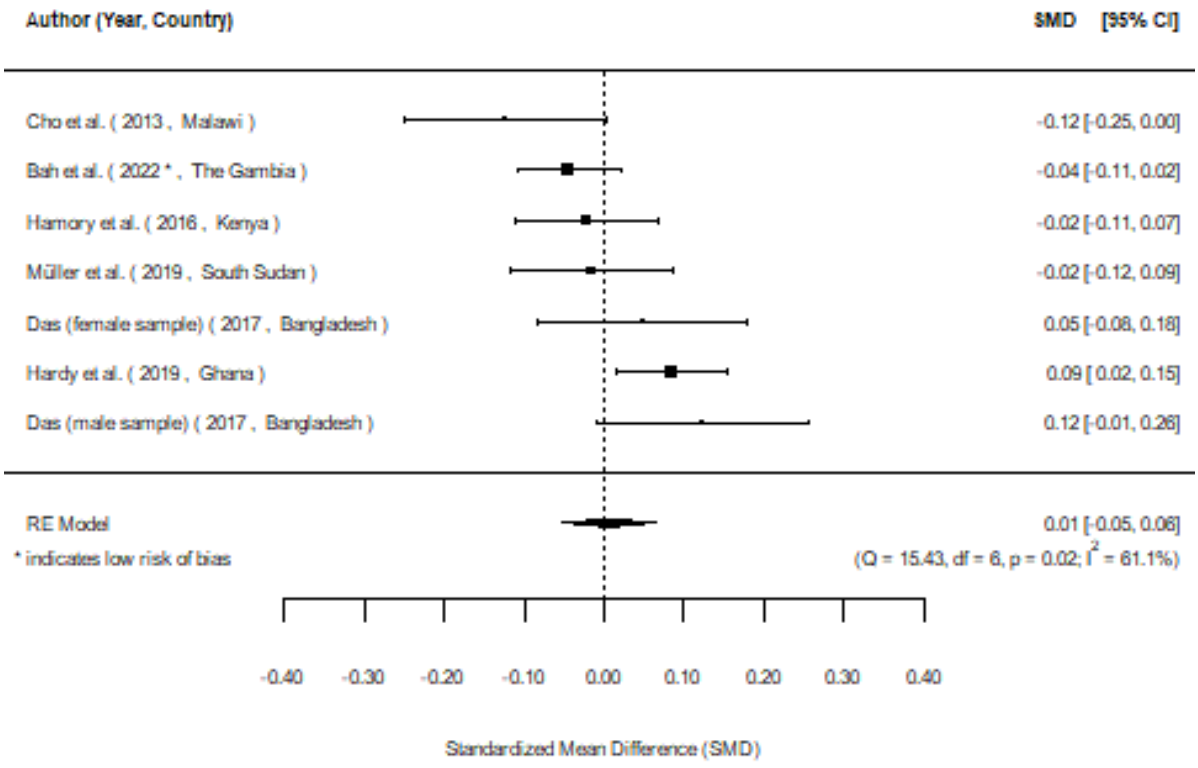
[11] We were not able to disentangle the effect of implementation in Latin America (LA) from on-the-job vs. classroom modality since there was perfect multicollinearity between the two variables. The studies that took place in LA also evaluated apprenticeship programs, while the studies elsewhere looked at classroom training programs.

[12] These five studies either indicated that the migration behavior measure being observed included both internal and international or did not specify the nature of it. For example, Das (2017) used the variable "Migrated (yes=1, no=0)" and indicated explicitly that it included international and internal migration, while Cho and colleagues (2013) used the variable "Migrated permanently or temporarily, for work, school or other" but did not define whether this migration behavior referred to internal, international or both. We included these studies because we cannot rule out that their results include international migration.

We included a total of seven independent effect sizes from six studies in the analysis of skills-based ALMPs on migration, either international or unspecified as to its nature. The study by Das (2017) evaluating the STAR program in Bangladesh reported the impacts for men and women separately. Given that estimates used different samples, we included both in the model.

The observed outcomes ranged from -0.12 to 0.12, and migration was on average slightly higher among intervention participants than the control group by $SMD = 0.01$ (95% CI: -0.05 to 0.607), but this difference was very small in magnitude and was not statistically significant ($p = .83$)[13]. A forest plot showing the observed outcomes and the estimate based on the random-effects model is shown in Figure 6. Positive SMDs reflect a higher probability of migrating for participants in the treatment group than for those in the control group. One study found an effect on migration behavior significantly different from zero (Hardy et al. 2019), who evaluated the impact of the National Apprenticeship Program (NAP) in Ghana and found that it was associated with a significant increase of 0.09 SMD in the probability that beneficiaries had migrated.

Figure 7: Forest plot showing the observed outcomes and the estimate of the random-effects model of skills-based ALMPs on migration



The outcomes appear to be heterogeneous ($Q(6) = 15.43$, $p = .02$, $\hat{\tau}^2 = 0.004$, $I^2 = 61.12\%$). There was no indication of outliers in the context of this model.

The effect size used in this model from the study by Müller and colleagues (2019) who evaluated the cancelation of an unconditional cash grant combined with business- and life-

[13] Results did not change when we only pooled effects sizes from the five studies reporting outcomes on migration behavior unspecified as to whether it was international or not ($SMD = 0.03$, 95% CI: -0.05 to 0.11, $p = .51$)

skills training in South Sudan, is an average effect for all treatment groups receiving the training on business- and life-skills (*training, no grant*, and *training and grant*). However, the authors also report the effects on migration for the groups of beneficiaries that were assigned to receive the grant, attended the training as a pre-requisite for receiving the grant, but did not receive the grant because the program was cancelled prior to the grant distribution. The results show that, in contrast to the average effect, participants who mainly expected to receive the grant were less likely to migrate ($g = -0.09, [95\%\ \text{CI}: -0.19\ \text{to}\ 0.02]$), though the difference was not statistically significant. The authors interpret this as an anticipation of receiving the grant, rather than an actual effect of the training component specifically. Sensitivity analysis leaving this study out in the context of this model does not change our results.

The results of studies looking specifically at international migration did not differ significantly from the results of studies that did not specify the nature of migration ($\widehat{\text{B}} = -0.62,\ p = .30$ [95% CI: -0.18 to 0.05]). None of the other moderators we were able to test in the context of this model were statistically significant. This included exposure to intervention in months ($\widehat{\text{B}} = 0.003,\ p = .12$ [95% CI: -0.001 to 0.01]), evaluation period in months ($\widehat{\text{B}} = 0.01,\ p = .25$ [95% CI: -0.005 to 0.02]), whether the training took place on the job instead of in a classroom ($\widehat{\text{B}} = 0.07,\ p = 0.17$ [95% CI: -0.03 to 0.17]), whether the program used components other than training ($\widehat{\text{B}} = -0.06,\ p = .36$ [95% CI: -0.18 to 0.06]), mean age of participants ($\widehat{B} = -0.01, p = 0.33$ [95% CI: -0.03 to 0.01]), and the mean of the control group ($\widehat{\text{B}} = 0.01,\ p = .95$ [95% CI: -0.45 to 0.48])[14].

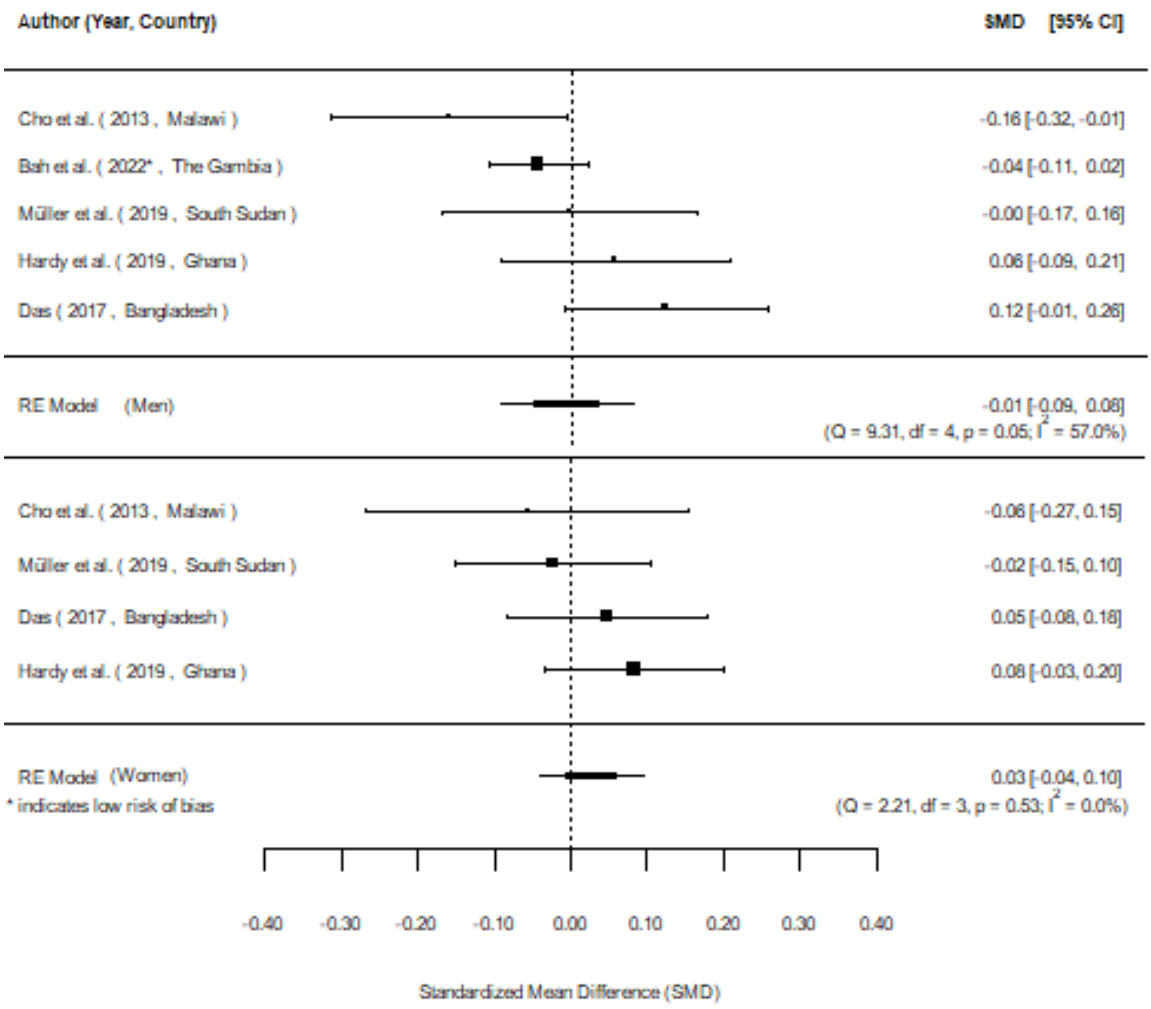
Subgroup analysis by gender

We were able to run an additional subgroup analysis by gender for the migration outcome. Since the program evaluated by Bah and colleagues (2022) in The Gambia targeted only male participants, the analysis for men included one more study ($k = 5$) than the analysis for women ($k = 4$).

For the effects of training on male migration, the observed outcomes ranged from -0.16 to 0.12 and male participants migrated less frequently than control participants by $\text{SMDm} = -0.01$ (95% CI: -0.09 to 0.08). For female migration the observed outcomes ranged from -0.06 to 0.08 and female intervention participants migrated more than female non-participants by $SMDf = 0.03$ (95% CI: -0.04 to 0.10). However, in both cases, the group differences were not statistically different from zero ($p_m = .87$; $p_f = .41$). There were no indications of outliers for either group of studies. Forests plots showing the observed outcomes and the pooled estimates based on the random-effects model for both samples are shown in Figure 7. The only study finding an effect on migration behavior significantly different from zero was Cho and colleagues (2013), who evaluated the technical and vocational education program in Malawi and found that male participants were on average 0.16 SMD less likely to have migrated permanently or temporarily, for work, school or other, after participating in the program.

[14] We could not run moderator analyses for the design (all experimental) and the risk of bias rating of included studies (only one at low risk of bias). We were also not able to test whether targeting migration exclusively was a source of potential heterogeneity because only one of the included studies in this analysis did. Neither were we able to test differences by geographic region given that none of the studies in the context of this model took place in Latin America.

Figure 8: Forest plots showing the observed outcomes and the estimates of the random-effects models of skills-based ALMPs on migration by gender



There was no significant amount of heterogeneity in the true outcomes for women ($Q(3) = 2.21$, $p = .53$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). For men, heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q(4) = 9.31$, $p = .05$, $\hat{\tau}^2 = 0.01$, $I^2 = 57.05\%$). Once again, the moderator analyses were exploratory given the small number of studies in each analysis. That said, none of the moderators we tested for males were statistically significant. For the sample of studies evaluating the impacts for men, exposure to intervention in months ($\hat{B} = 0.003, p = .44$ [95% CI: -0.01 to 0.01]), evaluation period in months ($\hat{B} = 0.01, p = .32$ [95% CI: -0.01 to 0.03]), whether the training took place on the job instead of in a classroom ($\hat{B} = 0.04, p = .71$ [95% CI: -0.17 to 0.24]), whether the program included components other than training ($\hat{B} = -0.04, p = .71$ [95% CI: -0.24 to 0.17]), and the average migration rate of the control group ($\hat{B} = -0.35, p = .48$ [95% CI: -1.32 to 0.63]), did not help explain heterogeneity of results. For the studies reporting impacts for women, moderator analyses were not appropriate as there was no heterogeneity among the effects.

4.3 Qualitative findings (context, barriers, and enablers)

We conducted a thematic synthesis on the 13 studies reporting qualitative evidence related to the included skills-based active labor market interventions (7 qualitative studies and 6

impact evaluations using mixed-methods of detailed description of the context and implementation). As indicated above, this thematic synthesis identified themes related to the intervention design, intervention implementation, target population, and context. In total, we identified 17 descriptive themes, which we configured into six analytical themes (Table 7). These six analytical themes present the synthesis results and are discussed in more detail below.

Table 6: Overview table of descriptive and analytical themes

Descriptive themes based on the inductive coding of primary studies	Analytical themes derived from the configuration of descriptive themes
Outreach strategies Program targeting	**1. Design factors that may encourage program enrollment and uptake (design)** Skills-based active labor market policies may benefit from awareness-raising activities at the start of the program. Sensitizing target populations should be considered as this can facilitate the realization of desired impacts by encouraging an understanding of program objectives and uptake[15] of the programs. Recruitment procedures that help target participants that could benefit the most should also be considered.
Multifaceted communication platforms and activities Applied (as opposed to theory-based) training modalities tailored to local contexts Gender and equity-sensitive curriculum and activities	**2. Design factors that could promote engagement and learning (design)** Designing and providing multifaceted communication platforms and activities for training and communications including lectures, videos, group discussions, group roleplay exercises and sports might be considered in the design of the intervention as they may be useful for engagement and learning. Equally important is that training activities are tailored to the local context (e.g., language or geographic location) and are gender and equity sensitive.
Governance and accountability Communication channels with and within collaborators Timely disbursement of tools, funds and certificates	**3. Implementer and participant program design features that act as barriers and enablers to program implementation and success (design/implementation).** To promote the successful completion of skills-based active labor market policies, mechanisms to ensure that implementing partners clearly understand their authority and responsibilities are key. It is also important to establish detailed guidelines including communication channels. To facilitate successful programmatic outcomes, participants should be provided certificates of completion in a timely manner. Obstacles and delays to accessing tools and fees are also potential bottlenecks that should be addressed to ensure program completion and participant satisfaction.
Distance of training facilities Logistical support to attend training Timing and length of training	**4. Other design and implementation features that can encourage participation (design/ implementation).** The location, accessibility, and convenience of the training sessions appear to be important determinants of attendance. To encourage participation, providing safe and accessible training locations may be necessary. It is also important to ensure that the training is not conducted during periods that may interfere with other activities and responsibilities.

[15] Uptake refers to beneficiaries being incorporated into the program

Descriptive themes based on the inductive coding of primary studies	Analytical themes derived from the configuration of descriptive themes
Gender roles and norms Network and support systems National context	**5. Contextual factors that may affect the successful implementation of interventions (context)** Interventions may not be successfully implemented in some contexts due to disease outbreaks, violence and conflicts, political turmoil or structural constraints that reduce employment opportunities. Support from family and local authorities may be important in ensuring participation whilst networks may become key for participants to secure employment after the completion of training. Gender norms may mean that women spend more time taking care of the family and performing household and agriculture-related chores, potentially preventing women from taking full advantage of the training opportunities.
Socioeconomic status of participants Demographic/population characteristics Attitudes and interest	**6. Participants characteristics that may determine program take-up (population)** When designing interventions, implementers may need to consider how they will include participants from different socioeconomic statuses. This could help avoid, for instance, excluding those who are less educated, have no access to the internet and young individuals who might migrate during the program implementation. The health status of individuals can inhibit participation in program activities or securing employment.

4.3.1 Analytical themes related to the design of the program

Theme 1: Design factors that may encourage program enrollment and uptake

We found some qualitative evidence suggesting that community outreach via awareness-raising campaigns and sensitization campaigns may lead to increases in enrollment . For instance, the Youth Startup Business Grant Program in South Sudan which included an unconditional cash grant combined with business- and life-skills training targeting the youth, conducted various promotional activities to encourage applicants from the target population. These included banners, flyers, posters, t-shirts, radio and print ads, and other materials. This approach reportedly yielded robust engagement, with 8,240 submissions ultimately returned of the over 9,000 invitations distributed (Pape 2015). Some of these outreach strategies were also implemented by other programs included in this review. For instance, the Council for Technical and Vocational Education and Training (COTVET) publicized Ghana's youth-focused National Apprenticeship Program (NAP) through radio broadcasts. Community representatives, including the Technical and Vocational Education Training (TVET) coordinators from the Ghana Education Services, also distributed program materials along with flyers (Hardy et al. 2019). However, authors did not discuss the effectiveness of these strategies and we caveat that program implementers should limit using demand-driven enrollment methods which require participants to seek programs, as they are often subject to selection bias (Gertler et al. 2016). Participants who are able to access demand-driven programs could be those who are better off socioeconomically. Implementers and program designers may want to consider instead the approach adopted by the USAID/Cambodia Countering Trafficking in Persons (CTIP) Program. There, program implementers used community identification methods such as conducting program meetings with village leaders, village volunteers, key informants, schoolteachers, representatives of

community-based organizations, and other stakeholders to create a list of potential participants who preliminarily fit the criteria for eligibility, who were then invited to participate in the program (Ahn et al. 2020).

Clear program communication, not only during outreach activities, but throughout the implementation of the program might be necessary for its continuation. For example, Pape (2015) found that both beneficiaries and staff of the Youth Startup Business Grant Program in South Sudan identified that a lack of program clarity posed a barrier to participation. In particular, beneficiaries expressed uncertainty regarding the program's purpose.

The design of the application can also be used to target the population of interest and identify participants who could benefit the most from the intervention. For example, Pape (2015) described how the applications for the Youth Startup Business Grant Program in South Sudan were only available in English in order to filter applicants that could read and write and "therefore be better equipped to absorb the provided training" (p. 43). Further, applications required formal proof of age but were flexible to accept an assessment of age provided by a hospital as valid for applicants who did not possess a National ID or birth certificate.

Theme 2: Design factors that could promote engagement and learning

Some of the qualitative findings in the evidence base used in this review suggest that participants from classroom-based interventions appreciated multifaceted communication platforms not only for the training but also for additional communicational components. For example, the USAID/Cambodia Countering Trafficking in Persons (CTIP) program used various techniques to educate beneficiaries on the risk of migration and human trafficking and to deliver job-seeking and soft skills training, including lectures, videos, group discussions, and role-playing activities. Both beneficiaries and staff expressed a shared belief in the effectiveness of videos as educational tools. One beneficiary emphasized the value of videos stating "*Yes… If there were just words, it wouldn't be too engaging, because it's hard to remember. But with videos and pictures, the villagers find it easier to understand"* (Ahn et al. 2020, 40).

Other studies reported themes surrounding the pedagogical modality. Participants from the Bangladesh Skill Training for Advancing Resources (STAR) program, which provided apprenticeships and classroom-based vocational training to disadvantaged and unemployed/under-employed youth, suggested the creation of video modules specifically targeting technical trades such as mobile servicing and fridge repair, to enhance the quality of learning (Rahman et al. 2017). Participants in the A Ganar program in Honduras appreciated the role of sporting activities as a training method. The program utilized soccer and other team sports to engage with at-risk youth on topics of youth unemployment. One participant reported, *"I learned a great deal from A Ganar through sports. I remember they had us engage in team activities, like tying my foot to a classmate's, and we had to cooperate as we walked and ran together. This taught me that effective communication with a teammate can lead to successful outcomes, and I've carried this lesson with me. Even now, in my current job, not everyone may be fond of me, but I consistently provide support, just as I learned to do in sports"* (Duthie et al. 2018c, 22).

We also found evidence of participants reporting that pragmatic learning objectives were a priority, as opposed to theory-based learning. In the Cambodia CTIP program, some

beneficiaries found activities to lean too heavily towards theory rather than practical application. One beneficiary emphasized the importance of hands-on experience, stating, "*For me, I think they should give us the real practical [sic] after learning. If they don't give us any practices, there aren't any good results at all. Don't just only give a speech without any training."* Another participant from the animal raising program indicated, "*The problem with raising chicken is that there are lots of diseases, so I want a practical experience on how to deal with it. That is why both training and practicing are needed at the same time. I want them to teach us how to protect chickens from getting diseases" (Ahn et al. 2020, 43).* Similarly, a participant from the Bangladesh STAR program reported that her education in the trade was enriched by both hands-on and theoretical training. She contended that her mentorship under the Master Craft Person (MCP) not only covered the program's curriculum but also encompassed additional techniques and knowledge (Rahman et al. 2017).

For additional intervention components that communicate sensitive topics related to irregular migration, messages conveyed through activities and platforms can more effectively prepare participants for real-life scenarios if they are designed to facilitate trust and openness. For instance, one beneficiary, who participated in the initial phase of the Bong Pheak/Soft Skills training, described a job search and risk mitigation role-playing exercise conducted during the Cambodia CTIP program as both enjoyable and effective in reducing stress given the nature of the topic related to labor exploitation and trafficking. A program staff member echoed this sentiment, noting, *"The participants seem to have fun. Like I've said, when I joined, for those who lack the courage, they seem to be braver, more willing to talk and express opinions after they've watched the video"* (Ahn et al. 2020, 40).

We also found some evidence suggesting that curriculum and training could benefit if designed to reflect the local context and the characteristics of participants. For example, the Cambodia CTIP program's use of videos was beneficial for participants with low literacy levels compared to traditional instruction methods involving written content on a board (Ahn et al. 2020). In the Gambian Dakar migration information and vocational skill training program, all videos were translated and presented in the three primary languages spoken in the regions: Mandinka, Fula, and Wolof. Respondents were able to watch videos in their chosen language using tablets (Bah et al. 2022).

Finally, some evidence further suggested that program design could benefit from gender and equity sensitive approaches. Ayagiba and colleagues (2016) reported that the National Apprenticeship Programme (NAP) implemented by COTVET in Ghana did not actively promote inclusion of female apprentices in trades traditionally dominated by males which resulted in gender imbalances in the selection of trades offered by the program. The authors recommended providing gender-specific guidance and counselling to Master Craft Persons (MCPs) and apprentices to motivate or facilitate female NAP apprentices to enter male-dominated trades.

4.3.2 Analytical themes related to the design and implementation of the program

Theme 3: Implementer and participant program design features that act as barriers and enablers to program implementation and success

From the descriptive themes we had coded from the qualitative evidence, we identified the lack of adequate governance and accountability systems, and pre-implementation planning measures as potential determinants of the successful implementation of the interventions.

Audit findings for the Ghana's NAP program, for example, concluded that the first phase of NAP did not achieve its objectives because there was no regular feedback from coordinators and master craft persons (MCPs) to guide COTVET in addressing emerging challenges during implementation (Ayagiba et al. 2016). Although COTVET had a governance framework and strategy for the implementation of NAP, there were no clear communication channels or accountability mechanisms to ensure that partners clearly understood their authority and responsibilities. This includes providing partners with apprenticeship guidelines, which then affected recruitment, selection, training, and assessment of apprentices as initially planned. Without proper guidance and reinforcement mechanisms, Ghana Education Service (GES) officials working as NAP coordinators lost enthusiasm for the program, and eventually neglected their duties. There was also no structured system for overseeing plans, programs, and activities, which would enable timely corrective measures when needed.

Another potential barrier to program success is the lack of adequate resources for participant program completion or for practicing the trade on their own. Ayagiba and colleagues (2016) identified other implementation issues in the Ghana's NAP related to the supply of tools and equipment necessary for the types of trade during and after apprenticeships and with the payment of fees to MCPs. COTVET had not provided a sufficient number of tool sets for all apprentices, nor had it remitted apprenticeship fees to all MCPs. This seemed to have been caused by inaccurate record keeping of MCPs and apprentices. As a result, some apprentices indicated that they were unable to set up their own businesses after training. Others indicated that they stayed in the training even after the stipulated period had lapsed because in addition to the lack of tools for training, they still needed to acquire more practical skills. Some apprentices left the program before the one-year stipulated training period out of frustration because they were denied the use of their tools during training and were not allowed to use tools that belonged to their masters (Ayagiba et al. 2016). Relatedly, Rahman and colleagues (2017) found that participants in the Skills Training for Advancing Resources (STAR) program implemented by BRAC in 2012 remained unemployed, despite possessing the skills, because they did not have adequate start-up capital.

Finally, we found evidence to suggest that offering participant incentives such as timely provision of certificates may enhance the successful completion of skills-based ALMPs. Certificates can serve as an important credential to help participants secure employment quickly. Trainees of the Bangladesh STAR program indicated that the post-graduation certificate acquisition process was lengthy and cumbersome, and therefore, they requested to expedite this process. However, it was a useful tool for entry into the formal labor market. One participant mentioned that the certificate she received upon finishing the training helped her validate her qualifications to prospective employers (Rahman et al. 2017).

Theme 4: Other design and implementation features that can encourage participation

We found qualitative evidence suggesting that the facility in which training takes place, its location, and how accessible and convenient it is, appear to be important determinants of the program's uptake and attendance. Based on interviews with the participants of the Youth Startup Business Grant Program in South Sudan, Pape (2015) reported that the distance to the training venue as well as limited time for arranging transportation and logistics were some of the reasons for non-attendance. Similarly, Bah and colleagues (2022) found that distance to the training venue, along with the difficulty in finding accommodation, were reasons for not taking up the training offer. Cho and colleagues (2013) also found that girls in the Technical

Education and Vocational Education and Training Authority (TEVETA) program were more likely to drop out if they lived far away from the training center. This was not the case for boys who were more likely to attend regularly. One strategy to deliver the program close to participants' place of residency was observed in the NAP program in Ghana which encouraged applicants to list trainers within walking distance from their place of residence, while trainers who were further away were acceptable provided applicants had reliable means of transportation (Hardy et al. 2019).

Other design features That are important to consider include the timing of the training session and the length of the program. For instance, the training sessions in the study by Bah and colleagues (2022) were delayed at the participants' request because most of them were engaged in farming activities and were unable to leave their farms until the end of the rainy season. In Ghana, an apprenticeship program originally planned 12-months ended up lasting between 18 months to nearly four years based on what trainers decided was sufficient (Hardy et al., 2019). The Skill Training for Advancing Resources (STAR) program also identified recommendations by some participants to increase the length and frequency of the training. One such suggestion was to provide an advanced option for graduates who wished to specialize in a specific trade's intricacies (Rahman et al. 2017).

4.3.3 Analytical themes related to the context in which programs are implemented

Theme 5: Contextual factors may affect the successful implementation of interventions (context)

Gender norms

Understanding how gender norms affect women's participation may be key for ensuring women can take full advantage of training opportunities. Women's responsibilities linked to household chores, childcare and agricultural responsibilities can affect both the extent to which women can participate in training opportunities as well as their ability to find and engage in subsequent employment. For example, in the evaluations of the Technical Education and Vocational Education and Training Authority (TEVETA) program, Cho and colleagues (2013) found that women spent almost twice as much time as men on household and agricultural chores, preventing them from fully capitalizing on the training opportunities being offered. Meanwhile, men's responsibilities were more likely to carry market returns and the skills men developed outside the home may have also allowed them to make better use of the training. Further, in the USAID/Cambodia Countering Trafficking in Persons (CTIP) program, staff members reported that "*sometimes, husbands don't allow their wives to come because there is a baby that the wife needs to take care of. So even though we invite them to come, they cannot come because of the baby"* (Ahn et al. 2020, 47). Reasons listed by participants for not attending the program activities included family obligations such as taking care of younger family members, sick family members, or elders in their household (Ahn et al. 2020). This is similar to participants of the vocational skill training implemented in Gambia, where young people stated that they could not take up the training offer because they had too many responsibilities in their home villages (Bah 2022).

Not only do gender norms affect women's participation in the programs, but they also seem to affect their ability to eventually find and engage in employment opportunities. For instance, after taking part in the A Ganar program, one participant reported that she had *"lost faith in finding work"*, acknowledging that *"now with my daughter, it is difficult"* (Murphy-Graham 2018, 174).

Social networks support for program uptake and success

Support from fellow participants, friends, neighbors, family, and local authorities can impact the extent to which participants can benefit from training programs. For example, staff members from the CTIP program reported that the impact was enhanced when beneficiaries learned from one another (Ahn et al. 2020). Participants of the Cambodia CTIP program activities reported that authority figures were a key influence on their participation. One beneficiary stated that she attended a program meeting because *"the village chief only told me that I needed to go."* This is echoed in a statement by another beneficiary who decided to attend a program meeting *"because he's the chief of the village. Whatever he asks us to do, we should do it* (Ahn et al. 2020, 44)

Participants from the STAR project reported that they were able to take full advantage of the program because their families supported them morally and even financially (Rahman et al. 2017). Many trainees availed better jobs with the help and recommendation of their relatives and neighbors. Social networks were highlighted as an important pathway to materialize the program benefits as *"having friends or relatives close to the training center is a very strong predictor of whether trainees – both males and females – can complete training"* (Rahman et al. 2017, 15). Conversely, most unsuccessful respondents stated that their family members were unable to support them, and they did not have any influential relatives or neighbors who could help them to find a job or provide them capital to start their own venture (Rahman et al. 2017).

National contexts influence in programs success

National contexts in which skills-based ALMPs are implemented may also affect the successful implementation of these programs and the power of studies to measure the impact. For example, the COVID-19 pandemic led to a nine-month suspension of the Gambia Technical Training Institute (GTTI) training program, and when the program resumed only half of the participants returned (Bah et al. 2022). Study participants also reported numerous contextual reasons for reductions in irregular migration. First, the change in government in Gambia, which shifted the country away from an autocratic political regime, was seen to be linked to increased chances of asylum denial upon reaching Europe and higher risks of deportation upon detection without documents. Numerous news articles reported on the government's agreement with the European Union to repatriate unauthorized migrants in return for financial aid (Bah et al. 2022). Furthermore, financing migration costs became more challenging (which can be explained by the increased risks), prompting some individuals to pursue seasonal migration to urban centers and Senegal as a means to gather funds for their eventual migration through unofficial channels. All these contextual factors were listed by the authors as reasons for a decreased rate of migration during the time of the study, making it harder for the evaluation to find any impacts of the information campaign and training on migration.

In Ghana, the 2012 elections resulted in a change in the political regime, which delayed the implementation of the National Apprenticeship Programme by about a year (Hardy et al. 2019). The USAID/Cambodia Countering Trafficking In Persons (CTIP) program experienced unique challenges, as there was a disease outbreak among the chickens and this affected participation (Ahn et al. 2020).

The A Ganar program was implemented in Honduras and in Guatemala, where participants expressed frustration about the lack of formal work opportunities (Murphy-Graham 2018).

Participants shared that*, "…opportunities are what we need more than anything. I think that there are a lot of people here like me, that have potential, but they get lost"*, and that *"…the situation is a bit complicated in this country" (*Murphy-Graham 2018, 173). On top of the lack of opportunities, Duthie and colleagues (2018c) also found that insecurity prevented participants from finding employment, as observed by a participant who had turned down a good job offer because it was unsafe for her to walk back to her neighborhood at the time her shift would have ended. The lack of opportunities and high levels of crime and violence, coupled with poor transportation infrastructure and gang-imposed curfews, were beyond the scope of the program and ultimately meant that youth are not able to take full advantage of the training programs and eventually secure employment.

A final example of the importance of context, comes from the South Sudan Youth Business Start-Up Grant Program which could not be successfully implemented because the country is extremely fragile, with active conflict in parts of the country. South Sudan has challenges dealing with the legacy of over 50 years of conflict and continued instability which has led to development deficits evident with the country having some of the lowest human development indicators in the world, continuing violence and instability, and minimal levels of infrastructure development (Pape 2015). The violence in the country forced many of those who were a part of the program to migrate, reducing the number of participants (Pape et al. 2018). This also affected the evaluability of the program as the original control and treatment groups could not be located. Furthermore, the program was terminated prematurely: *"Escalating violence at the end of 2015 forced the program to terminate the disbursement of the grants before all participants had accessed them. Completion of the program was first postponed and finally cancelled to mitigate the perceived risk for beneficiaries to become the target of crime. In addition, it was seen as major risk that the conflict might be exacerbated if grant money got into the wrong hands and was used to purchase arms"* (Pape et al. 2018, 8).

4.3.4 Analytical themes related to the targeted population

Theme 6: Participants characteristics may determine program take-up

The socio-economic status of the participants may affect the successful implementation of skills-based ALMPs. Participants of the USAID/Cambodia Countering Trafficking in Persons (CTIP) program, for example, faced some barriers in terms of access to the internet and lack of experience with internet-based apps which hindered the success of the programs (Ahn et al. 2020). For the vocational training implemented in Malawi, Cho and colleagues (2013) reported that participating in the training was expensive. Trainees, especially girls, had to draw down their savings to participate. Some programs, however, were responsive to the barriers faced by participants. For instance, participants of the vocational skills training implemented in Gambia complained about financial pressures in sustaining themselves while studying, and in response the program added a monthly stipend of 1,000 GMD conditional on regular attendance (Bah et al. 2022). Participants from the Youth Business Start-Up program in South Sudan who had a formal bank account, found it easier to access the grant (Pape et al. 2018). Rahman and colleagues (2017) reported that less educated trainees from the STAR program struggled to understand the content covered in theory classes.

The health status of participants might also affect the extent to which they can benefit from the program. One participant of the Skills Training for Advancing Resources (STAR) program completed the training but could not look for wage employment due to her deteriorating health condition caused by a chronic illness (Rahman et al. 2017).

Participants in the A Ganar program expressed difficulties in finding employment associated with coming from marginalized neighborhoods, with presence of gangs, where transportation can pose a risk for youth. Yet, the authors of the impact evaluation found that the program had helped one participant who was part of a gang to abandon the path to violence, gangs, alcohol, and drug addiction (Duthie et al. 2018c).

Age of participants is another factor that may affect program participation. For instance, the CTIP program found that the job-seeking platform training was primarily attended by older participants; this was because many younger at-risk participants had already moved or migrated away from their villages and thus could not participate in these activities (Ahn et al. 2020). Further, younger trainees from the STAR program struggled to learn the trade and this experience led to an obstruction of their career progress (Rahman et al. 2017).

The attitudes and interests of participants can also affect their willingness to take part in the program. Participants in the CTIP program cited being interested in the specific topics offered as one of the reasons for participating in the program activities (Ahn et al. 2020), while some youth in the treatment group of the vocational training implemented in Gambia said they did not take part in the training because they were not interested in the training subjects (Bah et al. 2022).

In some instances, such as in the case of the A Ganar program (Duthie et al. 2018c), participants were not looking for employment because of external factors, namely, having young children that require care, having health issues and not having work qualifications or required employment documents. One of the participants of the program shared the following sentiments: "*The truth is that I want to find work, but I haven't been able to. I haven't been able to look because almost everywhere they ask for a certificate from Tercero básico (upper primary),"* yet at the moment he is not able to obtain his certificate (Duthie et al. 2018c, 30). Similarly, some participants from the STAR program were unemployed at the time of the follow-up survey because they were not interested in wage employment (Raham et al. 2017). Conversely, some self-motivated participants were active in finding jobs even without any support from the program or their social network.

5. Discussion

5.1 Summary and application of findings

Overall, the studies included in this review assessed programs that were initially not designed to target irregular migration but rather, targeted underlying economic insecurity issues that are conceived by our theory of change as a root cause of irregular migration. Therefore, we still consider their results relevant for informing future programs and research aiming to address the lack of economic opportunities, perceptions of migration, and, ultimately, migration aspirations and behavior. However, the evidence base of rigorous impact evaluations in this field is emerging and with the small number of effectiveness studies available we were not able to find evidence of an effect of skills-based ALMPs on migration outcomes. For the outcomes where we were able to conduct a quantitative meta-analysis, we found non-significant effects on the intermediate outcome of intention to migrate ($\mathrm{SMD} = -0.03$ (95% CI: -0.13 to 0.07), k = 3), as well as for the final outcomes of attempted migration ($SMD = 0.07$ (95% CI: -0.03 to 0.18), k = 4) and migration ($SMD = 0.01$ (95% CI: -0.05 to 0.07), k = 7). However, given limited sample sizes, with at most seven estimates

being pooled, and the majority of studies being rated as high-risk of bias, these findings should be taken as exploratory and interpreted with caution. Additional high-quality evidence will be needed to draw robust conclusions about the effectiveness of ALMPs in improving intermediate outcomes that influence eventual migration or migration behavior..

Insights gleaned from the qualitative evidence suggest that key assumptions in the theory of change may not have been met and that implementation challenges may diminish programmatic impact. Several studies reported low uptake, implementation fidelity issues, and disconnects between programmatic approaches and context (e.g., labor market requirements, gaps in appropriate design for participants, etc.). Encouraging enrollment and uptake are key to the realization of desired impacts. Program design and implementation elements that can encourage participation include outreach and sensitizing target populations in the early stages of the program, providing certificates that serve as credentials in a timely manner, ensuring programs are appropriately designed (location, language, program duration and intensity, accessible and safe) and planned for, and providing sufficient incentives. Context and population characteristics such as socio-economic and demographic characteristics of participants, and their attitudes and interest, may influence participation as well and should be considered.

Although most of the studies in this review did not focus on migration as a primary outcome of interest, the two studies that did can help us interpret our results. For instance, Bah and colleagues (2022) found that vocational training along with an information campaign on the risks of irregular migration deterred intention to migrate through unofficial channels in The Gambia, but they did not find any effect on actual irregular migration outcomes. However, this may be because contextual factors had reduced migration in the time of the evaluation so that only 1% of the control group had attempted irregular migration. The authors note that it is “unclear whether these programs [which offer the hope for better jobs at home, through building skills with a tuition-free vocational skills program] reduce migration, or instead promote it by giving youth marketable skills that they can use at destination” (Bah et al. 2022, 4). This is in line with our theory of change, where we hypothesize that training programs for active labor market participants could deter irregular migration by increasing opportunity costs of leaving countries of origin, or by increasing access to regular migration through increased access to labor opportunities in countries of destination. However, the available evidence does not allow us to explore whether either mechanism explains our findings.

Another possible explanation for the lack of evidence of significant effects may be a disconnect between the programs and the targeted populations. The studies in this review focused mainly on vulnerable youth populations. According to Card and colleagues (2015), these types of participants benefit more from “work first” programs such as job search assistance, rather than training and other human capital development interventions, which are more appropriate for long-term unemployed participants. Given the scope of our review and limited evidence, we are not able to test this, but one of the included studies looked at both types of interventions (a job-search platform and a soft and professionalism training in Cambodia) and did not find an effect of either (Anh et al. 2020). Further, Das (2017) explored whether age may have been a reason for the null effects detected for migration of the Skills Training for Advancing Resources (STAR) program in Bangladesh. Eligible participants in the study were mostly 18 years or younger and may not have been allowed to be away from their families. We attempted to examine this by conducting subgroup analysis by age, but results were still statistically insignificant for both male and female participants who were 18 years

old or older. Heckman and colleagues (1999) argue that the relatively unskilled and less able populations, like those targeted by this type of program, may be the reason for the generally small impacts of ALMPs.

Alternatively, the lack of results on migration outcomes could be a consequence of unmet assumptions about the nature of the problem and gaps in mechanistic pathways earlier in the theory of change. (See section 5.1.3 for more details of studies that discuss lack of changes in employment outcomes). As suggested by Carranza and McKenzie (forthcoming), the main issue is not that jobseekers lack the skills required by the economy but rather the "shortage of good wage jobs in economies dominated by small and informal firms. As a result, job training and job search policies by themselves are unlikely to generate a lot of new employment." (p. 23). Thus, skills-based ALMPs cannot address supply-side labor market gaps, making it unlikely that downstream outcomes such as migration decision-making will be affected.

Despite our efforts to explore potential explanations for the lack of significant effects, the small evidence base restricts our opportunities for quantitative synthesis using meta-analysis and our ability to draw generalizable conclusions. In addition, it limits our power to explore potential sources of heterogeneity in the outcomes, explore contextual barriers and facilitators for the design and implementation of programs, and identify population characteristics associated with program success or failure. Given the limitations of this narrow evidence base, results should be interpreted with caution. Rather than concluding that skills-based ALMPs are not effective in deterring irregular migration or intentions to migrate through irregular channels, we would emphasize the need for more evidence of programs with explicit focus on irregular migration and targeting of populations vulnerable to these risks before robust conclusions can be drawn.

A further challenge in drawing strong conclusions is the lack of high-quality evidence. Only the study by Bah and colleagues (2022) evaluating the information campaign and vocational skill training in The Gambia was assessed as having a low risk of bias and should be taken as an example of how to produce high quality evidence for irregular migration programing. In addition to having low-risk of bias, it is also an example of how to use creative ways to measure challenging outcomes such as irregular migration, including through the use of technology. The authors used WhatsApp location sharing, which was induced by mobile credit compensation, to complement geo-coded and time-stamped in-person follow-up surveys[16]. Further, the authors used a mixed-methods approach, complementing the analysis of their findings with qualitative discussions with respondents to identify possible reasons for the small effects found in their quantitative analysis of impacts on migration to Europe. The authors highlighted the outbreak of COVID-19 pandemic, which disrupted the implementation of training, but also a reported deterrent to migrating during the evaluation period, along with political context which increased the risks and costs of irregular migration.

[16] Researchers should be transparent about obtaining consent, anonymizing, blinding and other processes applied for the use of sensitive data such as geolocation. For this study, the authors reported obtaining ethical clearance but did not provide details about the strategy used to deal with potential identification risks in collecting and including such data in their analysis.

5.1.1 Gaps and limitations in included studies

Given the limited evidence base of rigorous impact evaluations, more research is needed in the field more broadly. Specific limitations include a lack of research that targets migration specifically, migration measures that do not allow us to understand whether migration occurred through regular or irregular channels, authors not being specific in the reporting of whether migration measures referred to national or international, a lack of mixed methods research, and a relatively short-term follow-up period on average for included interventions. These limitations are discussed in turn below.

Null effects on migration outcomes could be a result of programs not targeting migrants. Though our exploratory analyses did not indicate a significant difference between the two programs targeting migrants compared to the other studies not specifically targeting migrants, we were not sufficiently powered to draw strong conclusions from these results. The analysis was also not powered to explore this potential source of heterogeneity for any outcome beyond migration.

Further complicating our ability to disentangle effects on migration, the migration measures used by included studies did not specify whether it occurred through regular or irregular means, nor if the measure referred to internal or international migration or in which proportion if both. This was likely related to the fact the interventions were not targeting migration specifically. The lack of details on the composition of the migration measures being evaluated limits our ability to conclude something about the impact for irregular migration programming in particular.

We did not find any study using quasi-experimental designs (QEDs) to evaluate the impact of skills-based ALMPs on migration outcomes, however the broader ALMP literature covering other contexts and outcomes have been widely evaluated using QEDs (Heckman, Ichimura, and Todd 1997; Heckman, LaLonde, and Smith 1999). In instances where randomization is not possible, QEDs may be less costly and less intrusive (Smith & Todd, 2005), and may potentially facilitate retrospective evaluations of skills-based ALMPs.

Further, half of the impact evaluations in this review included a qualitative component; however, we would have expected a larger proportion of studies using mixed methods given that the evidence base is very recent, with all the included studies published in the last decade during a period in which the evaluation community has put a lot of emphasis into mixed-methods approaches (Bamberger 2012; Vaessen, Lemire, and Befani 2020) to complement the results of what works with how, when and for whom.

Migration is a complex outcome, and migration behavior is driven by many factors and root causes. It is ambitious to expect observable impacts, particularly in the short term, and especially if we consider the challenges of targeting potential migrants. Heckman and colleagues (1999) suggest that one of the reasons skills-based ALMPs have been intensively evaluated is because short-run measures of outcomes such as earnings and employment are easily obtained. However, this is not necessarily the case for migration-related outcomes. This may result in interventions being delivered to people who were unlikely to migrate anyway, which translates in having no room for change. On the other hand, systematic reviews on the effectiveness of ALMPs on employment have shown that in the short-run these type of interventions are not effective, and rather changes are observed two to three years after completion of the program (Card, Kluve, and Weber 2015). On average, the

amount of time between the end of the intervention and the outcome measurement in studies included in this review was less than a year. While measuring long term impacts poses unique challenges, the short follow-up periods may have prevented effects from being identified.

Finally, attrition rates ranged from as low as 7.65% in Das (2017), up to 30% in Cho and colleagues (2013), and 45% in Müller and colleagues (2019). Attrition is an issue for the identification of impacts because it may be related to program drop-out rates (i.e., selection into the program), but also because attritors may differ from non-attritors. Attrition in Müller and Colleagues (2019) was not associated with treatment, but for Cho and colleagues (2013) there was some evidence of differential attrition between female beneficiaries and non-beneficiaries, where the former were less likely to attrite though the difference was not statistically significant at conventional levels. When the likelihood of attrition is different between participants in the treatment and control group, their outcomes are no longer comparable as a valid measure of the effect of the program. Finally, for migration outcomes, attrition might, in part, be capturing migration behavior, which would result in an underestimation of the average effect.

5.1.2 How do our results compare to other reviews?

There are multiple efforts to synthesize the literature examining the impacts of training programs on employment and earnings (Heckman, LaLonde, and Smith 1999; Blattman and Ralston 2015; Card, Kluve, and Weber 2015; McKenzie 2017), yet we are not aware of any efforts to synthesize the evidence on migration-related outcomes. In line with our statistically non-significant findings, the reviews on ALMPs' effectiveness have found little to no effect on employment or poverty alleviation outcomes.

5.1.3 Results of cost evidence

We found limited cost evidence reported in the included impact evaluations. Two studies reported some information related to cost (cost of the training and amount of the beneficiary stipend). Unfortunately, information such as the amount of the stipend paid to participants is not particularly useful for determining whether an intervention is cost effective. Only one of the 10 included impact evaluations reported on cost effectiveness, which is below average given that approximately 20 per cent of impact evaluations in the development field are including formal cost analyses in their evaluation reports (Brown and Tanner 2019). Understanding the effectiveness of interventions is insufficient for evidence-informed decision making. Decision makers must also understand how much these effective interventions cost, allowing programs to realize the best possible outcomes within finite resources. We direct interested readers to Glandon and colleagues (2023) for a review of resources for researchers or program designers interested in incorporating cost-effectiveness into their evaluations.

Table 7: Synthesis of the quantitative and qualitative analysis answering the research questions of this review

	Research question	Answers based on quant and qual analysis of the evidence base
1	How effective are skills-based active labor market interventions in improving intermediate outcomes that influence eventual migration (intention to migrate, and knowledge, perceptions, attitudes and expectations) in low and middle-income countries (L&MICs)?	**Intention to migrate** (n = 3) No evidence of a significant effect The observed outcomes ranged from -0.11 to 0.04 standard deviations, only one was statistically different from zero and suggested a reduction in intention to migrate associated with the program. **Knowledge, perceptions, attitudes and expectations** (n = 2) Not enough studies for a quantitative synthesis Individual outcomes were not statistically different from zero.
2	How effective are skills-based active labor market interventions in improving migration behavior (attempted migration, any migration, international migration, migration flow, migration stock) in L&MICs?	**Attempted migration** (n = 4) No evidence of a significant effect The observed outcomes ranged from -0.02 to 0.15 standard deviations, only one was statistically different from zero and suggested an increase in attempted migration associated with the program. **Migration behavior** (n = 6) No evidence of a significant effect The observed outcomes ranged from -0.12 to 0.12 standard deviations, only one was statistically different from zero and suggested an increase in migration behavior associated with the program. Results did not differ by gender.
3	Are there any unintended consequences of such interventions?	Authors of impact evaluations and linked programmatic studies did not report any unintended consequences.
4	Do effects vary by context, intervention type, design or population characteristics (e.g., age, sex, socio-economic status, etc.)?	The only moderators that helped explain differences in effects were the duration of the interventions and the average age of participants, which increased and reduced the effect of training on attempting migration respectively. Other characteristics of the context, the programs, and the population were not statistically significant predictors of the differences in the effects of the included studies.
5	What are contextual barriers to and facilitators of intervention effectiveness?	The qualitative evidence base shed some light on potential program design components, contextual factors, and population characteristics that may hinder or encourage enrollment and uptake of the training interventions included in this systematic review. Program uptake is one of the underlying assumptions in our theory of change. With limited uptake, the causal path between the intervention and outcomes will be broken. Program design and implementation elements that can encourage uptake include outreach and sensitizing target populations in the early stages of the program, providing certificates that serve as credentials in a timely manner, ensuring programs are appropriately designed (location, language, program duration and intensity, accessible and safe) and planned for, and providing sufficient incentives.

	Research question	Answers based on quant and qual analysis of the evidence base
		Context and population characteristics such as social norms around gender and equity, socio-economic and demographic characteristics of participants, and their attitudes and interest, may influence uptake as well and should be considered.
6	What is the cost-effectiveness of these interventions?	No evidence was found on the cost-effectiveness of ALMPs on migration outcomes.
7	How can future research enrich the evidence on the effects of active labor market interventions designed to improve migration in L&MICs?	Impact evaluations are needed on the effects of ALMPs targeting potential migrants, that use international migration measures in particular and creative methods such as mobile apps to measure irregular migration behavior, or that inquire about intentions to migrate through irregular channels.

5.2 Implications for future research, policy, and practice

Conducting rigorous research on migration programming can be challenging for a variety of reasons. While some evidence exists on the effects of skills-based ALMPs on migration-related outcomes, the current evidence base is insufficient to understand whether, how, or why these programs are working. Researchers, policymakers and practitioners have an important role to play in addressing these evidence gaps by identifying opportunities for evaluation in current and future investments. We suggest that when designing or commissioning skills-based ALMP programs, decisionmakers should consider:

Implications

Research

Fill geographic gaps

Commissioning more research across a range of geographies will help fill gaps in the current evidence base.

Measure long term impacts

Measuring longer term program impacts to help to determine whether these programs have long term effects for beneficiaries.

Utilize technology

Utilizing technology to assist with data collection and mitigate response issues, for example mobile apps with geolocation.

Consider a mixed-methods approach

Adopting a mixed-methods approach in which both quantitative (experimental or non-experimental) and qualitative data collection can contribute to our understanding of the mechanisms through which interventions achieve success.

Policy

Establish acceptability and confirm TOC assumptions

Collecting primary data to establish acceptability and appropriateness of interventions and confirm that assumptions in any program theories of change reflect the concerns and needs of program participants and are context appropriate. Migration decision-making is often driven by a combination of factors in addition to economic insecurity. A program that fails to account for all primary drivers may lead to low uptake and diminish the rigor of impact evaluations.

Include IEs in future programming

Identifying opportunities to include impact evaluations in future programming. There is a dearth of impact evaluations on programs explicitly attempting to address migration and target populations considering irregular migration as a means for a better life. However, this requires intentionality and commitment. For example, identify activities or phases of a program that could facilitate random assignment and testing of specific program components for impact, or use eligibility criteria for quasi-experimental approaches (e.g., the pilot or initial phases of rolling out a large program may present a natural opportunity to use implementation scaling delays for randomization).

Include cost data

Including cost data in order to provide useful information for cost-effectiveness analyses, which can help funders make the most out of limited resources.

Champion better reporting practices

Producing reports that contain the information needed to make meaningful contributions to our knowledge base. Impact evaluation studies should follow rigorous methodological designs that allow for causal attribution and report their research in a transparent and complete manner which would allow for a full assessment of study quality. Good impact evaluations will register and report a pre-analysis plan which is publicly accessible, establish an appropriate counterfactual (i.e., through random assignment or through an appropriate quasi-experimental design) and collect baseline data from all participants prior to the start of the program. They will collect outcome data from all participants (both treatment and control) at equivalent intervals. In reporting their findings authors will appropriately account for multiple hypothesis testing in their analysis (e.g., through use of a Bonferroni correction). Authors should avoid *p*-hacking by reporting on all outcomes, regardless of statistical significance. Finally, when reporting on these evaluations, authors should also report information related to the assignment mechanism, baseline balance among treatment and control groups, attrition and differential attrition, deviations from the intended intervention, potential spill-over effects, and potential for performance bias.

Ensure sensitivity to vulnerable groups

Ensuring that studies are sensitive to the needs of vulnerable groups by reporting clearly all ethical approvals received from relevant review boards. Studies may also consider weaving gender and equity sensitive approaches throughout the project lifecycle (e.g., project design, data collection, etc.) to account for social norms which might reinforce inequities, and taking care that data is collected in way which allows vulnerable groups, particularly women, to feel comfortable speaking.

Ensure adequate resources for evaluation

If evidence generation is appropriate and a priority for the investment, ensure there are adequate resources for evaluation. This means that evaluated programs are of adequate scale so that impact evaluations are sufficiently powered for minimum detectable effect sizes. Relatedly, ensure there is adequate funding for impact evaluations within the larger investments.

Closely collaborate with researchers

Closely collaborating with researchers to help ensure the success of an impact evaluation. This includes coordination of data collection prior to implementation if a baseline is needed, resources and planning for follow up measures to reduce non-response or attrition, and advocating for explicit testing of irregular migration outcomes while reducing social desirability bias.

Migration decision making is complex - plan accordingly

Anticipating that migration decision-making is complex and that successful programs may not change an individual's desire to migrate, but may actually increase their capacity to do so. If there are few legal channels to do so, individuals may choose irregular pathways. Ensure that programs do not unintentionally exacerbate these dynamics by anticipating and planning during program design.

Consider how impacts can be sustained

Consider how potential investment impacts can be sustained after the life of the program. Successfully building skills and strengthening local workforce capability may address short-term "demand-side" issues but will fail to achieve longer-term impact on irregular migration if solutions for "supply-side" economic root causes (e.g., lack of local labor market and economic opportunities) are left unaddressed.

5.3 Limitations of this review

The small number of included studies limited our ability to draw strong conclusions about intervention effectiveness. Relatedly, in all cases moderator analyses are considered exploratory as the small number of studies limited our ability to fully explore potential sources of variation in the outcomes. We were also unable to test for publication bias in the literature as these tests require a minimum of 10 studies to be valid.

The qualitative evidence we included in this synthesis was limited to programs from included impact evaluations, using the effectiveness plus approach (Snilstveit 2012). This approach is particularly limiting when interventions are small in scale and program information is not publicly available. Future studies may want to consider qualitative research beyond just included programs. In addition, we did not complete critical appraisals of qualitative studies.

Our risk of bias assessments are based solely on what is reported in the included studies, and any associated pre-analysis plans. This means that some domains may have been assessed as “some concerns” when information was not reported.

6. Acknowledgements

6.1 Contributions of authors

María Daniela Anda León (DA), Shannon Shisler (SS), and Carolyn Huang (CH) are the core team for this review. DA is a research associate with experience in development economics. SS is a systematic review and quantitative methods expert, with over a decade of experience in designing, managing and analyzing quantitative research, including meta-analyses. CH is an expert in causal research who leads and implements impact evaluation and evidence review program.

- Content: DA, SS, Promise Nduku (PN) and Andile Madonsela (AM) developed the content of the review with inputs and quality assurance from CH.
- Systematic review methods: DA, PN and AM drafted the review methods with inputs and support from SS (Senior Evaluation Specialist).
- Qualitative evidence synthesis: PN and AM conducted the search for qualitative evidence, screening, and thematic synthesis.
- Statistical analysis: SS oversaw the statistical analysis by DA who extracted effects data in duplicate with a consultant, with quality assurance from SS.
- Quality appraisal: DA assessed the risk of bias of included impact evaluations in duplicate with a consultant, with quality assurance from SS.
- Information retrieval: Miriam Berretta (MB) developed the initial search strings with quality assurance by information specialists. MB and DA led screening, grey literature review, snowballing and references checks with the support of a pool of consultants.

6.2 Declarations of interest

There are no conflicts of interest to report on this review.

6.3 Plans for updating this review

There are no plans for an update to this review to date.

Appendix A: Search strategy

The search for the impact evaluations included in this systematic review was implemented as part of the Evidence Gap Map on addressing root causes and drivers of irregular migrations (Berretta et al., forthcoming). This appendix summarizes that search strategy for the EGM. More details can be found in the EGM report along with an example of a search string.

Authors of the EGM followed two different approaches depending on whether the intervention domain had been explored recently by other evidence mapping efforts or not. For the former, the authors leveraged pre-existing search strategies, while for the later they devised a search strategy comprising key words and Boolean operators.

Updated searches

For the domain on strengthening resilience against shocks and stressors, numerous categories were taken from the Mapping evidence of what works to strengthen resilience to shocks and stressors (Berretta et al., 2022). To update the search, the following databases were used:

- CAB Abstracts (EBSCO)
- CAB Global Health (OVID)
- Africa-Wide (EBSCO)
- Academic Search Complete (EBSCO)
- APA PsycInfo (OVID)
- Web of Science (SSCI)
- Econlit (EBSCO)
- Social Science Research Network (SSRN)
- World Bank (EBSCO Discovery)
- Agris (EBSCO Discovery)
- RePEc (EBSCO Discovery)
- Campbell library

For the domain on violence prevention, numerous categories were taken from The effects of rule of law interventions on justice outcomes: an evidence gap map (Sonnenfeld et al. 2023). To update the search, the following databases were used:

- Scopus
- Social Science Citations Index
- International Political Science Abstracts
- Communication & Mass Media Complete
- Research Papers in Economics (RePEc)

New search strategies

Two domains in the EGM by Berretta and colleagues (Forthcoming) had not been covered by previous EGMs: Economic opportunities and Orderly and safe migration management. Given the nature of interventions within those domains, reported changes in outcomes are expected to occur in a number of development sectors. As such, the strategy considered sector specific databases where appropriate. The following databases were searched using:

- Scopus
- Social Science Citations Index

- International Political Science Abstracts
- Research Papers in Economics (RePEc)
- CAB Abstracts
- Africa-Wide
- Academic Search Complete
- Web of Science
- Econlit
- Social Science Research Network (SSRN)
- World Bank
- Campbell library

Grey literature searches

Berretta and Colleagues (Forthcoming) searched for grey literature on the websites of 102 organizations. These organizations were selected on the basis of their action and work in migration related matters such as the International Organization of Migration (IOM), the Center for Migrant Studies, the Global Forum on Migration and Development, and IZA World of Labor, among others. Other website from referential international development and research organizations were also searched including Abdul Latif Jameel Poverty Action Lab (J-Pal), the United Nations Evaluation Group, the United States – Development Experience Clearing House, the AEA RCT Registry protocols, and others. A complete list of organizations and websites are presented in the appendix of the EGM report.

Other searches

Berretta and Colleagues (Forthcoming) also implemented forward and backward citation tracking of included papers. The authors used the software Publish and Perish and Citation Tracer to facilitate this search. In backward citation tracking, they reviewed eligible studies from the bibliographies of included studies. Finally, a public call for relevant papers was published via blog.

Appendix B: Data extraction tools

1. Quantitative data extraction tool for effect size calculation

VARIABLE LABEL	EXPLANATION
Study ID	This is the study ID - it should match the study ID from the Outcome Mapping Sheet (e.g., 946578)
Estimate ID	The estimate ID will provide a specific number for each effect size extracted and should include the original study number, underscore, then the unique ID number (e.g., 946578_1, 946578_2 and so on)
STUDY DESCRIPTIVE INFORMATION	
Author	Author last name For 1 author: leading author last name (e.g. Gomez) For 2 authors: both author last names with ampersand in between (e.g. Smith & Bahn) For 3 or more authors: leading author last name followed by et al. (e.g. Gupta et al.)
Year	Year published
Location	Country of intervention. If it is an intervention for an specific location within a country, like city, write down the city instead.
Design	0=Experimental Design , 1=Quasi-Experimental Design
How Counterfactual is Chosen	Free text (e.g., RCT, Cluster RCT, propensity score matching, Instrumental variable, Fixed effects, etc.) - Multiple codes are ok
Estimate Type	Type of data for this effect size: 1 = Continuous - means and SDs, 2 = Continuous - mean difference and SD, 3 = Dichotomous outcome - proportions, 4 = Regression data - dichotomous outcome, 5 = Regression data - continuous outcome
Population	Drop down menu
Subgroup	Is this analysis of a subgroup? 0=no, 1=yes
If yes to subgroup, describe	Free text, describe the subgroup if applicable (e.g., boys, girls). If no subgroup, type N/A
Source	Note the page number, table number, column, and row you used to extract the data
Intervention description	Provide detailed description of the intervention such that a reader could easily understand what happened. Avoid copying text directly from the article as it is likely to be verbose. Summarize in your own words but include page numbers for quick reference. If more than two or more interventions are being evaluated, please provide descriptions for each intervention arm under separate rows, e.g. description of cash transfer (in all rows where estimate id's evaluate the cash transfer), description of cash transfer + community mobilization (in all rows where estimate id's evaluate the multicomponent intervention).

VARIABLE LABEL	EXPLANATION
Intervention code	Dropdown menu with intervention codes
Exposure to intervention (in months)	How long is the intervention exposure itself? If time series is used, indicate the length of the period covering data points when the intervention was going on.
Evaluation period (in months)	The total number of months elapsed between the end of an intervention and the point at which an outcome measure is taken post intervention, or as a follow-up measurement. If less than one month, use decimals (e.g., measurement immediately after the intervention end would be coded as 0, one week would be .25, etc.)
Post-intervention or change from baseline?	0 = Post-intervention, 1 = Change from baseline
OUTCOMES	
Outcome description	Record the outcome for the corresponding effect size. Use this open answer field to enter, in the author's own words, a description of the outcome. Be selective and concise with the excerpts being transcribed here to ensure accurate and precise descriptions of the outcome. To the extent possible, be sure to include numbers, units, population, and comparators. Include page numbers with every excerpt extracted.
Outcome codes	Dropdown menu with outcome codes
Dataset	Record if data for this outcome comes from an identified dataset
EFFECT SIZE DATA EXTRACTION	
Reverse Sign (i.e., decrease is good)	Record no if an increase is good, record yes if a decrease is good and the sign needs to be reversed.
Unit of analysis	What is the unit of analysis? UOA for this effect size: 1= Individual, 2= Household, 3= Group (e.g. community organization), 4= Village, 5 = Other, 6 = Not clear
mean_t	Outcome mean for the treatment group
sd_t	Outcome standard deviation for treatment group
mean_c	Outcome mean for the comparison group
sd_c	Outcome standard deviation for control group
mean_overall_diff	Overall mean difference (treatment - control)
diff_se	Standard error of the overall mean difference
diff _t	t-statistic of mean difference
diff_p-value	p-value of mean difference
Odds ratio	Odds ratio reported in the study

VARIABLE LABEL	EXPLANATION
OR_se	Odds ratio standard error reported in the study
Risk ratio	Risk ratio reported in study
RR_se	Risk ratio standard error
reg_coeff	Report the regression coefficient of the treatment effect
reg_SE	Report the associated standard error of the regression coefficient.
reg_t	Report the associated t statistic of the effect size (coefficient/SE)
reg_CI_LB	Report the associated Lower bound of the 95% Confidence interval of the effect size. If CI is reported for a different confidence level, indicate that in the notes section.
reg_CI_UP	Report the associated Upper bound of the 95% Confidence interval of the effect size. If CI is reported for a different confidence level, indicate that in the notes section.
Exact p value	Exact p value if given, if not, record as written in the manuscript (e.g., p < .001, or p > .05)
clust_t	Number of clusters - treatment group
clust_c	Number of clusters - control group
clust_T	Number of clusters - total sample
n_t	Sample size - treatment group
n_c	Sample size - control group
n_T	Sample size - total sample
periods (1 if cross sectional)	Record how many periods of evaluation there are (e.g., cross section is 1, panel data with 3 measurements is 3)
Does the sample size need to be corrected?	Often in panel data, models will report the number of observations rather than number of participants. In this column you will indicate 1="Yes" if the sample size needs to be divided by the number of periods, and 0="No" if either it is cross-sectional data, or if the authors have already divided the number of observations by the number of panel assessments and thus no correction is necessary.
Treatment Variable	Record the treatment variable as written in the model (e.g., the variable name the author uses, such as ("Intervention x Time")
CODING RECORDS	
coder	Record your name
Notes	Record any notes important for the team

2. Qualitative data extraction tool

Analytical factor	Description
Context	Any variable related to external factors beyond the program's control that affect the impact. This can refer to political factors such as types of governance, societal factors such as norms, economic factors such as a recession, and cultural factors such as beliefs.
Intervention design	Any variable that is related to the design and planning of the applied intervention. Design and planning of an intervention refer to the blueprint or schedule of the intervention and will typically outline what components the intervention consists of and in what sequence they will be applied. Examples of design variables refer to: outreach strategy, posters; reminders; type of training.
Intervention implementation	Any variable that is related to the implementation of the intervention in practice. This refers to variables that emerge while the intervention is applied and are usually not known in advance. Examples of implementation variables refer to the lack of attendance or uptake, payment difficulties, corruption, elite capture.
Population characteristics	Any variable related to the population targeted by the intervention or the population in which the effects are measured (in cases where these differ). This can refer to the socio-economic status of the population, its educational status, and asset ownership

Appendix C: Critical appraisal tool

1. Full Appraisal of Risk of Bias for Impact Evaluations using RCT Designs

The following table provides a provisional tool to guide the risk of bias assessment for quantitative impact evaluations. If necessary, we could amend the tool to better inform the appraisal of primary studies.

Provisional risk of bias assessment tool (RCT)

Code	Question	Coding format	Criteria	Decision rule
General	ID	EPPI ID		
General	Study first author	Open answer		
General	Time taken to complete assessment	Minutes		
General	Design type: What type of study design is used?	1= Randomized controlled trial (RCT) (random assignment to households/individuals) or quasi-RCT 2= Cluster-RCT (quasi-RCT)	-	
General	Methods used for analysis: Which methods are used to control for selection bias and confounding?	1 = Statistical matching (PSM, CEM, covariate matching) 2 = Difference in differences (DID) estimation methods 3 = IV-regression (2stage least squares or bivariate probit) 4 = Heckman selection model 5 = Fixed effects regression 6 = Covariate adjusted estimation 7 = Propensity weighted regression 8 = Comparison of means = Other (please state)	-	
General	Design and analysis method description	Open answer	Briefly describe the study design and analysis method undertaken by the authors.	

Code	Question	Coding format	Criteria	Decision rule
General	Study population	Open answer	Provide any details in the paper that describe how the study population was selected, covering: a) How is the population selected? What is the sampling strategy to recruit participants from that population into the study? b) What are the characteristics of that study participants? Was this a pilot program aimed at being scaled up? d) Were there specific factors of success or failure in the implementation?	
General	Type of comparison group	1=No intervention (service delivery as usual) 2=Other intervention 3=Pipeline (wait-list) control (still service delivery as usual)	Indicate type of comparison group	
General	Type of comparison group (if other)	Open answer		
General	Ethical clearance	Open answer	Provide any details of ethical research clearances granted. Report unclear if this information is not available.	
General	Study registration	Open answer	Provide any details of study registration, including registry IDs, etc.	
1: Assignment mechanism - Assessment	Assignment mechanism: Was the allocation or identification mechanism random or as good as random?	1= Yes, 2 = Probably Yes, 3 = Probably No, 4 = No, 8 = Unclear	a) The authors describe a random component in sequence generation/ randomization method (e.g. lottery, coin toss, random number generator) and assignment is performed for all units at the start of the study centrally or using a method concealed from participants and intervention delivery	Score “Yes” if all criterion a), b), c) and d) are satisfied. Score "Probably Yes" if only criterion a) and b) are not satisfied OR if only criteria c) is not satisfied.

Code	Question	Coding format	Criteria	Decision rule
			b) If public lottery is used for the sequence generation, authors provide detail on the exact settings and participants attending the lottery. c) If a special randomization procedure is used to ensure balance, it is well described and justified given the study setting (stratification, pairwise matching, unique random draw, multiple random draws etch). d) A balance table is reported suggesting that allocation was random between all groups including subgroup receiving different treatment within control or treatment groups (if the comparison is relevant for this assessment).	Score “Unclear” if d) is not satisfied because no balance table is reported. Score "Probably No" if d) is not satisfied because there is no balance table reported and there is evidence suggesting a problem in the randomization, such as baseline coefficients in a diff-in-diff regression table are very different or sample size is too small for the procedure used (using stratification when there are less than two units for each intervention and control group in each strata can lead to imbalance). Score “No” if d) is not satisfied because there are large imbalances concerning a large number of variables, providing evidence that the assignment was not random. If this is scored as no, use the NRS tool.
1: Assignment mechanism - Justification	Assignment justification	Open answer	Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).	

Code	Question	Coding format	Criteria	Decision rule
2: Unit of analysis - Assessment	Unit of analysis: Is unit of analysis in cluster allocation addressed in standard error calculation?	1=Yes 2=No 3=Not reported/unclear 4=Not applicable	Score "Yes" if UoA = UoR OR if UoA ≠ UoR and standard errors are clustered at the UoR level OR data is collapsed to the UoR level Score "Not reported/unclear" if not enough information is provided on the way the standard errors were calculated or what the unit of analysis is. Score "Not applicable" if it is not a cluster RCT. Score "No" otherwise.	
3: Selection bias - Assessment	Selection bias Was any differential selection into or out of the study (attrition bias) adequately resolved?	1= Yes, 2 = Probably Yes, 3 = Probably No, 4 = No, 8 = Unclear	Score "Yes" if there is no attrition or attrition falls into the green zone and the study establishes that attrition is randomly distributed (e.g. by presenting balance by key characteristics across groups) AND if survey respondents were randomly sampled. Score "Probably yes" if attrition falls into the green zone AND if survey respondents were randomly sampled. Score "Unclear" if there is an attrition problem but no information provided on the relationship between attrition and treatment status, OR if there is not enough information on how the population surveyed was sampled. Score "Probably no" if there is attrition which is likely to be related to the intervention OR there is some indication that the survey	

Code	Question	Coding format	Criteria	Decision rule
			respondents were purposely sampled in a way that might have led the sampling to be different between treatment and control groups, or attrition falls into the yellow zone. Score "No" if attrition falls into the red zone.	
3: Selection bias - Justification	Selection bias justification	Open answer	Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).	
4: Confounding - Assessment	Confounding and group equivalence: Was the method of analysis executed adequately to ensure comparability of groups throughout the study and prevent confounding	1= Yes, 2 = Probably Yes, 3 = Probably No, 4 = No, 8 = Unclear	a) Baseline characteristics are similar in magnitude; b) Unbalanced covariates at the individual and cluster level are controlled in adjusted analysis; c) Adjustments to the randomization were taken into account in the analysis (stratum fixed effects, pairwise matching variables)? (Bruhn and McKenzie 2009)	Score “Yes” if criterion a) and b) are satisfied; Score "Probably yes" if a) is not satisfied but b) is satisfied and imbalances are small in magnitude OR if only a) is satisfied. Score “Unclear” if no balance table is provided or if imbalances are controlled for but they are very large in magnitude and assignment mechanism is not coded as "Yes" or "Probably yes" Score "Probably no" if a) and b) are not satisfied and the magnitude of imbalances are small.

Code	Question	Coding format	Criteria	Decision rule
				Score “No” if a) and b) are not satisfied and the magnitude of imbalances are large, and covariates are clear determinant of the outcomes.
4: Confounding - Justification	Confounding justification	Open answer	Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).	
5: Deviations from intended interventions - Assessment	Deviations from intended interventions: Spillovers, crossovers and contamination: was the study adequately protected against spill-overs, crossovers and contamination?	1= Yes, 2 = Probably Yes, 3 = Probably No, 4 = No, 8 = Unclear	a) There were no implementation issues that might have led the control participants to receive the treatment (implementer's mistake). b) The intervention is unlikely to spillover to comparisons (e.g., participants and non-participants are geographically and/or socially separated from one another and general equilibrium effects are not likely) or the potential effects of spill overs were measured (e.g. variation in the % of unit within a cluster receiving the treatment). There is no risk of contamination by external programs: the treatment and comparisons are isolated from other interventions which might explain changes in outcomes. d) There is nothing in the surveys that might have given the control participants an idea of what the other group might receive OR they did but there is no risk that this has changed their behaviors; AND the survey process did not reveal information to the control group that they did not have before (e.g. the study	Score “Yes” if criterion a), b), c) and d) are satisfied. Score "Probably yes" if there is no obvious problem but there is no information reported on potential risks related to spill overs, contamination, or survey effects in the control group OR if there were issues with spillovers but they were controlled for or measured. Score “Unclear” if spillovers, crossovers, survey effects and/or contamination are not addressed clearly. Score "Probably no" if any of the criterion a), b), c) or d) are not satisfied but the scale of the issue is not clear.

Code	Question	Coding format	Criteria	Decision rule
			aims to measure increase in take up of a service or product that participants might not know about) Authors might put something in place in the design of the study that allows to control for that survey effect (e.g. a pure control with no monitoring except baseline end line)	Score “No” if any of the criterion a), b), c) or d) are not satisfied and happened at a large scale in the study.
5: Deviations from intended interventions - Justification	Deviations justification	Open answer	Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages). For example, intervention groups are geographically separated, authors use intention to treat estimation or instrumental variables to account for non-adherence, and survey questions are not likely to expose individuals in the control group to information about desirable behaviors (‘survey effects’).	
6. Performance bias - Assessment	Performance bias: Was the process of monitoring individuals unlikely to introduce motivation bias among participants?	1= Yes, 2 = Probably Yes, 3 = Probably No, 4 = No, 8 = Unclear	a) The authors state explicitly that the process of monitoring the intervention and outcome measurement is blinded and conducted in the same frequency for treatment and control groups, or argue convincingly why it is not likely that being monitored could affect the performance of participants in treatment and comparison groups in different ways (such as resulting in Hawthorne or John Henry effects). b) The outcome is based on data collected in the context of a survey, and not associated with a particular intervention trial, or data are	Score “Yes” if either criterion a) or b) are satisfied. Score "Probably yes" if the study is based on data collected during a trial and there is no obvious issue with the monitoring processes, but authors do not mention potential risks. Score “Unclear” if it is not clear whether the authors use an appropriate method to prevent

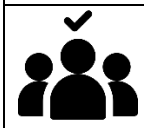
Code	Question	Coding format	Criteria	Decision rule
			collected from administrative records or in the context of a retrospective (ex post) evaluation.	Hawthorne and John Henry Effects (e.g., blinding of outcomes and, or enumerators, other methods to ensure consistent monitoring across groups). Hawthorne effects may result where participants know that they are being observed and John Henry Effects may result from participant knowledge of being compared. Score "Probably no" if there was imbalance in the frequency of monitoring in intervention groups, which might have influenced participants' behaviors. Score "No" if neither criterion a) or b) are satisfied.
6. Performance bias - Justification	Performance bias justification	Open answer	Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).	
7. Outcome measurement bias - Assessment	Outcome measurement bias: Was the study free from biases in outcome measurement?	1= Yes, 2 = Probably Yes, 3 = Probably No, 4 = No, 8 = Unclear	a) Outcome assessors are blinded, or the outcome measures are not likely to be biased by their judgement. b) For self-reported outcomes: respondents in the intervention group are not more likely to have accurate answers due to	Score “Yes” if criterion a), b), c) and d) are satisfied: Score "Probably yes" if there is a small risk related to any of a), b), c) or d) and there is no more information provided to justify

Code	Question	Coding format	Criteria	Decision rule
			recall bias. c) For self-reported outcomes: respondents do not have incentives to over/under report something related to their performance or actions, OR researchers put in place mechanisms to reduce the risk of reporting bias (researchers not strongly involved in the implementation of the program and it is clear that their answers to the survey will not affect what they receive in the future) OR authors have measured the risks of bias through falsification tests or measuring the effect on placebo outcomes in cases where there was a risk of reporting bias. d) Timing issue: the data collection period did not differ between intervention and comparison group; the baseline data is not likely to be affected by the beginning of the intervention or affects a small percentage of the study participants.	the absence of bias OR if there was a high risk of bias, but authors have either controlled it in their design or measured it with a placebo outcome. Score “Unclear” if there is a high risk related to any of a), b), c) or d) and there is no more information provided to justify the absence of bias. Score "Probably no" if there are high risk related to a), b), c) or d) and it is clear that authors were not able to control this bias. Score “No” if there is evidence of bias.
7. Outcome measurement bias - Justification	Outcome measurement justification	Open answer	Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).	
8. Reporting bias - Assessment	Analysis reporting: Was the study free from selective analysis reporting?	1= Yes, 2 = Probably Yes, 3 = Probably No, 4 = No, 8 = Unclear	a) A pre-analysis plan or trial protocol is published and referred to or the trial was preregistered, or the outcomes were preregistered. b) Authors report results corresponding to the outcomes announced in the method section (there is no outcome reporting bias);	Score "Yes" if all the criterion a), b), c), d), and e) are satisfied; Score "Probably yes" if all the conditions are met except a), or if all the conditions are met but there is some element missing that could

Code	Question	Coding format	Criteria	Decision rule
			c) Authors report results of unadjusted analysis and intention to treat (ITT) estimation, alongside any adjusted and treatment-on-the treated/complier average-causal effects analysis.) d) Authors use the appropriate analysis method (use baseline data when available), and different treatment arms are differentiated in the analysis e) Authors have reported all the analysis which could help understand the results and no other bias is assessed as unclear due to the lack of an important analysis (e.g., a balance table or a subgroup analysis)	have helped understand the results better (e). Score "Unclear" if there is not enough information to determine that there is an analysis missing; Score "Probably no" if any of the criterion b), c) or d) are not satisfied; Score "No" if any of the criterion b), c) or d) are not satisfied and there is evidence that the analysis results would be different because large imbalances were not controlled for, compliance was very low and ITT estimation was not reported or different treatment arms were pooled.
8. Reporting bias - Justification	Analysis reporting justification	Open answer	Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).	
9. Other bias - Assessment	Other risks of bias Is the study free from other sources of bias?	1= Yes, 4 = No		
9. Other bias - Justification	Other bias justification	Open answer	Justification for coding decision	
10. Blinding - observers - Assessment	Blinding of participants?	1=Yes 2=No 8=unclear 9= N/A	If there is no information, code NO. If there is information but it is ambiguous, code UNCLEAR.	

Code	Question	Coding format	Criteria	Decision rule
10. Blinding - observers - Assessment	Blinding of outcome assessors?	1=Yes 2=No 8=unclear 9= N/A	If there is no information, code NO. If there is information but it is ambiguous, code UNCLEAR.	
10. Blinding - analysts - Assessment	Blinding of data analysts?	1=Yes 2=No 8=unclear 9= N/A	If there is no information, code NO. If there is information but it is ambiguous, code UNCLEAR.	
10. Blinding - method(s)	Method(s) used to blind	Open answer (including describe method of placebo control) No 9= N/A	Describe method(s) used to blind	
11. External validity - Assessment	External validity	Open answer	a) What do authors say about external validity?	Include all information that can help assess the external validity of the results.
			Summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).	

Appendix D: Risk of bias criteria definitions

	Assignment mechanism In randomized controlled trials, the random component in sequence generation or randomization method needs to be clearly reported along with proof of effectiveness of this process in balancing treatment and control groups. This would ensure that groups are comparable. However, if the authors fail to report details on the randomization process, or if the random allocation of the intervention is not performed upon all participants or groups of participants in the evaluation, we would not be able to trust the results of the study as valid measures of the impact of the intervention being assessed.
	**Selection** Selection bias is the difference between an impact estimate and the real effect of a program that occurs when participants or groups of participants differ systematically from the comparison group or population of interest. This difference leads to a systematic error in the association between the program and the observed changes in the outcome. A possible source of selection bias is differential attrition, that is when missing outcome data is related to a factor or group of factors associated with the intervention assignment. Possible sources of missing outcome data are withdrawal of participants from the study, participants refusing to provide information, participants not being able to experience the outcome, for example diseased participants.
	Confounding Confounding occurs when a factor, be it a contextual factor or characteristic of the participants or group of participants in an evaluation, is independently associated with the exposure to the intervention and the outcome of interest. When such a factor is not controlled for in the analysis, the association between the intervention and the observed change in outcome is distorted.
	Implementation Fidelity When there are deviations from intended interventions the implementation fidelity is broken. Issues such as non-adherence of participants to treatment and control groups as intended in the protocol, suggest that the observed changes in the outcome can be associated with the reasons for noncompliance instead of the intervention.
	Performance Performance bias occurs when the participants of group of participants know the intervention allocation, know they are being observed or know they are being compared. For example, participants in the control group might seek to receive the intervention elsewhere or other interventions that address the same underlying issues. This bias can be reduced or avoided when participants, implementers and researchers are blinded to the experiment and/or to the assignment of participants or groups of participants.
	Outcome measurement Measurement errors refer to inaccurate measures of continuous outcomes or misclassification of categorial outcomes. When measurement errors differ systematically among participants or group of participants, the association between the program and the observed change in the outcome also suffers from systematic errors. This could happen, for example, when participants know the intervention assignment and the outcome is self-reported, or when the timing of data collection processes differs among participants or groups of participants.

Reporting

Reporting the results of the study based on impartial selection of outcomes, specifications, or samples, can occur through different ways:

- Omitting outcomes deemed to be unfavorable or statistically insignificant
- Adding outcomes based on collected data favoring statistical significance
- Including only a subset of the analyzed data in the published study
- Failing to report data that was analyzed in the trial (e.g., adverse effects)
- Changing outcomes of interest (from primary outcomes to secondary outcomes if they do not yield significant results)

References

Included impact evaluations

**Studies preceded by an asterisk were also used in the qualitative analysis.*

*Ahn, Roy, Ali Protik, Brian Kirchhoff, and Phoebe Bui. 2020. “Impact Evaluation USAID/Cambodia Countering Trafficking in Person (CTIP) Program: Final Report.” Washington, DC: USAID Democracy, Human Rights, and Governance Learning, Evaluation, and Research (DRG-LER) Activity. https://pdf.usaid.gov/pdf_docs/PA00X25V.pdf

*Bah, Tijan L., Catia Batista, Flore Gubert, and David McKenzie. 2022. Can information and alternatives to irregular migration reduce “backway” migration from The Gambia?. World Bank Group, Development Economics, Development Research Group. - Washington, DC: World Bank Group, 2022. - 65 S. Policy research working paper; 10146. https://openknowledge.worldbank.org/bitstream/handle/10986/37911/IDU044e4c11c01f53041f408aec06a 6adb79c705.pdf

Bandiera, Oriana, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali. 2021. The Search for Good Jobs: Evidence from a Six-year Field Experiment in Uganda. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3910330.

*Cho, Yoonyoung, Davie Kalomba, Ahmed Mushfiq Mobarak, Victor and Orozco-Olvera. 2013. Gender Differences in the Effects of Vocational Training: Constraints on Women and Drop-Out Behavior. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2266835.

Das, Narayan Chandra. 2017. Training the Disadvantaged Youth and Labor Market Outcomes: Evidence from Bangladesh, in Essays on Labor and Credit Markets in Bangladesh. UC Berkeley. UC Berkeley Electronic Theses and Dissertations. https://escholarship.org/uc/item/1v71f82c

*Duthie, Mike, Mateusz Pucilowski, Lisette Anzoategui, Braden Agpoon, and Erin Murphy-Graham. 2018a. A Ganar Alliance Impact Evaluation Endline Report Honduras. Social Impact, inc. under contract No. AID-OAA-M-11-00019 and produced for review by the United States Agency for International Development (USAID). https://pdf.usaid.gov/pdf_docs/PA00T78R.pdf

* Duthie, Mike, Mateusz Pucilowski, Lisette Anzoategui, Braden Agpoon, and Erin Murphy-Graham. 2018b. A Ganar Alliance Impact Evaluation Endline Report Guatemala. Social Impact, inc. under contract No. AID-OAA-M-11-00019 and produced for review by the United States Agency for International Development (USAID). https://pdf.usaid.gov/pdf_docs/PA00T78S.pdf

Hamory, Joan, Michael Kremer, Isaac Mbiti, and Edward Miguel. 2015. Evaluating the Impact of Vocational education voucherson out-of-school youth in Kenya. 3ie Impact Evaluation Report 37. New Dehli: International Initiative for Impact Evalution. https://www.3ieimpact.org/sites/default/files/2019-01/ie37-kenya-vocational-education.pdf

*Hardy, Morgan, Issac Mbiti, Jamie Lee McCasland, and Isabelle Salcher. 2019. The Apprenticeship-to-Work Transition Experimental Evidence from Ghana. World Bank Group. Social Protection and Jobs Global Practice. Policy Research Working Paper 8851. https://ssrn.com/abstract=3385859

Müller, Angelica, Utz Pape, and Laura Ralston. 2019. Broken Promises: Evaluating an Incomplete Cash Transfer Program. World Bank Group. Poverty and Equity Global Practice & Social Protection and Jobs Global Practice. Policy Research Working Paper 9016. https://doi.org/10.1596/1813-9450-9016

Additional qualitative studies

Ayagiba, Lawrence, Monalisa Lotsu, Kwabena SarfoSarkwa, and Theophilus Eghan. 2016. Performance audit report of the auditor-general on the governance of the national apprenticeship programme by the council for technical and vocational education and training. Ghana Audit Service. https://audit.gov.gh/files/audit_reports/Performance_Audit_Report_of_the_Auditor-General_on_the_Governance_of_the_National_Apprenticeship_Programme_by_the_Council_for_Technical_and_Vocational_Education_and_Training.pdf

Duthie, Mike, Mateusz Pucilowski, Lisette Anzoategui, Braden Agpoon, and Erin Murphy-Graham. 2018c. A Ganar alliance impact evaluation synthesis report Guatemala and Honduras. Social Impact. Washington, DC. https://pdf.usaid.gov/pdf_docs/PA00T78T.pdf

Pape, Utz. 2015. Republic of South Sudan Youth Startup Business Grant Program-Volume I: Program rationale, design and implementation. Report No: ACS14502. URL: http://documents. worldbank. org/curated/en/647811498587425293/South-Sudan-Youth-startup-business-grant-program-program-rationale-designand-implementation.

Pape, Utz, Laura Ralston, Angelika Mueller, Mollie Foust, Luca Parisotto, Nadia Selim, Jeremy Shapiro, James Walsh, and Nicola Pontara. 2018. Impact of program cancellation due to conflict in South Sudan: a chapter of the South Sudan poverty assessment 2017. Disclosure. https://documents1.worldbank.org/curated/en/861351529986243437/pdf/Impact-of-Program-Cancellation-Due-To-Conflict-in-South-Sudan.pdf

Mbiti, Isaac, Jamie McCasland, Morgan Hardy, Kym Cole, Mark Enriquez, and Isabelle Salcher. 2019. Training for success: targeting and incentives in apprenticeship training in Ghana. International Initiative for Impact Evaluation. 3ie Grantee Final Report. https://www.3ieimpact.org/sites/default/files/2019-11/GFR-TW1.1063-Ghana-apprenticeship.pdf

Murphy-Graham, Erin. 2021. Life Skills Education for Urban Youth in Honduras and Guatemala: A Capability Analysis of the Sports-Based Job Training Program A Ganar. In Life Skills Education for Youth: Critical Perspectives (pp. 147-167). Cham: Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-030-85214-6_7

Rahman, Rehnuma, Atiya Rahman, Zion Rabi Samadder, and Abdul Bayes. 2017. The effects of skill training on livelihoods: Evidence from BRAC's Intervention on School Dropout Adolescents (No. 2). Skills Development Working paper. https://bigd.bracu.ac.bd/publications/the-effects-of-skill-training-on-livelihoods-evidence-from-bracs-intervention-on-school-dropout-adolescents/

Other references

Anda Leon, Maria Daniela, Shannon Shisler, Carolyn Huang, Miriam Berretta, Promise Nduku, and Andile Madonsela. 2023. “Systematic Review of Skills-Based Active Labor Market Interventions on Intermediate and Final Migration Outcome: Protocol.” New Delhi: International Initiative for Impact Evaluation (3ie). https://3ieimpact.org/sites/default/files/2023-09/IOM-Irregular-Migration-SR-Protocol.pdf.

Bamberger, Michael. 2012. “Introduction to Mixed Methods in Impact Evaluation.” *Impact Evaluation Notes*, No. 3. www.interaction.org/wp-content/uploads/2019/03/Mixed-Methods-in-Impact-Evaluation-English.pdf.

Berretta, Miriam. 2022. “Mapping Evidence of What Works to Strengthen Resilience to Shocks and Stressors.” 2022. https://www.3ieimpact.org/blogs/mapping-evidence-what-works-strengthen-resilience-shocks-and-stressors.

Berretta, Miriam, Maria Daniela Anda Leon, Carolyn Huang, and Sanghwa Lee. 2023. “Addressing Root Causes and Drivers of Irregular Migration – an Evidence Gap Map.” New Delhi: International Initiative for Impact Evaluation (3ie).

Berretta, Miriam, Carolyn Huang, Maria Daniela Anda Leon, and Sanghwa Lee. 2023. “Protocol: Addressing Root Causes and Drivers of Irregular Migration – an Evidence Gap Map.” New Delhi: International Initiative for Impact Evaluation (3ie). https://3ieimpact.org/sites/default/files/2023-05/IOM-Irregular-Migration-EGM-Protocol.pdf.

Betcherman, Gordon, Karina Olivas, and Amit Dar. 2004. “‘Impacts of Active Labor Market Programs: New Evidence from Evaluations with Particular Attention to Developing and Transition Countries.” *Social Protection Discussion Paper Series*, 0402. Social Protection Unit.Human Development Network. The World Bank. https://documents1.worldbank.org/curated/en/426901468779104515/text/29142.txt

Blattman, Christopher, and Laura Ralston. 2015. “Generating Employment in Poor and Fragile States: Evidence from Labor Market and Entrepreneurship Programs.” *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2622220.

Borenstein, Michael, ed. 2009. *Introduction to Meta-Analysis*. Chichester, U.K: John Wiley & Sons.

Brown, Elizabeth, and Jeffery Tanner. 2019. “Integrating Value for Money and Impact Evaluations: Issues, Institutions, and Opportunities.” *Policy Research Working Paper* 9041. http://hdl.handle.net/10986/32586.

Card, David, Jochen Kluve, and Andrea Weber. 2015. “What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations.” *IZA Discussion Paper* 9236. https://docs.iza.org/dp9236.pdf

Carling, Jørgen. 2002. “Migration in the Age of Involuntary Immobility: Theoretical Reflections and Cape Verdean Experiences.” *Journal of Ethnic and Migration Studies* 28 (1): 5–42. https://doi.org/10.1080/13691830120103912.

Carling, Jørgen, and Kerilyn Schewel. 2018. “Revisiting Aspiration and Ability in International Migration.” *Journal of Ethnic and Migration Studies* 44 (6): 945–63. https://doi.org/10.1080/1369183X.2017.1384146.

Carling, Jørgen, and Cathrine Talleraas. 2016. “Root Causes and Drivers of Migration Implications for Humanitarian Efforts and Development Cooperation.” Peace Research Institute Oslo (PRIO). https://www.prio.org/publications/9229

Carranza, Eliana, and David McKenzie. Forthcoming. “Job Training and Job Search Assistance Policies in Developing Countries.”

Dalton, Jarrod E., Shari D. Bolen, and Edward J. Mascha. 2016. “Publication Bias: The Elephant in the Review.” *Anesthesia & Analgesia* 123 (4): 812–13. https://doi.org/10.1213/ANE.0000000000001596.

Dar, Amit, and Zafiris Tzannatos. 1999. “Active Labor Market Programs: A Review of the Evidence from Evaluations.” World Bank Social Protection Working Paper. The World Bank. https://documents1.worldbank.org/curated/en/246161468740363969/pdf/multi-page.pdf

ECORYS. 2020. “Addressing Root Causes (ARC) Programme Final Report.” Rotterdam: The Netherlands. https://www.kpsrl.org/publication/addressing-root-causes-arc-programme-final-report.

European Commission. Directorate General for International Partnerships. 2022. *EU Emergency Trust Fund for Africa: 2022 Annual Report.* LU: Publications Office. https://data.europa.eu/doi/10.2841/9748.

Faist, Thomas. 2008. “Migrants as Transnational Development Agents: An Inquiry into the Newest Round of the Migration-Development Nexus: Migrants as Transnational Development Agents.” *Population, Space and Place* 14 (1): 21–42. https://doi.org/10.1002/psp.471.

Gent, Saskia. 2002. “The Root Causes of Migration: Criticising the Approach and Finding a Way Forward.” Sussex Migration Working Paper 11. https://www.researchgate.net/publication/237735548_The_Root_Causes_of_Migration_Criticising_the_Approach_and_Finding_a_Way_Forward

Gertler, Paul J., Sebastian Martinez, Patrick Premand, Laura B. Rawlings, and Christel M. J. Vermeersch. 2016. *Impact Evaluation in Practice*. Second Edition. Washington, DC: Inter-American Development Bank and World Bank. https://www.worldbank.org/en/programs/sief-trust-fund/publication/impact-evaluation-in-practice

Ghosh, Bimal. 2006. *Migrants’ Remittances and Development: Myths, Rhetoric and Realities*. Geneva, Switzerland: International Organization for Migration. https://publications.iom.int/books/migrants-remittances-and-development-myths-rhetoric-and-realities

Glandon, Douglas, Sam Fishman, Caitlin Tulloch, Radhika Bhula, Grace Morgan, Shahana Hirji, and Liz Brown. 2023. “The State of Cost-Effectiveness Guidance: Ten Best Resources for CEA in Impact Evaluations.” *Journal of Development Effectiveness* 15 (1): 5–16. https://doi.org/10.1080/19439342.2022.2034916.

Goldin, Ian, Andrew Pitt, Benjamin Nabarro, and Kathleen Boyle. 2018. “Migration and the Economy Economic Realities, Social Impacts & Political Choices.” https://www.citivelocity.com/citigps/migration-and-the-economy/.

Heckman, James, Hidehiko Ichimura, and Petra Todd. 1997. “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program.” *Review of Economic Studies*, 1997. https://doi.org/10.2307/2971733

Heckman, James, Robert LaLonde, and Jeffrey Smith. 1999. “The Economics and Econometrics of Active Labor Market Programs.” In *Handbook of Labor Economics*. Vol. III. O. Ashenfelter and D. Card. https://doi.org/10.1016/S1573-4463(99)03012-6

Higgins, Julian P.T., Jonathan A. Sterne, Jelena Savovic, Mathew J. Page, Asbjørn Hróbjartsson, Isabelle Boutron, Barney Reeves, and Sandra Eldridge. 2016. “A Revised Tool for Assessing Risk of Bias in Randomized Trials.” *Cochrane Database of Systematic Reviews* 10 (Suppl 1): 29–31.

Higgins, Julian P.T., James Thomas, Jacqueline Chandler, Miranda Cumpston, Tianjing Li, Matthew J. Page, and Vivian A. Welch, eds. 2019. *Cochrane Handbook for Systematic Reviews of Interventions*. 1st ed. Wiley. https://doi.org/10.1002/9781119536604.

Hossain, Munshi Israil. 2022. “Impacts of Social Remittances on Economic Activities: Labour Migration from a Village of Bangladesh to Malaysia.” *Migration and Development* 11 (3): 273–90. https://doi.org/10.1080/21632324.2020.1753962.

ILO. 2022. *Global Estimates of Modern Slavery Forced Labour and Forced Marriage*. Geneva: International Labour Office. https://www.ilo.org/global/topics/forced-labour/publications/WCMS_854733/lang--en/index.htm

IMF. 2015. “International Migration: Recent Trends, Economic Impacts, and Policy Implications.” https://www.imf.org/external/np/g20/pdf/2015/111515background.pdf.

Kleemans, Marieke. 2015. “Migration Choice under Risk and Liquidity Constraints.” https://doi.org/10.22004/AG.ECON.200702.

Knoll, Anna, and Andrew Sheriff. 2017. “Making Waves: Implications of the Irregular Migration and Refugee Situation on Official Development Assistance Spending and Practices in Europe.” Expertgruppen för biståndsanalys (EBA). https://www.oecd.org/derec/sweden/201701-ECDPM-rapport.pdf.

Kugley, Shannon, Anne Wade, James Thomas, Quenby Mahood, Anne-Marie Klint Jørgensen, Karianne Hammerstrøm, and Nila Sathe. 2017. “Searching for Studies: A Guide to Information Retrieval for Campbell Systematic Reviews.” *Campbell Systematic Reviews* 13 (1): 1–73. https://doi.org/10.4073/cmg.2016.1.

Loschmann, Craig, Katie Kuschminder, and Melissa Siegel. 2014. “The Root Causes of Movement: Exploring the Determinants of Irregular Migration from Afghanistan.” Irregular Migration Research Programme Occasional Paper Series. Maastricht Graduate School of Governance (MGSoG) | UNU-MERIT. https://www.homeaffairs.gov.au/research-and-stats/files/irregular-migration-afghanistan.pdf.

Lwamba, Etienne, Shannon Shisler, Will Ridlehoover, Meital Kupfer, Nkululeko Tshabalala, Promise Nduku, Laurenz Langer, Sean Grant, Ada Sonnenfeld, Daniela Anda, John Eyers and Birte Snilstveit. 2021. “Strengthening Women’s Empowerment and Gender Equality in Fragile Contexts towards Peaceful and Inclusive Societies.” 2021st ed. International Initiative for Impact Evaluation (3ie). https://doi.org/10.23846/SR00047.

Massey, Douglas S., Joaquin Arango, Graeme Hugo, Ali Kouaouci, Adela Pellegrino, and Edward Taylor. 1993. “Theories of International Migration: A Review and Appraisal.” *Population and Development Review* 19 (3). https://www.jstor.org/stable/2938462.

McKenzie, David. 2017. “How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence.” *The World Bank Research Observer* 32 (2): 127–54. https://doi.org/10.1093/wbro/lkx001.

National Security Council, USG. 2021. “U.S. Strategy for Addressing the Root Causes of Migration in Central America.” National security council. https://www.whitehouse.gov/wp-content/uploads/2021/07/Root-Causes-Strategy.pdf.

Noyes, Jane, Andrew Booth, Graham Moore, Kate Flemming, Özge Tunçalp, and Elham Shakibazadeh. 2019. “Synthesising Quantitative and Qualitative Evidence to Inform Guidelines on Complex Interventions: Clarifying the Purposes, Designs and Outlining Some Methods.” *BMJ Global Health* 4 (Suppl 1): e000893. https://doi.org/10.1136/bmjgh-2018-000893.

Obokata, Reiko, Luisa Veronis, and Robert McLeman. 2014. “Empirical Research on International Environmental Migration: A Systematic Review,” 2014. https://pubmed.ncbi.nlm.nih.gov/25132701/.

Office of Management and Budget. 2022. “Budget of the U.S. Government Fiscal Year 2023.” U.S. Government Publishing Office. https://www.whitehouse.gov/wp-content/uploads/2022/03/budget_fy2023.pdf.

Pitoski, Dino, Thomas J. Lampoltshammer, and Peter Parycek. 2021. “Drivers of Human Migration: A Review of Scientific Evidence.” https://www.mdpi.com/2076-0760/10/1/21.

Rose, Sarah, Reva Resstack, Helen Dempster, Elisa Cascardi, and William Weinstein. 2021. “Addressing the ‘Root Causes’ of Irregular Migration from Central America: An Evidence Agenda for USAID.” 243. Immigration Policy Lab. Centre for Global Development. https://www.cgdev.org/publication/addressing-root-causes-irregular-migration-central-america-evidence-agenda-usaid.

Shemilt, Ian, David McDaid, Kevin Marsh, Catherine Henderson, Evelina Bertranou, Jacqueline Mallander, Mike Drummond, Miranda Mugford, and Luke Vale. 2013. “Issues in the Incorporation of Economic Perspectives and Evidence into Cochrane Reviews.” *Systematic Reviews* 2 (1): 83. https://doi.org/10.1186/2046-4053-2-83.

Snilstveit, Birte. 2012. “Systematic Reviews: From ‘Bare Bones’ Reviews to Policy Relevance.” *Journal of Development Effectiveness* 4 (3): 388–408. https://doi.org/10.1080/19439342.2012.709875.

Snilstveit, Birte, Sandy Oliver, and Martina Vojtkova. 2012. “Narrative Approaches to Systematic Review and Synthesis of Evidence for International Development Policy and Practice.” *Journal of Development Effectiveness* 4 (3): 409–29. https://doi.org/10.1080/19439342.2012.710641.

Sonnenfeld, Ada, Tomasz Kozakiewicz, Lina Khan, Jennifer Doherty, Katherine Garcia, John Eyers, Rami Zalfou and Douglas Glandon. 2023. “Rule of Law and Justice: An Evidence Gap Map.” International Initiative for Impact Evaluation (3ie). https://www.3ieimpact.org/evidence-hub/publications/evidence-gap-maps/rule-law-and-justice-evidence-gap-map.

Soto, Ariel G Ruiz, Rossella Bottone, Jaret Waters, Sarah Williams, Ashley Louie, and Yuehan Wang. 2021. “Charting a New Regional Course of Action: The Complex Motivations and Costs of Central American Migration.” *Migration Policy Institute*. https://www.migrationpolicy.org/research/motivations-costs-central-american-migration

Sterne, Jonathan A C, Jelena Savović, Matthew J Page, Roy G Elbers, Natalie S Blencowe, Isabelle Boutron, Christopher J Cates, et al. 2019. “RoB 2: A Revised Tool for Assessing Risk of Bias in Randomised Trials.” *BMJ*, August, l4898. https://doi.org/10.1136/bmj.l4898.

Thomas, James, and Angela Harden. 2008. “Methods for the Thematic Synthesis of Qualitative Research in Systematic Reviews.” *BMC Medical Research Methodology* 8 (1): 45. https://doi.org/10.1186/1471-2288-8-45.

UNHCR. 2022a. “Budget and Expenditure.” Global Focus. 2022a. http://reporting.unhcr.org/financial.

United Nations Office on Drugs and Crime (UNODC). 2021. “COVID 19 and the Smuggling of Migrants.” Vienna: United Nation Office on Drugs and Crime. https://www.unodc.org/documents/human-trafficking/SOM_and_COVID-19_Publication_final_EN_final.pdf.

UNSD. 2022. “Global Indicator Framework for the Sustainable Development Goals and Targets of the 2030 Agenda for Sustainable Development.” https://unstats.un.org/sdgs/indicators/Global%20Indicator%20Framework%20after%202022%20refinement_Eng.pdf.

Vaessen, Jos, Sebastian Lemire, and Barbara Befani. 2020. “Evaluation of International Development Interventions An Overview of Approaches and Methods.” Independent Evaluation Group. https://documents1.worldbank.org/curated/en/942511608652015232/pdf/Evaluation-of-International-Development-Interventions-An-Overview-of-Approaches-and-Methods.pdf.

Vutha, Hing, Lun Pide, and Phann Dalis. 2011. “Irregular Migration from Cambodia: Characteristics, Challenges, and Regulatory Approach,” PIDS Discussion Paper Series, , no. 2011–26. https://www.econstor.eu/handle/10419/126861.

Yayboke, Erol K., and Carmen Garcia Gallego. 2019. “Out of the Shadows Shining a Light on Irregular Migration.” CSIS Project on Prosperity and Development. https://reliefweb.int/report/world/out-shadows-shining-light-irregular-migration.

Other publications in the 3ie Systematic Review Series

The following reviews are available at
http://www.3ieimpact.org/evidence-hub/publications/systematic-reviews

Lwamba, E, Shisler, S, Ridlehoover, W, Kupfer, M, Tshabalala, N, Nduku, P, Langer, L, Grant, S, Sonnenfeld, A, Anda, D, Eyers, J and Snilstveit, B, 2021. *Strengthening women's empowerment and gender equality in fragile contexts towards peaceful and inclusive societies*. 3ie Systematic Review 47.

Strengthening intergroup social cohesion in fragile situations. 3ie Systematic Review 46. Sonnenfeld, A, Doherty, J, Berretta, M, Shisler, S, Snilstveit, B, Eyers, J, Castaman, K, Gupta, R, Anda Leon, MD, Franich, A, Yavuz, C, Baafi, A and Obaid, R, 2021.

Effects of electricity access interventions on socio-economic outcomes in low- and middle-income countries. 3ie Systematic Review 45. Moore, N, Glandon, D, Tripney, J, Kozakiewicz, T, Shisler, S, Eyers, J, Zalfou, R, Leon, MDA, Kurkjian, V, Snilstveit, B and Perdana, A. 2020.

Incentives for climate mitigation in the land use sector – the effects of payment for environmental services (PES) on environmental and socio-economic outcomes in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 44. Snilstveit, B, Stevenson, J, Langer, L, da Silva, N, Rabat, Z, Nduku, P, Polanin, J, Shemilt, I, Eyers, J, Ferraro, PJ, 2019.

Does incorporating participation and accountability improve development outcomes? Meta-analysis and framework synthesis. 3ie Systematic Review 43. Waddington, H, Sonnenfeld, A, Finetti, J, Gaarder, M and Stevenson, J, 2019.

Impact of financial inclusion in low- and middle-income countries: a systematic review of reviews. 3ie Systematic Review 42. Duvendack, M and Mader, P, 2019.

Agricultural input subsidies for improving productivity, farm income, consumer welfare and wider growth in low- and middle-income countries: a systematic review. 3ie Systematic Review 41. Hemming, DJ, Chirwa, EW, Ruffhead, HJ, Hill, R, Osborn, J, Langer, L, Harman, L, Coffey, C, Dorward, A and Phillips, D, 2018.

Vocational and business training to improve women's labour market outcomes in low- and middle-income countries: a systematic review. 3ie Systematic Review 40. Chinen, M, De Hoop, T, Balarin, M, Alcázar, L, Sennett, J, and Mezarina, J, 2018.

Interventions to improve the labour market for adults living with physical and/or sensory disabilities in low- and middle-income countries: a systematic review. 3ie Systematic Review 39. Tripney, J, Roulstone, A, Vigurs, C, Hogrebe, N, Schmidt, E and Stewart, R, 2017.

The effectiveness of contract farming in improving smallholder income and food security in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 38. Ton, G, Desiere, S, Vellema, W, Weituschat, S and D'Haese, M (2017)

Interventions to improve the labour market outcomes of youth: a systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions. 3ie Systematic Review 37. Kluve J, Puerto S, Robalino D, Romero JM, Rother F, Stöterau J,

Weidenkaff F and Witte M (2017)

Promoting handwashing and sanitation behaviour change in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 36. Buck, ED, Remoortel, HV, Hannes, K, Govender, T, Naidoo, S, Avau, B, Veegaete, AV, Musekiwa, A, Lutje, V, Cargo, M, Mosler, HJ, Vandekerckhove, P and Young T (2017)

Incorporating the life cycle approach into WASH policies and programmes: A systematic review. 3ie Systematic Review 35. Annamalai, TR, Narayanan, S, Devkar, G, Kumar, VS, Devaraj, R, Ayyangar, A and Mahalingam, A (2017)

Effects of certification schemes for agricultural production on socio-economic outcomes in low- and middle-income countries: a systematic review 34. Oya, C, Schaefer, F, Skalidou, D, McCosker, C and Langer, L (2017)

Short-term WASH interventions in emergency response: a systematic review. 3ie Systematic Review 33. Yates, T, Allen, J, Joseph, ML and Lantagne, D (2017)

Community monitoring interventions to curb corruption and increase access and quality of service delivery in low- and middle-income countries. 3ie Systematic Review 32. Molina E, Carella L, Pacheco A, Cruces, G and Gasparini, L (2016)

Effects and mechanisms of market-based reforms on access to electricity in developing countries: a systematic review. 3ie Systematic Review 31. Bensch, G, Sievert, M, Langbein, J, Kneppel, N (2016)

Youth gang violence and preventative measures in low- and middle-income countries: a systematic review (Part II), 3ie Systematic Review 30. Higginson, A, Benier, K, Shenderovich, Y, Bedford, L, Mazerolle, L, Murray, J (2016)

Youth gang membership and violence in low- and middle-income countries: a systematic review (Part I), 3ie Systematic Review 29. Higginson, A, Benier, K, Shenderovich, Y, Bedford, L, Mazerolle, L, Murray, J (2016)

Cash-based approaches in humanitarian emergencies: a systematic review, 3ie Systematic Review Report 28. Doocy, S and Tappis, H (2016)

Factors affecting uptake of voluntary and community-based health insurance schemes in low- and middle-income countries: a systematic review, 3ie Systematic Review 27. Panda, P, Dror, IH, Koehlmoos, TP, Hossain, SAS, John, D, Khan, JAM and Dror, DM (2016)

Parental, community and familial support interventions to improve children's literacy in developing countries: a systematic review, 3ie Systematic Review 26. Spier, E, Britto, P, Pigott, T, Roehlkapartain, E, McCarthy, M, Kidron, Y, Song, M, Scales, P, Wagner, D, Lane, J and Glover, J (2016)

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

Interventions for improving learning outcomes and access to education in low- and middle-income countries: a systematic review, 3ie Systematic Review 24. Snilstveit, B, Stevenson, J, Phillips, D, Vojtkova, M, Gallagher, E, Schmidt, T, Jobse, H, Geelen, M, Pastorello, M, and Eyers, J (2015)

Economic self-help group programmes for improving women's empowerment: a systematic review, 3ie Systematic Review 23. Brody, C, De Hoop, T, Vojtkova, M, Warnock, R, Dunbar, M, Murthy, P and Dworkin, SL (2016)

The identification and measurement of health-related spillovers in impact evaluations: a systematic review, *3ie Systematic Review 22*. Benjamin-Chung, J, Abedin, J, Berger, D, Clark, A, Falcao, L, Jimenez, V, Konagaya, E, Tran, D, Arnold, B, Hubbard, A, Luby, S, Miguel, E and Colford, J (2015)

The effects of school-based decision-making on educational outcomes in low- and middle-income countries: a systematic review, 3ie Systematic Review Report 21. Carr-Hill, R, Rolleston, C, Pherali, T and Schendel, R, with Peart, E, and Jones, E (2015)

Policing interventions for targeting interpersonal violence in developing countries: a systematic review, *3ie Systematic Review 20*. Higginson, A, Mazerolle, L, Sydes, M, Davis, J, and Mengersen, K (2015)

The effects of training, innovation and new technology on African smallholder farmers' wealth and food security: a systematic review, *3ie Systematic Review 19*. Stewart, R, Langer, L, Rebelo Da Silva N, Muchiri, E, Zaranyika, H, Erasmus, Y, Randall, N, Rafferty, S, Korth, M, Madinga, N and de Wet, T (2015)

Community based rehabilitation for people with disabilities in low- and middle-income countries: a systematic review, *3ie Systematic Review 18*. Iemmi, V, Gibson, L, Blanchet, K, Kumar, KS, Rath, S, Hartley, S, Murthy, GVS, Patel, V, Weber, J and Kuper H (2015)

Payment for environmental services for reducing deforestation and poverty in low- and middle-income countries: a systematic review, *3ie Systematic Review 17.* Samii, C, Lisiecki, M, Kulkarni, P, Paler, L and Chavis, L (2015)

Decentralised forest management for reducing deforestation and poverty in low- and middle-income countries: a systematic review, *3ie Systematic Review 16.* Samii, C, Lisiecki, M, Kulkarni, P, Paler, L and Chavis, L (2015)

Supplementary feeding for improving the health of disadvantaged infants and young children: a systematic and realist review, *3ie Systematic Review 15.* Kristjansson, E, Francis, D, Liberato, S, Greenhalgh, T, Welch, V, Jandu, MB, Batal, M, Rader, T, Noonan, E, Janzen, L, Shea, B, Wells, GA and Petticrew, M (2015)

The impact of land property rights interventions on investment and agricultural productivity in developing countries: a systematic review, 3ie Systematic Review Report 14. Lawry, S, Samii, C, Hall, R, Leopold, A, Hornby, D and Mtero, F, 2014.

Slum upgrading strategies and their effects on health and socio-economic outcomes: a systematic review, 3ie Systematic Review 13. Turley, R, Saith, R., Bhan, N, Rehfuess, E, and Carter, B (2013)

Services for street-connected children and young people in low- and middle-income countries: a thematic synthesis, 3ie Systematic Review 12. Coren, E, Hossain, R, Ramsbotham, K, Martin, AJ and Pardo, JP (2014)

Why targeting matters: examining the relationship between selection, participation and outcomes in farmer field school programmes, 3ie Systematic Review 11. Phillips, D, Waddington, H and White, H (2015)

The impact of export processing zones on employment, wages and labour conditions in developing countries, 3ie Systematic Review 10. Cirera, X and Lakshman, R (2014)

Interventions to reduce the prevalence of female genital mutilation/cutting in African countries, 3ie Systematic Review 9. Berg, RC and Denision, E (2013)

Behaviour change interventions to prevent HIV among women living in low and middle income countries, 3ie Systematic Review 8. McCoy, S, Kangwende, RA and Padian, NS (2009)

The impact of daycare programs on child health, nutrition and development in developing countries, 3ie Systematic Review 7. Leroy, JL, Gadsden, P and Guijarro, M (2011)

Willingness to pay for cleaner water in less developed countries: Systematic review of experimental evidence, 3ie Systematic Review 6. Null, C, Hombrados, JG, Kremer, M, Meeks, R, Miguel, E and Zwane, AP (2012)

Community-based intervention packages for reducing maternal morbidity and mortality and improving neonatal outcomes, 3ie Systematic Review 5. Lassi, ZS, Haider, BA and Langou, GD (2011)

The effects of microcredit on women's control over household spending: a systematic review, 3ie Systematic Review 4. Vaessen, J, Rivas, A, Duvendack, M, Jones, RP, Leeuw, F, van Gils, G, Lukach, R, Holvoet, N, Bastiaensen, J, Hombrados, JG and Waddington, H, (2013).

Interventions in developing nations for improving primary and secondary school enrolment of children: a systematic review, 3ie Systematic Review 3. Petrosino, A, Morgan, C, Fronius, T, Tanner-Smith, E, and Boruch, R, 2016.

Interventions to promote social cohesion in Sub-Saharan Africa, 3ie Systematic Review 2. King, E, Samii, C and Snilstveit, B (2010)

Water, sanitation and hygiene interventions to combat childhood diarrhoea in developing countries, 3ie Systematic Review 1. Waddington, H, Snilstveit, B, White, H and Fewtrell, L (2009)

International Initiative for Impact Evaluation
1 Poultry London
EC2R 8EJ
United Kingdom

3ieuk@3ieimpact.org
Tel: +44 20 3695 7895

www.3ieimpact.org

USAID
FROM THE AMERICAN PEOPLE

International Initiative for Impact Evaluation

IOM
UN MIGRATION